Structure-preserving model order reduction for drilling automation

Harshit Bansal

Structure-preserving model order reduction for drilling automation

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op dinsdag 13 oktober 2020 om 16:00 uur

door

Harshit Bansal

geboren te Agra, India

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter: prof.dr. J. Lukkien
1e promotor: prof.dr. W.H.A. Schilders
2e promotor: prof.dr.ir. N. van de Wouw
co-promotor: dr. L. Iapichino
leden: prof.dr. M. Ohlberger (University of Münster)
                prof.dr.ir. J.M.A. Scherpen (Rijksuniversiteit Groningen)
                prof.dr. S. Weiland
                prof.dr. F. di Meglio (MINES ParisTech)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background, Motivation and Research Objectives

The world witnessed a disastrous event with the Deepwater Horizon drilling rig explosion in 2010 (see Figure 1.1), which had large-scale repercussions on the environment and the economy. There was a chain of events which led to the disaster. Primarily, the incident occurred due to some errors in pressure management of the drilling system under gas influx from the reservoir, and this led to detrimental effects. The natural question that arises is whether such an unwanted incident could have been avoided. The answer to this question is affirmative. Certainly, such incident could be avoided through a better down-hole pressure control, automation and real-time drilling decision making.

Furthermore, future sustainable harvesting of geothermal energy requires exploitation of difficult-to-access, unconventional reserves. The drilling of deep wells for these purposes is characterized by high complexity, high uncertainty, high risk and high cost. As an example, data from drilling operations on the Norwegian Continental Shelf (NCS) from the last 10 years show that well construction costs have increased, while drilling efficiency has reduced [137]. This is a general trend that applies throughout the drilling industry world-wide. The cause for this undesirable trend is complex, but an important factor is the lack of automation of the drilling operations. The aforementioned discussion clearly demonstrates the lack of automation and proper down-hole pressure management.

Figure 1.1: Deepwater Horizon in flames after the explosion [Ref: Wikipedia].

Development of technology, such as Managed Pressure Drilling (MPD), is gaining relevance for offering improved automatic pressure control. MPD offers a better down-hole pressure control and aids in mitigating gas influx incidents. In view of virtual drilling scenario testing, and for simulation and control purposes, the foundation of drilling automation strategies for automating pressure management needs to be based upon a hydraulic model.

To this date, no hydraulic models exist that are both *(i)* accurate enough and *(ii)* simple enough to be employed in the context of real-time estimation and control in case of gas influx. On top of that, the sensor measurements are either not available or have an embedded delay during various drilling operations. The effectiveness of an automated system or a control system may be jeopardized under the above mentioned circumstances. Moreover, a large amount of data needs to be transferred from the down-hole to the surface in a reasonable amount of time. Wireless data transfer techniques, such as mud pulse, are generally employed for this purpose. However, such techniques require data to be stored in memory and downloaded during a certain drilling phase. This option, in which data is stored and then evaluated after some drilling operation, makes real-time drilling decisions nearly impossible if there does not exist an accurate

hydraulic model to predict a characteristic behaviour/response. To summarize, sensor measurements along with the hydraulic model need to be utilized to make real-time decisions while performing drilling operations. The fundamental prerequisite is hydraulic model development which, in the future, will serve as the basis for controller and estimator design. Currently, hydraulic models are simple, but not accurate enough as the distributed nature of the drilling systems is usually (totally or partially) neglected [3, 91, 133]. Ignoring these distributed effects can be detrimental to an automated MPD system. Furthermore, there exist hydraulic models that are highly accurate [14, 57–61, 67, 128, 167], but are not employed in the scope of drilling operations since they are too complex for virtual drilling scenario testing and controller design.

The main motivation of this research is to develop a much needed novel computational framework for (multi-phase) hydraulic modelling and model complexity reduction for drilling operations, thus supporting the wide-spread deployment of automation strategies for down-hole pressure management.

This dissertation is a stepping stone towards:

- Developing high-fidelity two-phase hydraulic models for drilling operations that accurately reproduce the dynamics of gas-liquid flow in the well and enable accurate prediction of down-hole pressure;

- Developing model order reduction techniques for the construction of fit-for-purpose hydraulic models that enable effective handling of distributed non-linearities and delays (due to pressure wave propagation), guarantee the preservation of key system properties (such as stability, multiple time- and length-scales and input-output behaviour), and preserve the dependency on key physical parameters in the reduced-order model.

In the scope of the aforementioned high-level objectives, this dissertation will discuss novel results in the field of modelling, structure-preserving discretization and model order reduction framework to deal with (large-scale) non-linear dynamical systems as in MPD.

1.2 Scientific challenges and research goals of this dissertation

In this section, we will first discuss scientific research challenges and then formulate research objectives that have been treated in this dissertation.

1.2.1   Scientific challenges

The scientific challenges associated with the aforementioned objectives can be identified in two broad themes.

- The first theme revolves around how to develop models that include essential features of the multi-phase flow dynamics in drilling, such as transition between one-phase (gas or liquid) flow and multi-phase flow, propagation of fast pressure waves and slow mass transport phenomena across a (variable) geometrical cross-section, gas expansion and acceleration, reservoir-well interaction, and accurate modelling of the friction losses; see Chapter 2 for physical insights.

- The second theme revolves around how to develop methods for model complexity reduction that, firstly, handle distributed non-linearities, delays and guarantee the preservation of key system properties and, secondly, can be used to construct low-complexity models suitable for both extensive, reliable drilling scenario testing and the design of automation strategies for down-hole pressure management.

First theme - Modelling and (structure-preserving) discretization of hydraulic systems
Multi-phase hydraulic models, which are multi-scale and evolutionary in nature, are popularly governed by non-linear partial differential equations. Such hydraulic models pose several theoretical and computational challenges. The physics governing the hydraulic dynamics in the wellbore involves several phenomena across different time scales. The slow time scale is related to the mass transport and the fast time scale is related to the propagation of the acoustic (or pressure) waves. Capturing these wave propagation phenomena, induced by slow or fast transients, is a central challenge in hydraulic modelling. The wave structure of (multi-phase) hydraulic models is composed of shock waves, rarefaction waves, and, stationary and moving contact discontinuities. Great progress has been made in the last few decades to implement a stable, robust and accurate numerical scheme to capture such discontinuous physical features [9–11, 37, 47, 60, 68, 92, 132, 181, 197]. The development of robust and accurate numerical methods for non-linear and compressible multi-phase flow problems, characterized by conditional hyperbolicity, remains a great challenge even after several decades of successful research.

Several aspects need to be taken into account when solving these equations numerically, such as: *(i)* efficiently capturing the relevant phenomena over the

various spatio-temporal scales, *(ii)* guaranteeing the monotonicity of a numerical scheme, positivity of a physical quantity, and the preservation of steady-state behaviour, conserved physical quantities and key system properties (such as *stability*[1] and input-output behaviour), and *(iii)* enabling effective handling of distributed non-linearities and non-conservative terms in the governing partial differential equation.

There exist several variants of multi-phase hydraulic models. We will only focus on two variants, namely, the Two-Fluid Model (TFM) [56] and the Drift Flux Model (DFM) [57, 90], in the scope of modelling for MPD systems. As explained in Chapter 2, the main interest lies in the modelling of two-phase flow dynamics which is of practical relevance under gas-influx. The TFM and the DFM provide a physically realistic description for modelling the two-phase flow behaviour of interest. A framework already exists for high-fidelity computations of shocks, expansion waves and contact discontinuities for both the TFM and the DFM [57–61, 67, 128, 167]. The aspect of preserving the steady state of two-phase flow models has also been addressed via *well-balanced* numerical schemes in [174]. Furthermore, the positivity of a physical quantity is not trivially guaranteed upon employing any numerical scheme. However, there exist a few selected numerical schemes that retain the positivity of a physical quantity [57, 59]. There also exists a rich literature to deal with non-conservative terms, particularly in the scope of the TFM [58, 167]. Such non-conservative terms also arise for modelling fluid flow behaviour across a variable geometrical cross-section, and have been dealt recently; see [7] in the scope of the DFM.

While some numerical schemes are difficult to derive and expensive to use [67], few other numerical schemes are simple to implement, but at the cost of a desirable numerical property [128]. On the one hand, some numerical schemes are able to capture fast-propagation effects, but strongly smear slow-propagating contact discontinuities [57]. On the other hand, other numerical schemes fail to effectively resolve the fast-propagating effects, but capture the

---

[1]It is important to distinguish between different kinds of (in)stability. By physical (in)stability, we refer to the intrinsic behaviour of the system that is simulated. While numerical (in)stability is a consequence of a numerical technique/algorithm that is used to simulate the system. A physically stable system can be numerically unstable. While there exist methodologies, such as Von Neumann stability analysis [85], total variation analysis [85], etc., for linear partial differential equations, a consistent definition of numerical stability is complicated for non-linear partial differential equations. Numerical algorithms for non-linear partial differential equations are required to possess numerical stability in the sense that the CFL criteria is not violated and that the numerical algorithm does not amplify errors [85]. Preservation of stability implies that the numerical stable regime is (almost) equal to the physical stable regime.

contact discontinuities [57]. Numerical schemes, such as AUSMV, have been developed in the past to resolve both fast moving wavefronts and slow moving contact discontinuities in an accurate manner with no or minimal smearing. While numerical schemes may satisfy Abgrall's principle[2], they may not possess the property of "Total Variation Diminishing". For instance, AUSMV and AUSMD satisfy Abgrall's principle, but lose monotonicity or the "Total Variation Diminishing" property [58]. Furthermore, several conserved physical quantities such as mass, momentum and energy are not preserved exactly in a discrete sense. To summarize, each numerical scheme comes with its own merits and demerits, and there is still a need to develop a suitable numerical tool to describe two-phase transport problems.

To add, usually no analytical results exist for either the TFM and the DFM, and a certain numerical scheme at an extremely fine spatial and temporal resolution is used to serve as a benchmark for a "truth" simulation response. Furthermore, under grid refinement, numerical solutions computed from various numerical schemes are theoretically expected to converge to the same solution. However, different numerical schemes yield different physical representations at a finer grid resolution (see Chapter 6) and, hence, a challenge is faced to arrive at a consensus regarding the best numerical scheme to simulate the TFM and the DFM. Moreover, for instance, the authors in [57] consider the solution generated by AUSMV (on a finer grid) as a reference solution of the DFM. However, as discussed earlier, AUSMV is not "Total Variation Diminishing" by nature. Furthermore, there is no sound physical ground to discard a sharp spike just behind a moving front generated by other numerical schemes, such as FVS [57]. Also, a surrogate of the sound speed model for the two-phase mixture (gas and liquid) is generally used; see [2, 57] in the scope of the DFM. Such a surrogate model assumes one of the phases to be incompressible, which is not physically realistic in the realm of drilling applications. The (mixture) sound speed model is known to play a crucial role in developing numerical schemes, and has significant implications in the location and speed of the moving wavefront, numerical dissipation, etc. Hence, the (mixture) sound speed model needs to be accurate and deserves careful attention. There is clearly a scope of further detailed numerical analysis, and the following natural questions arise:

Q1. How to rigorously certify a numerical scheme and establish the relative order of the merit of the numerical schemes?

---

[2]A flow, uniform in pressure and velocity, must remain uniform in the same variables during its time evolution.

Q2. Do exact/analytical solutions exist or can be developed for the two-phase flow models of interest in order to test the model and the numerical method?
Q3. How important is the role of the sound speed of a two-phase mixture?
Q4. How to preserve conserved physical quantities in a discrete setting?

In general, Finite-Volume (FV) Methods are employed for developing a high-fidelity numerical hydraulic model. These methods are popular as they are inherently conservative and can easily be formulated to deal with unstructured mesh of an underlying spatial domain. It is important to ensure the *numerical stability* (an aspect related to Q1. posed above) of the FV-based high-fidelity simulation. However, the *numerical stability* of such a (high-fidelity) complex model is not rigorously certified in the existing literature [57–60, 67, 128, 167]. Moreover, the results about the *physical (in)stability* of (multi-phase) hydraulic models do not exist. Most of the literature related to multi-phase hydraulic models, such as in the DFM, to the best of our knowledge, involve performing 'successful' numerical experiments within a certain restricted problem setting [57, 60]. These methods could generate (bounded) stable solutions until a certain time t and then run into numerical issues. Alternatively, non-stable (non-physical) solutions could be observed in the first few time-steps of the numerical simulation which then converges to physically meaningful results over a finite time-horizon. This can be caused by the used numerical scheme giving poor performance (by construction) for some combinations of parameters/operating conditions. Alternatively, the observed *numerical instability* could have arisen due to inherent properties of the physical model as explained next. The underlying multi-phase models for MPD systems are conditionally hyperbolic and, consequently, the eigenvalues of associated systems may become complex. Various potential issues, such as correct specification of boundary data, computation of stable numerical solutions, correct determination of propagation of wavefront, etc., arise if the eigenvalues become complex. Given the fact that we work with a model governed by non-linear partial differential equations, the model can theoretically/physically admit both stable and/or unstable solutions for a certain set of operating conditions. Moreover, it is not clear whether the hyperbolicity should be regarded as a necessary condition for well-posedness of the model. In other words, the physical validity of the model outside the hyperbolic regime remains an open question. Some works [51, 149, 157, 166, 170] modify few terms or add extra equations in the model of interest to ensure the hyperbolicity of the model across all or vast range of operating conditions. On the contrary, in principle, the numerical implementation for (conditionally) hyperbolic multi-phase flow model could encompass a switching behaviour to

enforce suitable numerical scheme as per the type of the partial differential equation for a certain state of the system. To summarize, numerical simulations work 'successfully' within a restrictive setting without concrete (theoretical) understanding which is aggravated by the lack of exact/analytical solutions.

To this end, we propose to adopt a port-Hamiltonian modelling framework with the vision to develop a map quantifying range of initial conditions, boundary conditions, several involved parameters, etc., for which the high-fidelity solutions can be guaranteed to be *physically stable*. A port-Hamiltonian framework provides a sound theoretical underpinning ultimately helping to decouple the physical effects and the numerical effects. Furthermore, port-Hamiltonian formulations have several key properties that include passivity, shifted passivity, existence of Casimirs and compositionality. Moreover, such a framework ensures certain dissipativity properties which are strongly linked to passivity of the system [185]. Also, stability can often be deduced from a dissipation inequality if the Hamiltonian satisfies some further conditions which make it a Lyapunov function [94]. Furthermore, the system-theoretic properties of the infinite-dimensional realization can be preserved in an equivalent finite-dimensional realization via structure-preserving discretization and reduction methods. A port-Hamiltonian framework can, hence, definitely be of great use for developing an automated (and stable) model order reduction framework for two-phase flow models of interest. A natural question that arises now is:

Q5. How to develop a port-Hamiltonian modelling and structure-preserving discretization (and reduction) framework for the MPD systems?

Structure-preserving discretization methods are known to play a crucial role in preserving the structural invariants (Casimirs) of an infinite-dimensional system at the finite-dimensional level. Hence, such a framework can also be an answer to Q4. posed above.

We have discussed the state-of-the-art, scientific challenges and open research questions in the scope of first theme dedicated to (high-fidelity) hydraulics modelling. We now turn our attention to delve into the second theme that hinges around model complexity reduction.

Second theme - Model order reduction for transport-dominated problems

Model Order Reduction (MOR) is an indispensable tool to reduce the complexity of a model. It enables multi-query simulations, (faster than) real-time estimation and control, and optimization of operations in real scenarios, in par-

ticular for multi-scale and multi-physics problems. The focus of this dissertation in the scope of model complexity reduction lies in the class of unsteady (parametrized) hyperbolic partial differential equations (for, e.g., multi-phase hydraulic models). Such mathematical models are characterized by solutions possessing distinct behaviours, such as different wave speeds, and locations of moving and continuously evolving discontinuous front(s). Standard MOR approaches, such as, the Reduced Basis Method (RBM) [81], the Proper Orthogonal Decomposition (POD) [25, 195], etc., in conjunction with the Discrete Empirical Interpolation Method (DEIM) [22, 44], rely on using spatially fixed modes as a basis for reduction. These methods fail to strike an acceptable balance between model fidelity and complexity to obtain computationally compact representations of several phenomena of interest, such as, for instance, multiple-transport phenomena with sharp gradient features.

Such an inefficiency of standard MOR approaches holds both for parametric and non-parametric (evolutionary) transport-dominated problems (for, e.g., problems with wave-propagation effects as in multi-phase hydraulic model) [77, 135]. The underlying challenge is briefly explained next. Suppose, we have an initial configuration of a certain shape profile. Such initial configuration can be reproduced (up to a required level of representation accuracy) using a combination of certain basis functions. Given the travelling nature of the problem, the initial configuration profile will translate to a different spatial location as time advances. Reproducing a new shifted profile demands for constructing a new set of basis functions. Basically, a new set of approximating functions needs to be constructed at every time instant. Hence, a large number of basis functions are required to capture even a single moving wavefront. In the scope of the TFM and the DFM, multiple wavefronts propagate in different directions with different strengths and, hence, the expected reduction via standard MOR approaches would be challenged even more. Moreover, the shape of the solution undergoes topological changes under wavefront interactions and then continues to evolve over the spatial domain. The aforementioned issues are further compounded by the presence of sharp gradient features, for instance, shocks. Such sharp jumps make the reduced-order model representation(s) susceptible to oscillatory effects, in the vicinity of sharp gradient in the solution, which can escalate into a *numerical instability* in finite time. The limitations of standard MOR techniques instigate a drive to develop novel and automated techniques tailored for preservation of key (physical) model properties, such as, positivity, evolving discontinuous fronts, etc., and system-theoretic properties (for instance, port-Hamiltonian structure, propagation delays, stability, etc.) of associated reduced-order models, while still guaranteeing to offer robust and

stable approximation with lowest possible reduced-order representations.

Furthermore, while reduced-order model realizations built on top of port-Hamiltonian framework (via structure-preserving discretization and model reduction methods) can be inherently stable/passive, the resulting reduced realizations are not necessarily of lowest possible dimension. Hence, in order to obtain an accurate, stable and lowest possible reduced model representation, in particular for problems with moving sharp gradient features, a novel ingredient is sought to be developed which in conjunction with the port-Hamiltonian framework will take us a step further to realize the envisioned dream.

While standard reduction approaches do not exploit the underlying hyperbolic structure/physics/dynamics of the problem, we strive to answer the following questions:

Q6. How do the standard MOR approaches fare in the context of two-phase flow models of interest?
Q7. How to exploit the underlying structure of the system dynamics and how to construct a model order reduction framework in order to reduce the dimensionality (compared to standard reduction) of problems admitting (discontinuous) wavefront interactions?

Having identified the open scientific challenges, we now formulate the research objectives addressed in this dissertation.

1.2.2 Research Goals

Though we earlier identified two broad research themes, it is clear that both themes are closely coupled to each other. We pursue both themes in this dissertation and address the open challenges given by the questions Q1 - Q9 identified above. The research goals pertaining to the challenges in the first theme and addressed in **Part I** of this dissertation are devoted to **Modelling and (structure-preserving) discretization of hydraulic systems**, and are summarized below.

O1 (related to Q5.) To develop port-Hamiltonian formulations for single- and two-phase flow models across a constant, and a spatially and temporally varying geometrical cross-section of the well;

O2 (related to Q5.) To identify whether a two-phase flow model, in particular the Drift Flux Model with the Zuber-Findlay slip law, is an energy-consistent model;

O3 (related to Q5.) To construct a compositional port-Hamiltonian model for MPD systems (under influx from the reservoir) by interconnecting specific drilling components (drill string, bit and the annulus), and accounting for energy exchange at the reservoir boundary;

O4 (related to Q4. and Q5.) To develop a structure-preserving discretization framework for the above obtained infinite-dimensional port-Hamiltonian realizations;

O5 (related to Q2.) To test the model and numerical method with exact reference solutions which are available only for a specific equation of the state for the phases under consideration;

O6 (related to Q1.) To estimate numerical discretization error via an error transport approach and use it along with other performance metrics on several challenging test cases (for, e.g., fast-pressure transients, slow mass transport phenomena, appearance and disappearance of phases, etc.) to establish the relative order of merit of numerical schemes;

O7 (related to Q3.) To study the sensitivity (in a numerical sense) of the sound speed model of a two-phase flow mixture;

The goals pertaining to the second theme, addressed in **Part II** of this dissertation, are devoted to **Model order reduction for transport-dominated problems**, and are presented next.

O8 (related to Q6.) To assess the performance of standard model order reduction techniques in reducing the (non-linear) two-phase flow models of interest along with simpler but challenging non-linear problems such as the Burgers' equation;

O9 (related to Q7.) To construct a (pre-processing) framework to deal with interacting and merging fronts (shocks);

O10 (related to Q7.) To reduce the dimensionality (compared to standard reduction) of problems admitting (discontinuous) wavefront interactions.

It is worth stating that the newly developed model order reduction framework is in a phase of theoretical development and is quite far from being applied to MPD systems. Though we restrict ourselves to simpler academic examples for assessing the numerical performance of the proposed approach, we do embody

the critical features observed in the scope of the MPD models in order to design simpler yet challenging simulation studies.

The main contributions along with the structure of this dissertation are outlined next.

1.3 Outline of the dissertation & Main contributions

The dissertation consists of an introductory chapter (Chapter 2) on MPD systems and two main parts. Chapter 2 discusses the need of an MPD system along with its basic configuration and, furthermore, highlights the relevant properties that need to captured by (multi-phase) hydraulic systems for MPD operations. The first part (Chapters 3-6) of this dissertation is dedicated to the contribution on the *modelling and (structure-preserving) discretization* of hydraulic models in the scope of MPD. In particular, we strive to address questions Q1 - Q5 by fulfilling outlined research objectives O1 - O7. The second part (Chapters 7-9) of this dissertation deals, firstly, with the application of standard *model order reduction* techniques on multi-phase hydraulic models and, secondly, discusses novel, efficient and automated *model order reduction* techniques, particularly in the scope of *transport-dominated problems*. In particular, we strive to address questions Q6 - Q7 by fulfilling outlined research objectives O8 - O10. A brief overview of the topics covered in each chapter, and the related key contributions, of this dissertation is provided below.

Introduction to Managed Pressure Drilling set-up

Chapter 2
In this chapter, we discuss the issue of pressure control during conventional drilling operations. The aspect of pressure control and its criticality is extensively discussed by using a specific drilling well: an Extended Reach Drilling well. We then review a basic configuration of the MPD system and highlight operational problems during drilling. In order to create a virtual environment for drilling scenario testing and to automate down-hole pressure management, a hydraulic model is required. Some requirements need to be imposed on the hydraulic model in order to mimic the scenarios realistically. These requirements are classified into two sub-classes of model requirements, namely, hydraulic model requirements, and controller and estimator requirements. These require-

ments are thereby described to outline the features of interest that need to be resolved in the high- and low-complexity models. Also, the mathematical formulation of the (multi-phase) hydraulic model(s) is introduced.

Part I: Modelling and (structure-preserving) discretization of hydraulic systems (Chapters 3-6)

Chapter 3
In this chapter, we focus on developing port-Hamiltonian formulations for modelling two-phase flow dynamics across a unit and constant cross-section. In particular, we develop dissipative Hamiltonian representations for the Two-Fluid Model and the Drift Flux Model, and propose corresponding novel and *state-dependent Stokes-Dirac structures* in order to obtain port-Hamiltonian formulations. This modelling effort is a stepping stone towards the simulation and control of two-phase flow models governed by conservation laws. This work is also significant as the existing theory in the scope of *linear distributed-parameter port-Hamiltonian systems* has been exploited to arrive at new results from an operator theoretic viewpoint, including further generalizations in the scope of *non-linear distributed-parameter port-Hamiltonian systems*.

Chapter 4
In this chapter, we extend the principles developed in Chapter 3 for modelling single- and two-phase flow dynamics across a spatially (and temporally) varying cross-section. These fluid dynamical systems are ultimately represented in port-Hamiltonian formulations with respect to newly proposed *state-dependent and extended Stokes- Dirac structures*. Moreover, we exploit the property of compositionality and compose a network of port-Hamiltonian systems for an MPD set-up via a power-preserving interconnection.

Chapter 5
In this chapter, we develop a framework for discretizing infinite-dimensional non-linear port-Hamiltonian representations in a structure-preserving manner. We employ the idea of mixed-finite-element methods for spatial discretization and the concept of Gauss-Legendre collocation methods for temporal discretization of port-Hamiltonian representations obtained in the scope of fluid dynamical systems studied in Chapter 3 and Chapter 4. The properties of the resulting (continuous-time and discrete-time) finite-dimensional realizations are assessed, and the conditions under which these are known to preserve the properties of a finite-dimensional Dirac structure are discussed. Moreover, we derive

the complete finite-dimensional interconnected port-Hamiltonian model by invoking the notion of power-preserving interconnection. Finally, we test the proposed discretization framework on a simple, yet representative test case.

Chapter 6
In this chapter, we account for compressibility effects of the phases and obtain a model for the speed of sound of a two-phase mixture. We also numerically study the importance of the role of the speed of sound of a two-phase mixture. Furthermore, the estimation of the discretization error is pursued in the scope of the Drift Flux Model. This numerical error is estimated by utilizing the concept of numerical error estimation for non-linear hyperbolic partial differential equations via non-linear error transport. Furthermore, the relative order of merit of finite-volume based numerical schemes is established by performing several challenging case studies.

Part II: Model order reduction for transport-dominated problems (Chapters 7-9)

Chapter 7
In this chapter, we lay the mathematical foundation for model order reduction by revisiting existing basic principles for model reduction in the scope of linear and non-linear dynamical systems. Standard (Galerkin-type) projection-based approach is then applied on the (non-linear) Burgers' equation and the Drift Flux Model. The numerical performance is assessed on some challenging test cases. These numerical experiments put forward the success and limitations of standard model order reduction techniques.

Chapters 8 & 9

Chapters 8 & 9 deal with the development of the model order reduction framework in the scope of *transport-dominated problems*, in particular for problems in which moving wavefronts hamper the efficacy of existing reduced-basis approaches.

In Chapter 8, we investigate the combined approach of the Method of Freezing/Symmetry Reduction and reduced basis approximations in dealing with merging (discontinuous) wavefronts via several numerical case studies. Furthermore, we propose a novel flux re-distribution approach in order to curb the additional travelling structures or numerical instabilities (induced by such

wavefront interactions) encountered in the Method of Freezing. Finally, we present numerical case studies and discuss the performance of our proposed approach (in conjunction with reduced basis approximations) in terms of computational accuracy compared with standard MOR techniques.

In Chapter 9, we propose a new MOR approach to obtain effective reduction for transport-dominated problems or hyperbolic partial differential equations. The main ingredient is a novel decomposition of the solution into a function that tracks the evolving discontinuity and a residual part that is devoid of shock features. This decomposition ansatz is then combined with Proper Orthogonal Decomposition applied only to the residual part to develop a reduced-order model representation for problems with multiple moving and possibly merging discontinuous features. We test our framework on challenging scenarios that embody the features of critical interest in the scope of multi-phase hydraulic models. Numerical case-studies show the potential of the approach in terms of computational accuracy compared with standard MOR techniques.

Conclusions and Future Works

Chapter 10
This is the concluding chapter of the dissertation. We briefly reflect on the theoretical and practical contributions. In this chapter, we also discuss limitations, and open-ended questions, and consequently provide insights on research directions that can be perceived as natural extensions of this work.

1.4   Scientific Publications

Some part of this dissertation is based on the results that have been published, accepted for publication or submitted in several peer-reviewed journals and conference proceedings. Another part of this dissertation is yet to be converted into a few scientific papers.

- Chapter 3 is an elaboration of the contents in the article: "Port-Hamiltonian Formulation of Two-Phase Flow Models", submitted to Systems & Control Letters (under review). Theorem 3.3.1 is an adaptation of the content in the contribution submitted to ACC 2020 (see below).

- Chapter 4 is based on the following two contributions:

1. "Port-Hamiltonian modelling of fluid dynamics models with variable cross-section", accepted for 24th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2020).

2. "Power-Preserving Interconnection of Single- and Two-Phase Flow Models for Managed Pressure Drilling", accepted for 2020 American Control Conference (ACC 2020). This publication has been co-authored with M.H. Abbasi. The theoretical section (Section 4.5) has been developed jointly, while the numerical test on the drilling well in Section 4.6 has been performed by M.H. Abbasi. We thank Kelda Drilling Controls, Norway for providing the data to M.H. Abbasi.

- Chapter 5 is an extensive elaboration of the contents in the conference contribution: "Structure-preserving Spatial Discretization of a Two-Fluid Model", accepted for 59th Conference on Decision and Control (CDC 2020). A journal article: "Structure-preserving Discretization of Two-Phase Flow Models" is envisioned based on the contents of this chapter, and is currently under preparation.

- Chapter 6 is based on a journal article that will soon be submitted: "Numerical analysis of the Drift Flux Model".

- Chapter 8 also constitutes novel research. However, we do not plan to submit the contents to any journal for publication. A related contribution, though not in the scope of the thesis, "Reduced order modelling for wafer heating with the Method of Freezing", has been submitted to Proceedings of 13th International Conference on Scientific Computing in Electrical Engineering (SCEE 2020).

- Chapter 9 is an extensive elaboration of the contents in the paper "Model order reduction framework for problems with moving discontinuities", accepted for European Numerical Mathematics and Advanced Applications Conference 2019 Proceedings (ENUMATH 2019). A journal article: "Title to be decided" is envisioned based on the contents of this chapter, and is currently under preparation.

Chapter 2

Managed Pressure Drilling system

In this chapter, we discuss the topic of pressure control during conventional drilling operations. The aspect of pressure control and its criticality is extensively discussed by using a specific drilling well: an Extended Reach Drilling well. We then review a basic configuration of the Managed Pressure Drilling (MPD) system and highlight operational problems during drilling. In order to create a virtual environment for drilling scenario testing and to automate down-hole pressure management, a hydraulic model is required. Some requirements need to be imposed on the hydraulic model and controller in order to represent the scenarios realistically. These requirements are classified into two sub-classes of model requirements, namely, hydraulic model requirements, and controller and estimator requirements. These requirements are thereby described to outline the features of interest that need to be present in the high- and low-complexity models for MPD. Also, the mathematical formulation of the (multi-phase) hydraulic model(s) is introduced.

2.1   Introduction: Issue of pressure control

Drilling processes for deep boreholes consist of different drilling operations, such as tripping (the procedure of pulling out or running in the drill string), choke-swapping/plugging (the contingency of having some cuttings stuck in the choke manifold), and connection (the operation of extending the length of
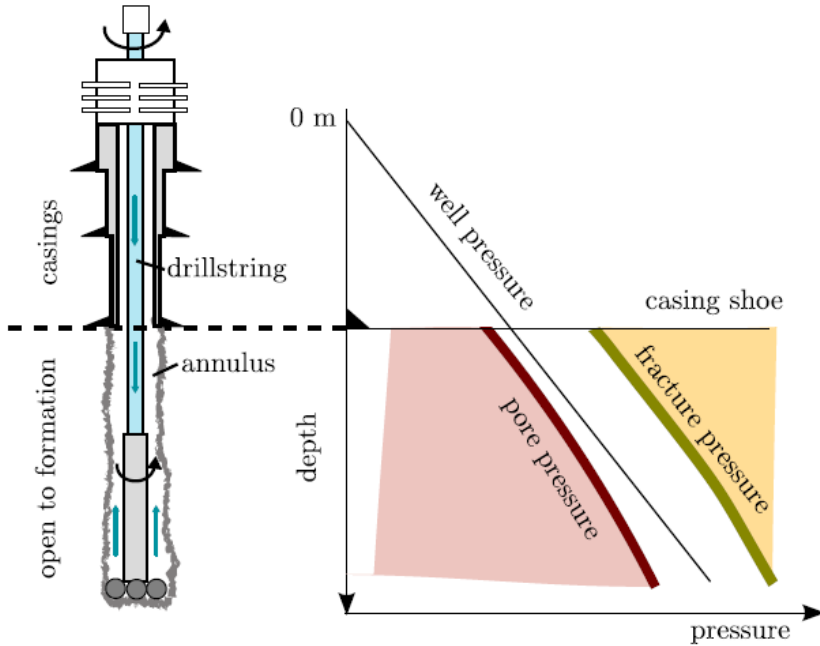
Figure 2.1: Overview of drilling with pressure margins [204].

the drill string through adding some new pieces of pipe). While performing these drilling operations, the bottomhole pressure (BHP), i.e., the pressure at the bottom of the wellbore, has to be maintained within the limits specified by geomechanists. Maintaining the BHP within the pressure limits of the well, i.e., between the pore and fracture pressure, is critical for the safety of drilling operations. If the down-hole pressure exceeds the strength of the formation, i.e., the fracture pressure, the wellbore will be fractured, causing a loss of drilling fluid to the formation, possibly damaging the reservoir. On the opposite side, if the down-hole pressure reduces below the formation pore pressure, it will cause an unwanted influx of formation fluid (liquid or gas) into the well. This is referred to as a well control incident, which in the worst case can escalate to a blow-out of hydrocarbons on the drilling rig. See Figure 2.1 for an overview of conventional/standard drilling set-up and admissible pressure margins. It can be observed that the well is open to the formation below the casing shoe, and

this open/exposed section forms the critical part during drilling operations. In conventional drilling, this exposed part must have a well pressure above the pore pressure and below the fracture pressure.

We consider a specific drilling setting: *Extended Reach Drilling (ERD) wells with long horizontal section* in the rest of this section to put forward critical pressure control issues observed during drilling. ERD wells reduce development costs and minimize the impact on marine ecosystems by avoiding the need for additional offshore structures. Due to their low environmental impact, ERD wells are widely used nowadays. For instance, such wells have been drilled in 3000 ft of water in the Mississippi Canyon of the Gulf of Mexico. Also, ERD wells with long horizontal section are economically beneficial as they result in a larger area for production of oil and gas. Nevertheless, ERD wells have a narrow pressure margin, and it is critical to maintain an accurate and precise pressure control during drilling operations, such as connection, choke-swapping, tripping, etc. Variations in BHP increase the risk of an underbalance (a state where BHP falls below pore pressure) during static wellbore conditions and lead to a higher chance of fracturing the formation in dynamic conditions. For instance, when a tripping operation is performed in an ERD well, there is a high risk of exceeding the fracture pressure because this type of well produces high Equivalent Circulating Density (ECD), which is an effective density exerted by the circulating fluid, if there is no backpressure, i.e., the pressure resulting from restrictions to fluid flow downstream, or friction in the annulus. Fracture pressure is usually the same in the horizontal part of the well but the ECD continues to rise. The risk of exceeding fracture pressure in ERD wells results from the contribution of the annular frictional pressure that increases continuously from the heel to the toe of the horizontal section of the ERD well. High ECD's are also a consequence of surge and swab effects which are frequently encountered during tripping. High ECD is a significant problem in ERD wells as it limits the distance to which the ERD well can be drilled. Annular pressure loss contributions and the contribution of surge and swab effects to ECD calculations/measurements need to be compensated in order to achieve the required length and reach of the well. Most ERD wells have been cut short prematurely because ECD approaches the fracture gradient. To obtain a maximum productivity from the ERD wells, it is necessary to reach the anticipated length and reach of the well.

Furthermore, oil-based muds or synthetic muds are generally used to drill ERD wells. These muds are used due to their lubrication capabilities; but they are highly compressible. The compressibility induces time delay in information propagation and complicates the pressure control in long wells. For instance, tripping operations (e.g., pipe movements and stopping or starting the pumps)

cause transient effects, which result in down-hole pressure variations. Dynamic pressures, such as surge and swab pressures, are detrimental to the wellbore. The criticality can be observed from the Figure 2.2. It can be observed that, while drilling, down-hole pressures (dark blue points) reasonably stay within the optimum window (green shaded area) between fracture gradient (14.1 ppg) and minimum mud weight for wellbore stability (13.0 ppg). Even though some equivalent mud weights (dark blue points) are as low as 12.8 ppg, this does not lead to serious issues for wellbore stability. However, while pulling out of hole (magenta points), pressures fall significantly below minimum mud weight requirements due to swab, triggering massive wellbore instability [1].

Remark 2.1.1 *Though the above discussion is for an ERD well, the cited issues are also relevant to other kind of wells. It is worth mentioning that drilling operations in ERD wells are more critical than carrying out drilling along conventional vertical wells. Accuracy and precision in down-hole pressure management becomes a necessity due to even narrower drilling windows for performing drilling operations in ERD wells.*
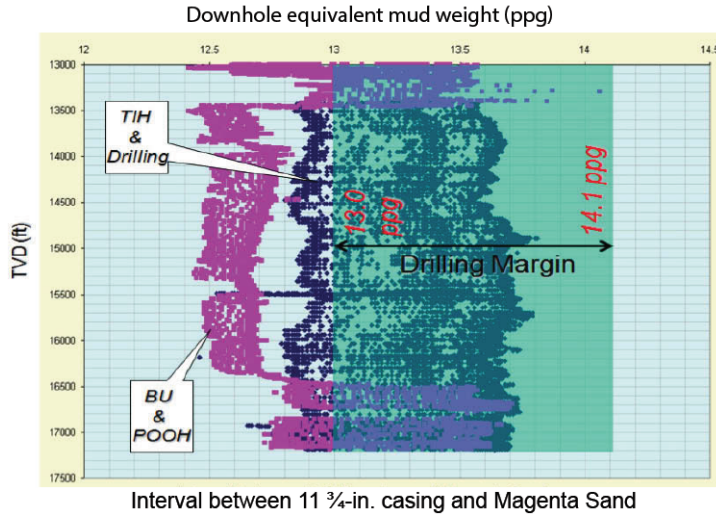
Figure 2.2: Down-hole equivalent mud weight behaviour on a deepwater ERD well. Here, TVD indicates true vertical depth, TIH stands for tripping in hole; and BU & POOH for (circulating) bottom-up and pulling out of hole [1].

The criticality of an accurate pressure control has been highlighted extensively in the introduction of this chapter. Furthermore, as mentioned in Chapter 1, the application of automation in the drilling industry lags behind other industries, for e.g., the aerospace and automotive industries, and currently a growing awareness of the potential benefits of automation is surfacing in the drilling community. Moreover, transient analysis is evidently crucial to drill wells with narrow mud weight windows. For instance, surge and swab analysis would help to control the tripping speed and other movement parameters of the drill string within the wellbore to avoid kicks (a state where formation fluids flow into the wellbore during drilling operations) or blowouts (a state where formation fluids flow into the wellbore in an uncontrolled manner) due to excessive swab or surge pressures. It is also clear that the evaluation of down-hole conditions is of paramount importance. The controller is required to be precise and accurate in actively manipulating the choke opening or in controlling dedicated surface backpressure pump to compensate for down-hole pressure variations. The control system is required to rapidly adjust the pressure in the entire wellbore and keep the BHP within allowable limits. Managed Pressure Drilling (MPD) (refer to Section 2.2) can be utilized in such a situation as it enables to remain within drilling window during drilling operations.

In view of the final goal of automating down-hole pressure management, in this chapter, we first describe a basic configuration of the MPD system. Then, in order to address the important question whether a hydraulic model (and a controller and estimator) should accurately capture fast transients/slow transients or both during several drilling scenarios, we extensively discuss the requirements that must be met by the to-be developed numerical hydraulics model, and a controller and estimator. Finally, we end this chapter with conclusions.

2.2   Basic configuration of MPD system

In addition to the standard/conventional drilling configuration, an MPD system consists of two critical components: a Rotating Control Device (RCD) and a choke manifold. The fluid flow path in MPD system is different from the standard drilling systems due to these additional components. Figure 2.3 represents the basic configuration of the MPD system along with the fluid flow path. The circulation path of the mud/fluid can be observed by following the brown arrows in Figure 2.3. A drilling liquid, known as mud, is pumped into a pipe, called the drill string, at high pressure. At the bottom of the drill string, the mud/fluid leaves the drill string through nozzles created inside the drill bit and

enters the area between the drill string and the wellbore, known as the annulus. It then flows up through the annulus and carries the rock cuttings out of the well. A seal assembly with the RCD enables the mud returns system to remain closed and pressurized, unlike in conventional well drilling situation where the annulus is exposed to the atmosphere. As a result, the RCD diverts the pressur-
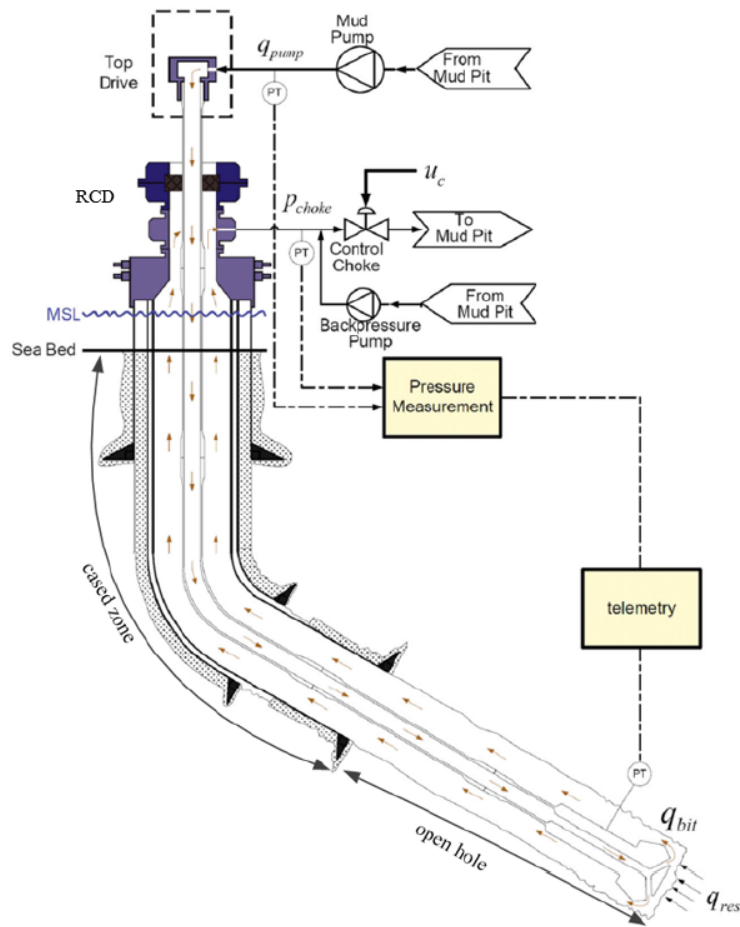
Figure 2.3: A typical drilling setup with MPD configuration [106].

ized mud/fluid returns from the annulus to the choke manifold/valve and, thus, makes closed-loop drilling (CLD) possible. The MPD configuration includes a choke manifold with (multiple) control choke valves. In case one of the two chokes gets plugged, the other choke can be utilized in order to continue the drilling operations. The choke manifold with the pressurized mud/fluid return system helps to apply backpressure across the wellhead. The choke manifold/-valve creates backpressure when the mud/fluid flows across it. The backpressure induced via the choke opening is a function of the amount of the mud/fluid flow across the choke manifold/valve. Hence, when the mudflow decreases, the amount of backpressure that the choke manifold/valve can induce, becomes limited. Also, the mud/fluid flow rate dictates how fast the choke manifold/-valve can act to provide the required backpressure. If the choke manifold/valve cannot provide sufficient backpressure, a backpressure pump (see Figure 2.3) connected to the choke and its control system can automatically be ramped up. A backpressure pump in the MPD configuration also serves as a redundancy in case of sudden loss of pressure caused by mud pump failure or by human errors.

To summarize, MPD allows to maintain the BHP within the drilling window and simultaneously to eliminate dynamical (down-hole) pressure variations by manipulating surface back-pressure and redirecting fluid flow along the alternate flow path. The latter goes across the MPD choke manifold to provide necessary surface backpressure across wellhead by manipulating choke opening.

2.3   Model Requirements

In order to create a virtual environment for drilling scenario testing and to support model-based automation of down-hole pressure management, a hydraulic model and a controller and estimator are required. Some model requirements are necessary to mimic the scenarios realistically. These requirements are classified into two sub-classes of model requirements namely: Hydraulic Model Requirements, and Controller and Estimator Requirements.

2.3.1   Hydraulic Model Requirements

Proper hydraulic modelling capabilities are extremely critical for MPD operations as loss of accuracy is unacceptable when it comes to down-hole pressure management. Hydraulic analysis will help in determining a progressive approach to support pressure control by calculating the down-hole pressures in real-time depending on the mud/fluid properties and the down-hole conditions.

The physics of the drilling process dictates that a dynamic model, based on the unsteady flow, must be developed for the prediction of variations in pressures across the wellbore. Fast changing pressure transients are caused by changes in fluid flow rate, drill string rotation rates or axial speed of the drillstring movement. Transients are also caused by unexpected events affecting the mud/fluid flow, such as plugging (typically nozzles of the drill bit and the choke valve get plugged) or due to (liquid or gas) influx or lost circulation. Slower changes to the wellbore pressure are caused by operational changes in mud/fluid properties (density or rheology), temperature effects, the amount of cuttings in suspension, building or erosion of cuttings beds or hole enlargement [76]. The hydraulic model must, hence, be effectively able to capture the pressure propagation, both in fast and slow transients, during all the phases of drilling operations irrespective of the circulation rates of the drilling mud/fluid. Also, hydraulics modelling needs to characterize the effect of the surface backpressure, which is induced by manipulation of the opening of the choke valve.

The drill string and the Bottomhole Assembly (BHA) are part of MPD systems through, and around, which single-phase and multi-phase fluids flows take place. These flow paths have different geometrical specifications. Hence, the flow area in the annular section of the well varies along the spatial location in the well. In addition, the flow area changes dynamically due to the axial movements of the integrated drill string and the BHA system. This dynamical change depends on the position of the drill string and the BHA inside the well. Hence, the dynamic model must take into account cross-sectional area variations. Cross-section variation affects the down-hole pressure. Namely, it alters the pressure transmission between the top and down-hole parts of the well, because part of the pressure wave is reflected and part of it is transmitted at the point where the geometrical cross-section changes. Oscillatory pressure profiles may be induced more frequently compared to the case where there are no geometrical cross-sectional changes along the well. The convergence to a steady-state situation may become slower with the inclusion of changes in the geometrical cross-section. Sudden pressure changes due to axial pipe (here, drill string) movements, as well as stopping and restarting the pumps, lead to so called water hammer effects. Travelling waves and standing waves result due to water hammer effects and dynamical cross-sectional area variations. These water hammer effects need to be incorporated in the hydraulic model. Capturing *wave propagation phenomenon* is a central concept in hydraulic modelling throughout this dissertation. The hydraulic model must be able to accurately capture the wave propagation for a better prediction of the down-hole characteristics. This motivates us to investigate single- and multi-phase models with

time-dependent and spatially varying cross-sections of the flow path.

We draw our attention to sub-modules, i.e., hydraulics across the drill string, annulus and the drill bit, in the MPD set-up in the rest of this dissertation. If no contingency happens during drilling, the flows inside the drill string and annulus are of a single-phase flow nature. However, if a gas influx occurs in the formation, the flow inside only the annulus involves two phases. Hence, a (complete) hydraulics model can be characterized by interconnection of subsystems (here, the drill string and the annulus) governed by single- and two-phase flow models, and by mathematical models governed by non-linear ordinary differential equations (for instance, a hydraulics model across a bit).

The (hydraulic) dynamics along the length of the wellbore are of crucial importance. Radial dimensions of the wellbore are much smaller than the length of the wellbore. Hence, the fluid flow is modelled via one-dimensional (1-D) single- and two-phase flow models. The one-dimensional single- and two-phase flow models are mathematically introduced next.

Single-phase flow model

A single-phase flow across a geometry with variable cross-section is mathematically modeled by isothermal Euler equations as in [105]. The model is basically governed by mass and momentum conservation laws along with the equation of state for the phase of interest. These equations read as follows:

$$\begin{cases} \partial_t(A\rho) + \partial_x(A\rho v) = 0, \\ \partial_t(A\rho v) + \partial_x(A\rho v^2 + Ap) = AS + p\partial_x A, \\ \rho = \rho_{\ell 0} + \frac{p - p_{\ell 0}}{c_\ell^2}, \end{cases} \tag{2.1}$$

where $S = -\rho g \sin\theta - \frac{32\mu v}{d^2}$. Here, $t \in \mathbb{R}_{\geq 0}$ and $x \in [a, b]$ are, respectively, the time and the spatial domain (a and b refer to the location of the left and the right boundary of the one-dimensional spatial domain). Variables ρ, v, p, $p_{\ell 0}$, $\rho_{\ell 0}$, A, g, μ, d and θ refer to density, velocity, pressure, reference pressure, reference density of the phase of interest, cross-section area, gravitational constant, fluid viscosity, the diameter of the well/pipe, and, the pipe inclination, respectively.

Two-phase flow model

Different variants of two-phase flow models exist in literature. We primarily focus on the following two variants of two-phase flow models: the Drift Flux

Model and the Two-Fluid Model. We first discuss the Drift Flux Model, and then the Two-Fluid Model.

Drift Flux Model:

The Drift Flux Model (DFM) [56, 57] is derived from a set of non-linear conservation laws, written in multi-fluid formulation in a certain asymptotic limit. It is based on a single pressure formulation and is an approximate macroscopic formulation compared to more rigorous two-fluid models [66, 107, 124, 129], which describe two-phase flow in more detail. However, the DFM caters to understand the physical phenomena of our interest. The (transient) DFM consists of a system of non-linear (conditional) hyperbolic partial differential equations (PDEs) and algebraic closure relations. It constitutes the mass conservation equation for liquid phase, mass conservation equation for the gaseous phase and the mixture momentum conservation equation. It is applicable to general multi-phase flow problems with arbitrary flow speeds and arbitrary levels of compressibility. It should also be mentioned that the hyperbolicity of the model is conditional. The flow patterns, dictated by the flow-dependent parameters in the slip relation, serve as constraining conditions and result in the loss of hyperbolicity of a convective subset, and, hence, in ill-posedness of the model equations.

A two-phase flow across a geometry with variable cross-section can be modelled by the 1-D DFM as in [5]. The partial differential equations (PDEs) read as follows:

$$\begin{cases} \partial_t(Am_\ell) + \partial_x(Am_\ell v_\ell) = 0, \\ \partial_t(Am_g) + \partial_x(Am_g v_g) = 0, \\ \partial_t(A(m_\ell v_\ell + m_g v_g)) + \partial_x(A(m_\ell v_\ell^2 + m_g v_g^2)) + A\partial_x p = -A\tilde{S}, \end{cases} \tag{2.2}$$

where $\tilde{S} = \left(g(m_g + m_\ell)\sin\theta + \frac{32\mu_m v_m}{d^2}\right)$, and $t \in \mathbb{R}_{\geq 0}$ and $x \in [a, b]$ are the time and the spatial variables, respectively. The abbreviations $m_\ell := \alpha_\ell \rho_\ell$ and $m_g := \alpha_g \rho_g$ have been used. Variables α_ℓ and α_g, respectively, denote liquid and gas void fraction, ρ_ℓ and ρ_g refer to the density of the liquid and the gas phase, respectively, v_g and v_ℓ refer to the velocity of the gaseous and liquid phase, respectively, μ_m is the mixture viscosity, and v_m is the mixture velocity.

There are seven unknowns: $\alpha_g, \alpha_\ell, \rho_g, \rho_\ell, v_g, v_\ell, p$, and only three governing partial differential equations. Hence, certain closure laws are required to close the model. The four closure equations are obtained using the property of volume conservation, equations of state for the liquid and the gaseous phase, and

the empirical slip relation. Using the property of volume conservation, i.e., any segment of the wellbore is occupied by the combination of the gaseous and the liquid phase, we have

$$\alpha_g + \alpha_\ell = 1. \tag{2.3}$$

Next, we use fluid density-pressure relationships (or equations of state). The equation of state for the gaseous phase is:

$$\rho_g = p/c_g{}^2, \tag{2.4}$$

where c_g is the speed of sound in the gaseous phase. Similarly, the equation of state for the liquid phase is:

$$\rho_\ell = \rho_{\ell 0} + (p - p_{\ell 0})/c_\ell{}^2, \tag{2.5}$$

where c_ℓ is the speed of sound in the liquid phase, $\rho_{\ell 0}$ is the reference density and P_{l0} is the reference pressure under nominal conditions.

The last closure relation is the slip law that has the following form:

$$v_g = (K v_\ell \alpha_\ell + S)/(1 - K \alpha_g), \tag{2.6}$$

where K and S are flow-dependent parameters, which, in principle, are functions of the gas void fraction. The slip law is the only relation that is empirical and uncertain in the description of the DFM. The slip law is mostly determined based on the experiments in certain setting for some set of fluids under certain operating conditions. The slip relation in (2.6), defined by Zuber and Findlay as in [56], is a constitutive equation to specify the relative motion between the two phases under consideration. Such a relation allows to consider unequal phase velocities. Such a concept plays an important role when the densities of the two phases are quite different in the presence of the gravity field as density difference is known to result in buoyancy effects. In principle, the slip law model should account for different flow regimes which exist across different sections of the wellbore. The flow-dependent parameters, i.e., K and S, take account of the non-uniform distribution of the gaseous phase and velocity profile over the wellbore cross section.

Remark 2.3.1 *The slip law* (2.6) *is valid for the slug and the bubbly flow regimes [24, 56, 69]. There is a singularity in the slip law when we approach the pure gas region. We note from* (2.6) *that the singularity in the slip law depends on flow parameter K. Several slip laws exist in the literature and each of them comes up with different hyperbolicity restrictions. See [46, 88, 93] for more exhaustive list of empirical correlations.*

In addition to the governing PDEs and the algebraic closure relations, the model needs to be complemented with appropriate initial and boundary conditions. Physical boundary conditions are independent of the numerical method used to solve the governing PDEs. These boundary conditions can be specified in terms of volumetric inflow rates of liquid and gaseous phase along with certain pressure restriction at the outlet. Inlet boundary conditions can also be expressed in terms of mass flow rates of liquid and gaseous phase. The general form of the physical boundary conditions can be given by

$$\begin{cases} (\rho_\ell \alpha_\ell v_\ell)(x=0,t) = f(t), \\ (\rho_g \alpha_g v_g)(x=0,t) = h(t), \quad \text{or} \\ p(x=L,t) = r(t), \end{cases} \quad \begin{cases} (\alpha_\ell v_\ell)(x=0,t) = f(t), \\ (\alpha_g v_g)(x=0,t) = h(t), \\ p(x=L,t) = r(t), \end{cases} \quad (2.7)$$

where functions $f(t)$, $h(t)$ and $r(t)$ in (2.7) are time-dependent functions.

Remark 2.3.2 *In principle, the conditions to be imposed at the given boundary depends on whether the flow is subsonic, transonic or supersonic. We assume that only subsonic conditions exist, which is mostly the case during drilling. The above form of the physical boundary conditions is only valid if the two-phase flow is in a sub-sonic state such that two boundary conditions need to be imposed on the left side of the spatial domain and one boundary condition needs to be imposed on the right side of the spatial domain.*

This two-phase single pressure model approach for describing the DFM is justified by the very short time-scale linked with the phenomena of the relaxation of two pressures towards an equilibrium. It is, however, worth mentioning that such a model is characterized by a system of non-linear evolution equations that change type from hyperbolic to elliptic or parabolic. In other words, the model is conditionally hyperbolic. This hyperbolicity is closely linked to the choice of the closure laws and is also impacted by the presence of spatial derivatives in the source terms.

Two-Fluid Model:

The Two-Fluid Model (TFM) is a set of Partial Differential Equations (PDEs) and algebraic closure relations. The PDEs expressing mass and momentum con-

servation for each phase are as follows:

$$\partial_t \left(\alpha_g \rho_g \right) + \partial_x \left(\alpha_g \rho_g v_g \right) = 0, \tag{2.8a}$$

$$\partial_t \left(\alpha_\ell \rho_\ell \right) + \partial_x \left(\alpha_\ell \rho_\ell v_\ell \right) = 0, \tag{2.8b}$$

$$\partial_t \left(\alpha_g \rho_g v_g \right) + \partial_x \left(\alpha_g \rho_g v_g^2 \right) = -\partial_x \left(\alpha_g p \right) + M_g - \alpha_g \rho_g g \sin\theta - 32 \frac{\mu_g \alpha_g v_g}{d^2}, \tag{2.8c}$$

$$\partial_t \left(\alpha_\ell \rho_\ell v_\ell \right) + \partial_x \left(\alpha_\ell \rho_\ell v_\ell^2 \right) = -\partial_x \left(\alpha_\ell p \right) + M_\ell - \alpha_\ell \rho_\ell g \sin\theta - 32 \frac{\mu_\ell \alpha_\ell v_\ell}{d^2}, \tag{2.8d}$$

where $t \in \mathbb{R}_{\geq 0}$ and $x \in [a, b]$ are, respectively, the temporal and spatial variables. The model contains seven unknown variables, namely, liquid and gas void fraction, α_ℓ and α_g, liquid and gas phase velocity, v_ℓ and v_g, liquid and gas phase density, ρ_ℓ and ρ_g, and the common pressure p. Here, μ_g and μ_ℓ refer to the viscosity of the gaseous phase and of the liquid phase, respectively.

To complete the model, we use one set of the most widely applied closure laws as in [56]:

$$\alpha_g + \alpha_\ell = 1, \tag{2.9a}$$

$$M_g + M_\ell = 0, \tag{2.9b}$$

$$M_g = p \partial_x \alpha_g + M_{ig}, \tag{2.9c}$$

$$M_{ig} = b_g^M (v_\ell - v_g), \quad \text{with } b_g^M \geq 0, \tag{2.9d}$$

$$\rho_g = \frac{p}{c_g^2}, \tag{2.9e}$$

$$\rho_\ell = \rho_{\ell 0} + \frac{p - p_{\ell 0}}{c_\ell^2}, \tag{2.9f}$$

where (2.9a) expresses that any pipe segment is occupied by the combination of gas and liquid. The terms M_g and M_ℓ with the constant b_g^M in (2.9b)–(2.9d) account for the force interaction between the phases. Finally, (2.9e)–(2.9f) define the equation of state of each phase with the reference density and pressure as $\rho_{\ell 0}$ and $p_{\ell 0}$, and c_g and c_ℓ are the constant speeds of sound in the gas and liquid phase, respectively.

Similar to the DFM, even the TFM can be complemented with appropriate initial and boundary conditions.

Remark 2.3.3 *We have presented the TFM for flows across a unit and constant cross-section. Such a model will be utilized in the next chapter. In principle, there exist two-fluid models for describing the fluid flow dynamics across a variable cross-section. However, we do not rely on the TFM admitting (spatial) cross-section vari-*

ations, and instead only rely on the DFM for numerically studying the behaviour across variable cross-section in this dissertation.

The fundamental difference between the TFM and the DFM lies in the formulation of momentum conservation equations. Similar to the DFM, the TFM is also conditionally hyperbolic. However, the hyperbolicity regimes of the TFM and the DFM differ. We refer to [56, 103, 110, 127, 138, 200, 206] for a detailed analysis of the hyperbolicity of these two models. Also, we refer to Chapter 6 for detailed discussion on the hyperbolicity of the DFM. The DFM is more widely used than the TFM due to its simplicity and low computational time required.

The above introduced mathematical models will be extensively used in the rest of this dissertation. At this stage, it worth mentioning that thermal and elastic effects also play a dominant role in addition to hydraulic effects in dictating the down-hole pressure. However, incorporating these effects are not the subject of research in the scope of this dissertation.

2.3.2 Controller and Estimator Requirements

A control system which is fast and accurate enough is necessary to control the pressure fluctuations due to dynamical effects during drilling. The control system must be able to prevent gas influx with the assistance of an automated MPD system (whose configuration was discussed in Section 2.2).

Automated down-hole pressure management relies heavily on real-time measurements. Real-time measurement of BHP, which is the pressure to be controlled, is not available during several drilling operations, such as tripping. Hence, the BHP needs to be estimated. It is worth mentioning that the estimation of down-hole pressures is sensitive to several factors: the uncertain wellbore position in long wells, the surface and down-hole equipment, the density and thermophysical properties of the drilling mud, the measurements, the well-flow models, the down-hole conditions, the operating procedures and their performance and the pressure control method [76]. The BHP can be estimated by using topside measurements such as pump pressure, choke pressure, pump flow rate, choke flow rate, etc., as these are reliable and available continuously. One can utilize these topside measurements along with the hydraulic model to make real-time decisions while performing drilling operations.

MPD pressure controllers are required to track a desired BHP when reliable real-time measurements are available; or, else, they are required to track a choke pressure set point which is deduced from BHP estimates. In order to achieve successful MPD operation during several drilling scenarios, the tracking

of the choke pressure set point is required as real-time measurements are not available and, hence, BHP needs to be estimated in real-time to enable appropriate choke pressure set point selection. Predicting and controlling, for instance, surge and swab pressures have to be based on some estimations due to the lack of the measurements. Reservoir pore pressure also needs to be estimated to have the dynamic knowledge of the drilling window.
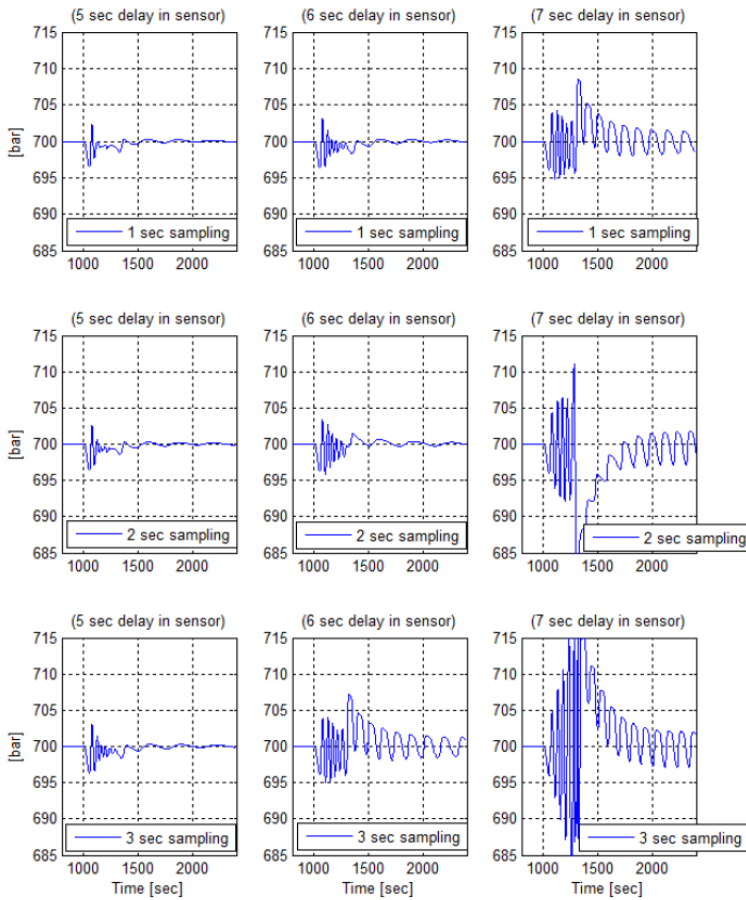
Figure 2.4: Consequences of uncertainties in sensor delay and bandwidth [76].

From a control point of view, gel effects in the drilling mud (which cause delay in pressure transmission), and the time delay imposed by the distance between the choke and the bottom of the well, make the control of the BHP challenging. Also, pump rates typically change in less than 10 seconds, and the MPD system controller operates at even shorter time-scales. An MPD controller operating at shorter time-scales than a pump rate change is also generally operating faster than the acoustic travel time in a typical well. To compound the challenge, for most wells the acoustic travel time for pressure waves in the fluid in the annulus is significant compared to the time for control transients. These critical factors need to be taken into account, otherwise the pressure at the choke can be over-perturbed and the system can become unstable. Moreover, uncertainties in sensor delay and bandwidth associated with the down-hole telemetry also degrade the controller performance. The quality of the sensory data also affects the pressure control. The closed-loop performance of MPD controller may worsen when there is a change in the dynamic properties of the well. Also, pressure sensor readings drift under harsh environmental conditions. Consequences of the delay in measurements and of the bandwidth can be understood from the Figure 2.4, where the BHP is plotted for different combinations of sampling times and measurement delays. It can be observed that the BHP varies up to ±15 bars. However, the drilling window for today's wells is ±2.5 bars and, therefore, the variations in BHP upto ±15 bars are not allowed. Hence, the knowledge of sampling times and the measurement delays are critical to keep the BHP within the drilling window. The longer is the wellbore, the lower is the sensor bandwidth as the pressure measurement would be updated less frequently which consequently means lower sampling time. The transmission rate deteriorates as the depth/length of the well increases, which is usually the case for deepwater wells. Under transmission delays, lower sampling times can result in adverse closed-loop system response. A robust controller must be able to account for these low sampling times and delays in transmission. Another major challenge is transmitting the copious amounts of data to the surface in a reasonable amount of time. When wireless data transfer techniques, such as mud pulse, are used, for example, much of this data must be stored in memory and downloaded during a certain drilling phase. This option, in which data is stored and then evaluated after some drilling operation, makes real-time drilling decisions nearly impossible if there does not exist an accurate hydraulic model to predict a characteristic behaviour/response. Hence, it is clear that the real-time prediction of quantities of interest (say pressure), while drilling, would be an enormous advantage for any well being drilled. The real-time down-hole pressure estimations would allow, for instance, to set the optimum tripping speed,

and also to assist in multivariable control of choke valve opening, mud pump and back pressure pumps simultaneously.

Though we outlined the controller and estimator requirements in the scope of the drilling application, the design of controller and estimator will not be treated in this dissertation as it lies outside the scope of this work. However, the controller and estimator requirements do dictate the need of an underlying accurate hydraulic model which is computable in real-time. The whole focus in this dissertation will, hence, be on (low-complexity) hydraulics model development which, in the future, will serve as the basis for controller and estimator design.

2.4   Conclusions

In the pursuit of obtaining an improved automatic pressure control in drilling operations, we discussed Managed Pressure Drilling (MPD) technology. We also provided fundamental insights to the physics of drilling hydraulics and the basic features that need to be mimicked by the high- (and low-) complexity model.

A fit-for-purpose hydraulic model, which needs to accurately capture the relevant properties of the complex multi-phase flow dynamics of the down-hole drilling process, while being fit (i.e., simple enough) for model-based control design, is a fundamental prerequisite for automatic pressure control in case of well control incidents. The model simplicity is also required to perform multi-query simulations for the planning of drilling operations. To this end, we need to develop low-complexity model representations from high-fidelity counterparts governing hydraulic systems in the scope of MPD.

Part I: Modelling and (structure-preserving) discretization of hydraulic systems

This part is dedicated to the contributions pertaining to the modelling and (structure-preserving) discretization of hydraulic systems. In particular, we strive to address questions Q1 - Q5 posed in Chapter 1 by fulfilling outlined research objectives O1 - O7. Chapter 3 is an effort in the direction of achieving research objectives O1 and O2. The research objective O1 is fully met in Chapter 4. This chapter also attains the research objective O3. The contributions in Chapter 5 help to successfully meet the research objective O4. Research objectives O5 - O7 are addressed in Chapter 6.

Chapter 3

Port-Hamiltonian Formulation of Two-Phase Flow Models

Two-phase flows are frequently modelled and simulated using the Two-Fluid Model (TFM) and the Drift Flux Model (DFM). This chapter proposes Stokes-Dirac structures with respect to which port-Hamiltonian representations for such two-phase flow models can be obtained. We introduce a non-quadratic candidate Hamiltonian function and present dissipative Hamiltonian representations for both models. We then use the structure of the corresponding formally skew-adjoint operator to derive a Stokes-Dirac structure in the scope of the two variants of multi-phase flow models. Moreover, we present a numerical counter example to demonstrate that only a special form of the DFM (without slip between the phases) can be cast in a port-Hamiltonian representation and that the DFM with the Zuber-Findlay slip conditions is not an energy-consistent model for two-phase flow.

3.1 Introduction

In this chapter, we develop a port-Hamiltonian formulation for modelling multi-phase flow dynamics in pipes. Multi-phase flow is important in a large range of industrial applications, such as those in the oil and gas industry, chemical and process industry (including heat-pumping systems) as well as in the safety analysis of nuclear power plants [2, 5, 145]. Within the oil and gas industry, such models are used for virtual drilling scenario testing [2, 5]. The multi-

phase aspect is particularly relevant in these applications in case of gas influx occurring from a reservoir.

Port-Hamiltonian systems have recently received a lot of attention for modelling physical phenomena governed by nonlinear Partial Differential Equations (PDEs) and ordinary differential equations [53, 182]. A port-Hamiltonian realization offers a suitable description for the modelling, analysis and controller design [53]. A port-Hamiltonian model formulation is known to provide a modular framework for multi-physics and interconnected systems [118]. The port-Hamiltonian structure allows for non-zero energy flow through the boundary and guarantees power preservation [184]. The advantage of this formulation is that structure-preserving methods for discretization and the model order reduction of infinite-dimensional port-Hamiltonian systems can preserve certain original system-theoretic properties such as stability and passivity [43, 140]. A port-Hamiltonian framework enables controller design based on energy considerations by different techniques such as energy-shaping [113], and interconnection and damping assignment [136]. In addition, the Hamiltonian defined in a port-Hamiltonian framework represents a good candidate for the Lyapunov function, rendering the physics-based control design and the stability analysis more tangible [114].

In the literature, the infinite-dimensional port-Hamiltonian structure has been exploited in several domains of science and engineering. For instance, some well-known fluid dynamical systems such as the shallow water equations [140], reactive Navier Stokes equations [13], and reaction diffusion processes [207] have already been formulated in the port-Hamiltonian formalism. Such a representation is also prevalent in the fields of structural dynamics [115] and fluid-structure interaction [34].

Multi-phase flows are mathematically governed by conservation laws. Several conservation laws have previously been converted to port-Hamiltonian representations [117, 178]. Some work on Hamiltonian modeling for multi-phase hydrodynamics has been done in [84]. However, (dissipative) Hamiltonian representations do not exist for the Two-Fluid Model (TFM) and the Drift Flux Model (DFM) [56]. Moreover, until now, to the best of our knowledge, port-Hamiltonian modeling for fluid dynamics only encompasses single-phase models [48].

Matrix/operator theory for *linear* distributed parameter port-Hamiltonian systems on one-dimensional domains is owed to some pioneering works [89, 101]. The central theme of the current chapter is to extend and propose modifications to the existing theory for *non-linear* distributed parameter systems. We exploit the existing theory in the scope of linear systems and arrive at new

results from an operator theoretic viewpoint, including further generalizations in the scope of non-linear distributed parameter port-Hamiltonian systems.

The main contributions of this chapter are as follows: *(i)* we obtain (dissipative) Hamiltonian representations of the TFM and the DFM, *(ii)* we propose state-dependent Stokes-Dirac structures for both the TFM and the DFM along with the proof of the corresponding representation obtained in the scope of the TFM, and *(iii)* we demonstrate that the DFM with the Zuber-Findlay slip conditions is not an energy-consistent model for two-phase flow.

The chapter is organized as follows. In Section 3.2, we introduce the two mathematical models governing 1-D multi-phase flow dynamics and mention under which conditions these are equivalent. The (dissipative) Hamiltonian representations of these models are presented in Section 3.3. Then, the corresponding geometrical properties are discussed and proved in Section 3.4. This section also includes a non-unique parametrization of the boundary port-variables. Afterwards, Section 3.5 deals with the reasons behind formulating the DFM without slip between the two phases in a port-Hamiltonian representation instead of a general DFM with the Zuber-Findlay slip conditions. Finally, Section 3.6 closes with conclusions.

3.2 Multi-phase flow models

In this section, we briefly recall the mathematical models for the TFM and the DFM. We refer the reader to Section 2.3.1 for a detailed discussion about these multi-phase flow models.

3.2.1 Two-Fluid Model (TFM)

In this chapter, we do not consider gravitational and frictional effects in the TFM description for the sake of simplicity. Hence, under this assumption, the PDEs expressing mass and momentum conservation for each phase are as follows:

$$\partial_t \left(\alpha_g \rho_g \right) + \partial_x \left(\alpha_g \rho_g v_g \right) = 0, \tag{3.1a}$$

$$\partial_t \left(\alpha_\ell \rho_\ell \right) + \partial_x \left(\alpha_\ell \rho_\ell v_\ell \right) = 0, \tag{3.1b}$$

$$\partial_t \left(\alpha_g \rho_g v_g \right) + \partial_x \left(\alpha_g \rho_g v_g^2 \right) = -\partial_x \left(\alpha_g p \right) + M_g, \tag{3.1c}$$

$$\partial_t \left(\alpha_\ell \rho_\ell v_\ell \right) + \partial_x \left(\alpha_\ell \rho_\ell v_\ell^2 \right) = -\partial_x \left(\alpha_\ell p \right) + M_\ell, \tag{3.1d}$$

where $t \in \mathbb{R}_{\geq 0}$ and $x \in [a, b]$ are, respectively, the temporal and spatial variables (a and b refer to the location of the left and the right boundary of the one-dimensional spatial domain). To complete the model, we use the algebraic closure laws given by:

$$\alpha_{\mathrm{g}} + \alpha_\ell = 1, \tag{3.2a}$$

$$M_{\mathrm{g}} + M_\ell = 0, \tag{3.2b}$$

$$M_{\mathrm{g}} = p \partial_x \alpha_{\mathrm{g}} + M_{ig}, \tag{3.2c}$$

$$M_{ig} = b_g^M (v_\ell - v_{\mathrm{g}}), \quad \text{with } b_g^M \geq 0, \tag{3.2d}$$

$$\rho_{\mathrm{g}} = \frac{p}{c_g^2}, \tag{3.2e}$$

$$\rho_\ell = \rho_{\ell 0} + \frac{p - p_{\ell 0}}{c_\ell^2}. \tag{3.2f}$$

The model contains seven unknown variables. A detailed explanation of the system governed by (3.1) and (3.2) is provided in Section 2.3.1.

Remark 3.2.1 *The ignored gravitational and frictional effects in the above TFM description will be accounted for in the next chapter while developing a composed port-Hamiltonian model formulation for MPD systems.*

The TFM, governed by the set of equations (3.1) and (3.2), can be written in terms of only four physical variables. We introduce the following shorthand notations that are used in the rest of this chapter:

$$m_g := \alpha_{\mathrm{g}} \rho_{\mathrm{g}}, \quad m_\ell := \alpha_\ell \rho_\ell. \tag{3.3}$$

We now adopt the following assumption in this chapter.

Assumption 1 *The gas void fraction α_{g}, the liquid void fraction α_ℓ, the liquid phase density ρ_ℓ, and the gaseous phase density ρ_{g} along with $\beta = \rho_{\ell 0} c_\ell^2 - p_{\ell 0}$ are positive.*

The statement in Assumption 1 is physically (always) true. However, the positivity of different physical quantities cannot be mathematically/numerically guaranteed and, hence, we need to pose this assumption.

The model of interest is a non-linear partial differential algebraic system. In the scope of developing a port-Hamiltonian model, such a model description

intuitively/naturally will call for a non-linear port-Hamiltonian descriptor realization in the sense of [123]. However, the area of non-linear port-Hamiltonian descriptor realizations is still in the initial stages of theoretical development, in particular for partial differential algebraic systems with non-quadratic Hamiltonian functionals and with index greater than 2. The two-phase flow model of interest possesses non-quadratic Hamiltonian functionals (see Section 3.3). Deferring to speak in the sense of the index of the system, it is worth to point out that the system constitutes a large number of (non-linear) algebraic equations which increases technical complexities. Furthermore, we encounter technical challenges while developing a non-linear port-Hamiltonian descriptor realization with the current state-of-the-art. Given the aforementioned reasons and that the model of interest, i.e., a partial differential algebraic system (3.1) and (3.2), can be rather straightforwardly re-written as a system of (only) partial differential equations, we strive to exploit the existing theory to develop a corresponding port-Hamiltonian realization. Next, we discuss Lemma 3.2.2 and 3.2.3 in order to provide insights on equivalent model realizations of the TFM governed by (3.1) and (3.2).

Lemma 3.2.2 *By considering m_g, m_ℓ, v_g and v_ℓ as state variables, the system of equations* (3.1) *and* (3.2) *can be re-written in the following form:*

$$\partial_t m_g + \partial_x \left(m_g v_g \right) = 0, \tag{3.4a}$$

$$\partial_t m_\ell + \partial_x \left(m_\ell v_\ell \right) = 0, \tag{3.4b}$$

$$\partial_t v_g + \partial_x \left(\frac{v_g^2}{2} \right) = -c_g^2 \partial_x \left(\ln \left(\frac{p}{c_g^2} \right) \right) + \frac{b_g^M}{m_g} \left(v_\ell - v_g \right), \tag{3.4c}$$

$$\partial_t v_\ell + \partial_x \left(\frac{v_\ell^2}{2} \right) = -c_\ell^2 \partial_x \left(\ln \left(\frac{p+\beta}{c_\ell^2} \right) \right) - \frac{b_g^M}{m_\ell} \left(v_\ell - v_g \right), \tag{3.4d}$$

where

$$p \left(m_g, m_\ell, \alpha_g \right) = m_g c_g^2 + m_\ell c_\ell^2 - \beta \left(1 - \alpha_g(m_g, m_\ell) \right), \tag{3.5}$$

$$\alpha_g \left(m_g, m_\ell \right) = -m_g \frac{c_g^2}{2\beta} - m_\ell \frac{c_\ell^2}{2\beta} + \frac{1}{2} + \sqrt{\left(m_g \frac{c_g^2}{2\beta} + m_\ell \frac{c_\ell^2}{2\beta} - \frac{1}{2} \right)^2 + m_g \frac{c_g^2}{\beta}}. \tag{3.6}$$

Proof: First, we note that (3.4a) and (3.4b) are exactly of the same form as (3.1a) and (3.1b), respectively. Next, we explain the steps to obtain (3.4c) and

(3.4d). Using (3.1a), we can express (3.1c) in the following manner:

$$\partial_t v_{\text{g}} + \partial_x \left(\frac{v_{\text{g}}^2}{2} \right) = -\frac{\alpha_{\text{g}}}{m_{\text{g}}} \partial_x p + \frac{b_{\text{g}}^M}{m_{\text{g}}} \left(v_\ell - v_{\text{g}} \right). \tag{3.7}$$

Equation (3.7) can be written as (3.9c) since the following identity holds:

$$c_g^2 \partial_x \left(\ln \left(\frac{p}{c_g^2} \right) \right) = \frac{1}{\rho_{\text{g}}} \partial_x p =: \frac{\alpha_{\text{g}}}{m_{\text{g}}} \partial_x p. \tag{3.8}$$

Equation (3.4d) can be obtained by following similar lines of reasoning as described above. In summary, the set of equations (3.4) is equivalent to (3.1) and (3.2). We refer the reader to [5] for the detailed proof of the expression for $\alpha_{\text{g}}(m_{\text{g}}, m_\ell)$. This completes the proof. ∎

Similar to (3.4), the TFM can be straightforwardly formulated with respect to another set of state variables. The corresponding model reformulation is discussed next.

Lemma 3.2.3 *By considering m_{g}, m_ℓ, $m_{\text{g}} v_{\text{g}}$ and $m_\ell v_\ell$ as state variables, the system of equations (3.1) and (3.2) can be re-written in the following form:*

$$\partial_t m_{\text{g}} + \partial_x (m_{\text{g}} v_{\text{g}}) = 0, \tag{3.9a}$$

$$\partial_t m_\ell + \partial_x (m_\ell v_\ell) = 0, \tag{3.9b}$$

$$\partial_t (m_{\text{g}} v_{\text{g}}) + \partial_x (m_{\text{g}} v_{\text{g}}^2) + m_{\text{g}} \partial_x \left(c_g^2 \ln \left(\frac{p}{c_g^2} \right) \right) = b_g^M (v_\ell - v_{\text{g}}), \tag{3.9c}$$

$$\partial_t (m_\ell v_\ell) + \partial_x (m_\ell v_\ell^2) + m_\ell \partial_x \left(c_\ell^2 \ln \left(\frac{p + \beta}{c_\ell^2} \right) \right) = -b_g^M (v_\ell - v_{\text{g}}), \tag{3.9d}$$

where p and α_{g} are given by the relations (3.5) and (3.6), respectively.

Proof: First, we note that (3.9a) and (3.9b) are exactly of the same form as (3.1a) and (3.1b), respectively. Next, we explain the steps to obtain (3.9c) and (3.9d). Using (3.2c) in (3.1c), we obtain:

$$\partial_t (m_{\text{g}} v_{\text{g}}) + \partial_x (m_{\text{g}} v_{\text{g}}^2) = -\alpha_g \partial_x p + b_g^M (v_\ell - v_{\text{g}}), \tag{3.10}$$

which can be equivalently expressed as follows:

$$\partial_t (m_{\text{g}} v_{\text{g}}) + \partial_x (m_{\text{g}} v_{\text{g}}^2) + \alpha_g \partial_x p = b_g^M (v_\ell - v_{\text{g}}). \tag{3.11}$$

Equation (3.11) can be written as (3.9c) since the following relation holds:

$$m_g \partial_x \left(c_g^2 \ln\left(\frac{p}{c_g^2} \right) \right) = m_g \left(\frac{c_g^2}{p/c_g^2} \right) \left(\frac{1}{c_g^2} \right) \partial_x p \overset{(3.2e)}{=} \alpha_g \partial_x p. \qquad (3.12)$$

Equation (3.9d) can be obtained by following similar lines of reasoning as described above. In summary, the set of equations (3.9) is equivalent to (3.1) and (3.2). This completes the proof. ∎

3.2.2 Drift Flux Model (DFM)

In this section, we briefly recall the DFM. However, in this chapter, we only consider fluid flow dynamics across a spatial domain with unit and constant cross-section and, hence, accordingly adapt the model description.

The DFM can be obtained from the TFM via a slip relation of the form

$$v_g - v_\ell = \Phi\left(m_g, m_\ell, v_g \right), \qquad (3.13)$$

where m_g and m_ℓ are as introduced above. Since the slip relation (3.13) determines the coupling between the velocities of the two phases, only one momentum equation is required contrary to the two momentum equations in the TFM (3.1). Several models of the form (3.13) exist depending on the choice of the function Φ [56]. In the simplest case, without slip, $\Phi := 0$. Another case is the Zuber-Findlay relation [56]:

$$\Phi := \frac{(K-1)v_g + S}{K\alpha_\ell} \rightarrow v_g = K(\alpha_g v_g + \alpha_\ell v_\ell) + S, \qquad (3.14)$$

where K and S are flow-regime dependent parameters, which are assumed to be constant in this study.

Using the abbreviations $I_g := m_g v_g$ and $I_\ell := m_\ell v_\ell$, the governing equations for the DFM (across a unit, constant cross-section) are:

$$\partial_t m_g + \partial_x I_g = 0, \qquad (3.15a)$$

$$\partial_t m_\ell + \partial_x I_\ell = 0, \qquad (3.15b)$$

$$\partial_t \left(I_g + I_\ell \right) + \partial_x \left(I_g v_g + I_\ell v_\ell \right) = -\partial_x p + Q_g + Q_v \qquad (3.15c)$$

completed with closure equations (3.2a), (3.2e), (3.2f), (3.13) and gravitational effects Q_g and frictional effects Q_v defined by [57]:

$$Q_g = -g\left(m_g + m_\ell \right)\sin\theta, \qquad (3.16a)$$

$$Q_v = -\frac{32\mu_m(\alpha_g v_g + \alpha_\ell v_\ell)}{d^2}, \qquad (3.16b)$$

with gravitational constant g, space-dependent pipe inclination $\theta(x)$, mixture viscosity $\mu_m > 0$, and pipe diameter d.

Remark 3.2.4 *Similar to Lemma 3.2.2 and Lemma 3.2.3, the governing equations* (3.15) *associated with* $v := v_g = v_\ell$ *(DFM without slip), together with the closure equations* (3.2a), (3.2e), (3.2f) *and* (3.16), *upon elimination of auxiliary variables, can be rewritten as a system of PDEs with as many unknowns as equations. The reason behind such a model reformulation is similar to that for the TFM.*

The TFM can be adapted to behave exactly like the DFM if the term M_{ig} in (3.2d) is replaced with the term stated in the following theorem. For the proof, we refer to [56].

Theorem 3.2.5 *Under zero gravitational and frictional effects, the DFM* (3.15) *together with* (3.2a) *and* (3.13) *is equivalent to the TFM* (3.1) *with* (3.2a)–(3.2c), *and*

$$
M_{ig} = -\alpha_g \alpha_\ell \frac{\rho_g - \zeta \rho_\ell}{m_g + \zeta m_\ell} \partial_x p - \frac{m_g m_\ell}{m_g + \zeta m_\ell} \Big(v_\ell \partial_x v_\ell -
$$
$$
\zeta v_g \partial_x v_g + \mu_g \partial_x (m_g v_g) + \mu_\ell \partial_x (m_\ell v_\ell) \Big), \quad (3.17)
$$

with $\mu_g := \frac{\partial \Phi}{\partial m_g}$, $\mu_\ell := \frac{\partial \Phi}{\partial m_\ell}$, $\zeta := 1 - \frac{\partial \Phi}{\partial v_g}$.

Remark 3.2.6 *The equivalence of the DFM and the TFM can also be shown in the presence of gravitational and frictional effects; see [56], for further details.*

The model equivalence, stated above, will play a crucial role in drawing a conclusion about the behaviour of the Hamiltonian along the solutions of the DFM by using the theoretical analysis conducted for the TFM (see Section 3.5).

3.3 Dissipative Hamiltonian Formulations

Port-Hamiltonian systems have several useful properties for system analysis and control. Basic properties of port-Hamiltonian systems include passivity and compositionality. The port-Hamiltonian model formulation is appealing as it helps to characterize the energy exchange across the boundaries and thus accounts for the interaction between the system and the environment. Such a framework generalizes the classical Hamiltonian framework by the definition of

boundary ports. We restrict ourselves to port-Hamiltonian systems (with state-variable z) of the form

$$\partial_t z = \Big(\mathcal{J}(z) - \mathcal{R}(z) \Big) \delta_z \mathcal{H}(z),$$
$$\begin{pmatrix} f_\partial \\ e_\partial \end{pmatrix} = \mathcal{M} \begin{pmatrix} (\delta_z \mathcal{H}(z))(b) \\ (\delta_z \mathcal{H}(z))(a) \end{pmatrix}, \tag{3.18}$$

where $\mathcal{H}$ is the Hamiltonian functional, $\delta_z \mathcal{H}$ its variational derivative, and $\mathcal{M}$ is a state-dependent bijective mapping. Furthermore, for every z, $\mathcal{J}(z)$ is formally skew-adjoint with respect to the $\mathcal{L}^2$ inner product, i.e., for e_1, e_2 sufficiently smooth and zero at the boundary there holds

$$\int_\Omega e_1^T \Big(\mathcal{J}(z) e_2 \Big) dx + \int_\Omega e_2^T \Big(\mathcal{J}(z) e_1 \Big) dx = 0, \tag{3.19}$$

where Ω refers to the spatial domain, and $\mathcal{R}$ is formally self-adjoint with respect to the $\mathcal{L}^2$ inner product and positive semi-definite.

Finally, f_∂, e_∂ are the boundary ports, and these are normally split into inputs, outputs, and homogeneous boundary conditions, see, e.g., [89]. Mathematically, we can express the inputs u, outputs y and homogeneous boundary conditions as follows:

$$\begin{bmatrix} u \\ y \\ 0 \end{bmatrix} = W \begin{bmatrix} f_\partial \\ e_\partial \end{bmatrix}, \tag{3.20}$$

where W is a mapping matrix. Here, the first two lines indicate the inputs and outputs, and the third line indicates the homogenous boundary condition. Splitting the boundary port variables as shown above helps to relate the boundary ports of a port-Hamiltonian formulation to the boundary conditions of a PDE system such as the TFM and the DFM.

The dissipation inequality, which expresses that energy cannot be generated within the system, is a property that directly follows from the definition of a port-Hamiltonian system. In particular, ignoring the boundary conditions,

$$\begin{aligned}
\frac{d\mathcal{H}}{dt} &= \int_\Omega (\delta_z \mathcal{H}(z))^T \partial_t z \, dx \\
&= \int_\Omega (\delta_z \mathcal{H}(z))^T \Big((\mathcal{J}(z) - \mathcal{R}(z)) \delta_z \mathcal{H}(z) \Big) dx \\
&= \int_\Omega (\delta_z \mathcal{H}(z))^T (-\mathcal{R}(z)) \delta_z \mathcal{H}(z) \, dx \le 0,
\end{aligned} \tag{3.21}$$

where the last inequality is due to the positive semi-definite nature of $\mathcal{R}(z)$.

Thus, $\mathcal{R}$ is the dissipative component of the system. In the presence of boundary conditions, the behaviour of the Hamiltonian along the solutions of the mathematical model is governed by the following balance equation:

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = \int_{\Omega} (\delta_z \mathcal{H}(z))^T \Big(-\mathcal{R}(z)\delta_z \mathcal{H}(z) \Big) \, \mathrm{d}x + \text{b.t.}, \tag{3.22}$$

where b.t. denotes the boundary terms. Normally f_∂, e_∂ are chosen such that the boundary terms equal $\langle f_\partial, e_\partial \rangle$ w.r.t. some inner product. In our case, this will be the standard inner product on the Euclidean space. Associated to the operators $\mathcal{J}$ and $\mathcal{R}$, we can identify an underlying geometric object called Stokes-Dirac structure. This is crucial as port-Hamiltonian systems can be defined with respect to these infinite-dimensional Stokes-Dirac structures [53]. Often, this structure is only associated to $\mathcal{J}$. This geometric object yields a manner to describe the boundary port variables, i.e., f_∂ and e_∂, see (3.18).

We first introduce (dissipative) Hamiltonian representations, i.e., without boundary effects for the mathematical models under consideration. The resulting formally skew-adjoint operators and formally self-adjoint operators are used as a tool to derive a non-canonical Stokes-Dirac structure, and hence the boundary port variables.

In the models discussed in Section 3.2, the Hamiltonian is dependent on the kinetic, gravitational potential and internal energy. To derive the internal energy of the system, consider the following remark.

Remark 3.3.1 *The internal energy $u_i, i \in \{\ell, g\}$, can be interpreted as the energy causing the expansion of the i-th compressed phase or compression of the i-th expanded phase. In order to derive this energy component, the Gibbs relation [41] under barotropic and isentropic flow considerations for an infinitesimal part of the phase is used, i.e.,*

$$\rho_i^2 \mathrm{d}u_i = p\mathrm{d}\rho_i, \qquad i \in \{\ell, g\}.$$

Using (3.2e)–(3.2f) *and integrating the above equation leads to*

$$u_\ell = -\frac{p_{\ell 0}}{\rho_\ell} + c_\ell^2 \ln \rho_\ell + \frac{\rho_{\ell 0} c_\ell^2}{\rho_\ell} + K_1, \tag{3.23a}$$

$$u_g = c_g^2 \ln \rho_g + K_2, \tag{3.23b}$$

where K_1 and K_2 are the integration constants.

Considering the total energy of the system (neglecting the gravitational potential energy), we define a candidate for the Hamiltonian as follows:

$$\mathcal{H} := \int_{\Omega} \left(\alpha_{\mathrm{g}} \rho_{\mathrm{g}} \frac{v_{\mathrm{g}}^2}{2} + \alpha_{\ell} \rho_{\ell} \frac{v_{\ell}^2}{2} + \alpha_{\mathrm{g}} \rho_{\mathrm{g}} u_{\mathrm{g}} + \alpha_{\ell} \rho_{\ell} u_{\ell} \right) \mathrm{d}x, \tag{3.24}$$

where $\Omega = [a, b]$ refers to the spatial domain.

Inserting (3.23) into (3.24), the Hamiltonian for a flow across a (unit) constant cross-section takes the following form:

$$\mathcal{H} := \int_{\Omega} \Big(\alpha_{\mathrm{g}} \rho_{\mathrm{g}} \frac{v_{\mathrm{g}}^2}{2} + \alpha_{\ell} \rho_{\ell} \frac{v_{\ell}^2}{2} + \alpha_{\mathrm{g}} \rho_{\mathrm{g}} \big(c_{\mathrm{g}}^2 \ln \rho_{\mathrm{g}} + K_2 \big) +$$
$$\alpha_{\ell} \rho_{\ell} \big(c_{\ell}^2 \ln \rho_{\ell} + K_1 \big) + \alpha_{\ell} \big(c_{\ell}^2 \rho_{\ell 0} - p_{\ell 0} \big) \Big) \mathrm{d}x. \tag{3.25}$$

It should be noted that when $\rho_i \to 0$, $\rho_i \ln \rho_i \to 0$. The term $\rho_i \ln \rho_i$ is bounded from below, i.e., $\rho_i \ln \rho_i \geq -1/e$. So, the Hamiltonian (3.25) is bounded from below. Due to the high bulk modulus of the liquid phase, we usually have $\rho_{\ell 0} c_{\ell}^2 \gg p_{\ell 0}$ [57]; therefore, the positivity of the Hamiltonian (3.25) can be ensured by appropriately choosing K_1 and K_2 or even adding some constants to the Hamiltonian. For simplicity, we set $K_1 := 0$ and $K_2 := 0$ henceforth.

Remark 3.3.2 *The discussion in the above paragraph is reasonable from a physical perspective. However, numerically, solutions of the TFM and DFM may not be guaranteed to have non-negative density and non-negative void fractions.*

3.3.1 Dissipative Hamiltonian Formulation for the TFM

We now present dissipative Hamiltonian formulations for the TFM under two different choices of state-variables.

Theorem 3.3.3 *The governing equations* (3.4) *can be written in the dissipative Hamiltonian formulation*

$$\partial_t z = \big(\bar{\mathcal{J}}_T - \bar{\mathcal{R}}_T(z) \big) \delta_z \mathcal{H}(z) \tag{3.26}$$

with $z := [m_{\mathrm{g}}, \, m_{\ell}, \, v_{\mathrm{g}}, \, v_{\ell}]^T$, *the Hamiltonian functional* $\mathcal{H}$ *as in* (3.25)*, and where*

$$\bar{\mathcal{J}}_T = - \begin{bmatrix} \boldsymbol{0} & I \\ I & \boldsymbol{0} \end{bmatrix} \frac{\partial}{\partial x}$$

is a formally skew-adjoint operator with respect to the $\mathscr{L}^2$ inner product, and

$$\bar{\mathscr{R}}_T(z) = \begin{bmatrix} 0 & 0 \\ 0 & \tau_r \end{bmatrix}$$

is a symmetric and positive semi-definite matrix, with $\tau_r = \begin{bmatrix} \dfrac{b_g^M}{m_g^2} & -\dfrac{b_g^M}{m_g m_\ell} \\ -\dfrac{b_g^M}{m_g m_\ell} & \dfrac{b_g^M}{m_\ell^2} \end{bmatrix}$,

and 0 and I a 2×2 zero and identity matrix, respectively.

Proof: The Hamiltonian functional (3.25) in terms of this set of state variables z is re-written as follows:

$$\mathscr{H}(m_g, m_\ell, v_g, v_\ell) := \int_\Omega m_g \frac{v_g^2}{2} + m_\ell \frac{v_\ell^2}{2} + m_g c_g^2 \ln\left(\frac{p}{c_g^2}\right) +$$
$$m_\ell c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right) + (1-\alpha_g)\beta \, dx, \quad (3.27)$$

where p and α_g can be replaced by the relations (3.5) and (3.6), respectively. The variational derivative of the Hamiltonian is given by:

$$\frac{\delta\mathscr{H}}{\delta m_g} = \frac{v_g^2}{2} + c_g^2 \ln\left(\frac{p}{c_g^2}\right) + c_g^2, \quad \frac{\delta\mathscr{H}}{\delta v_g} = m_g v_g,$$
$$\frac{\delta\mathscr{H}}{\delta m_\ell} = \frac{v_\ell^2}{2} + c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right) + c_\ell^2, \quad \frac{\delta\mathscr{H}}{\delta v_\ell} = m_\ell v_\ell.$$

It is straightforward to show that the system (3.26) is equivalent to (3.4) and, hence, we omit its detailed derivation here.

Moreover, $\bar{\mathscr{R}}_T$ is symmetric and positive semi-definite due to the positivity of b_g^M. It is straightforward to prove the formal skew-adjointness of $\bar{\mathscr{J}}_T$ with respect to the $\mathscr{L}^2$ inner product and, hence, we omit such proof here. ∎

Remark 3.3.4 *We can observe that τ_r will not be well defined if the masses in any of the phases, i.e., m_g or m_ℓ, become zero during the evolution of the states. This is basically a by-product of the model reformulation (3.4), which will not hold if m_g or m_ℓ become zero. At this stage, it is also worth mentioning that the masses of*

any of the phases will not become zero due to Assumption 1. The aforementioned observation serves as one of the motivations to consider other suitable choices of state variables, for instance, $m_g, m_\ell, m_g v_g$ and $m_\ell v_\ell$, and the corresponding model reformulations, for instance, as in Lemma 3.2.3.

The state vector in the previous theorem was composed of non-conservative variables. It is well known that any bijective map acting on the state variables still retains the (dissipative) Hamiltonian structure. However, such bijective map may not be always attainable in the scope of non-linear problems such as multi-phase flow models. Hence, as a first step, we define the state vector in terms of conservative variables and demonstrate the existence/non-existence of a corresponding dissipative Hamiltonian formalism for the TFM. Moreover, we assess the suitability of these conservative variables in the scope of dealing with vanishing phases.

Remark 3.3.5 *Deriving conditions under which bijective map would be attainable is outside the scope of this work.*

Theorem 3.3.6 *The governing equations* (3.1) *together with the closure equations* (3.2), *equivalent to* (3.9), *can be written in the following dissipative Hamiltonian form:*

$$\partial_t q = \left(\mathcal{J}_T(q) - \mathcal{R}_T \right) \delta_q \mathcal{H}(q) \tag{3.28}$$

with $q = [q_1, q_2, q_3, q_4]^T := [m_g, m_\ell, I_g, I_\ell]^T$, the Hamiltonian functional $\mathcal{H}$ as in (3.25), *and where*

$$\mathcal{J}_T(q) = - \begin{bmatrix} 0 & 0 & \partial_x(q_1 \cdot) & 0 \\ 0 & 0 & 0 & \partial_x(q_2 \cdot) \\ q_1 \partial_x(\cdot) & 0 & \partial_x(q_3 \cdot) + q_3 \partial_x(\cdot) & 0 \\ 0 & q_2 \partial_x(\cdot) & 0 & \partial_x(q_4 \cdot) + q_4 \partial_x(\cdot) \end{bmatrix}$$

is a formally skew-adjoint operator with respect to the $\mathcal{L}^2$ inner product, and

$$\mathcal{R}_T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & b_g^M & -b_g^M \\ 0 & 0 & -b_g^M & b_g^M \end{bmatrix}$$

is a symmetric and positive semi-definite matrix.

Remark 3.3.7 *Here, $\mathscr{J}_T(q)\delta_q\mathscr{H}(q)$ is not the standard (multiplicative) product of $\mathscr{J}_T(q)$ and $\delta_q\mathscr{H}(q)$. Instead, $\mathscr{J}_T(q)\delta_q\mathscr{H}(q)$ should be better viewed as composition of $\mathscr{J}_T(q)$ and $\delta_q\mathscr{H}(q)$, i.e., $\mathscr{J}_T(q)\circ\delta_q\mathscr{H}(q)$. For instance, $k\partial_x(\cdot)l :=$ $k\partial_x(\cdot)\circ l$. Such a composition operation is given by $k\partial_x(\cdot)\circ l = k\partial_x(l)$. Similarly, $\partial_x(k\cdot)l := \partial_x(k\cdot)\circ l = \partial_x(kl)$, where k and l are any functions such as q_1, q_2, q_3 and q_4. Moreover, $\partial_x(k\cdot) + k\partial_x(\cdot)$ is additive in standard sense, i.e., $\big(\partial_x(k\cdot) + k\partial_x(\cdot)\big)\circ l$ $= \partial_x(k\cdot)\circ l + k\partial_x(\cdot)\circ l$.*

Proof: The Hamiltonian (3.25) in terms of q_1, q_2, q_3 and q_4 is re-written as follows:

$$\mathscr{H}(q_1,q_2,q_3,q_4) := \int_\Omega \frac{q_3^2}{2q_1} + \frac{q_4^2}{2q_2} + q_1 c_g^2 \ln\left(\frac{p}{c_g^2}\right)$$

$$+ q_2 c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right) + (1-\alpha_g)\beta \; dx, \quad (3.29)$$

where p and α_g can be replaced by the relations (3.5) and (3.6), respectively.

The variational derivatives are:

$$\frac{\delta\mathscr{H}}{\delta q_1} = -\frac{1}{2}\frac{q_3^2}{q_1^2} + c_g^2\ln\left(\frac{p}{c_g^2}\right) + c_g^2, \quad \frac{\delta\mathscr{H}}{\delta q_3} = \frac{q_3}{q_1},$$

$$\frac{\delta\mathscr{H}}{\delta q_2} = -\frac{1}{2}\frac{q_4^2}{q_2^2} + c_\ell^2\ln\left(\frac{p+\beta}{c_\ell^2}\right) + c_\ell^2, \quad \frac{\delta\mathscr{H}}{\delta q_4} = \frac{q_4}{q_2}.$$

Next, we prove the claim equation by equation. The first line of (3.28) reads

$$\partial_t m_g = -\partial_x\left(m_g\left(\frac{m_g v_g}{m_g}\right)\right) = -\partial_x(m_g v_g). \quad (3.30)$$

Similarly, the second line reads

$$\partial_t m_\ell = -\partial_x\left(m_\ell\left(\frac{m_\ell v_\ell}{m_\ell}\right)\right) = -\partial_x(m_\ell v_\ell). \quad (3.31)$$

The third line yields

$$\partial_t(m_g v_g) = -m_g\partial_x\left(\frac{\delta\mathscr{H}}{\delta q_1}\right) - \partial_x\left(m_g v_g\left(\frac{\delta\mathscr{H}}{\delta q_3}\right)\right) - m_g v_g\partial_x\left(\frac{\delta\mathscr{H}}{\delta q_3}\right) - b_g^M\left(\frac{\delta\mathscr{H}}{\delta q_3}\right) + b_g^M\left(\frac{\delta\mathscr{H}}{\delta q_4}\right).$$

By substituting the variational derivatives into the above equation, we have

$$\partial_t(m_g v_g) = -m_g\partial_x\left(-\frac{1}{2}v_g^2 + c_g^2\ln\left(\frac{p}{c_g^2}\right) + c_g^2\right) - \partial_x(m_g v_g^2) - m_g v_g\partial_x v_g - b_g^M(v_g - v_\ell).$$

Simplifying the above equation leads to:

$$\partial_t(m_g v_g) = -\partial_x(m_g v_g^2) - m_g \partial_x\left(c_g^2 \ln\left(\frac{p}{c_g^2}\right)\right) - b_g^M(v_g - v_\ell).\qquad(3.32)$$

The fourth line yields of (3.28) reads

$$\partial_t(m_\ell v_\ell) = -m_\ell \partial_x\left(\frac{\delta \mathcal{H}}{\delta q_2}\right) - \partial_x\left(m_\ell v_\ell\left(\frac{\delta \mathcal{H}}{\delta q_4}\right)\right) - m_\ell v_\ell \partial_x\left(\frac{\delta \mathcal{H}}{\delta q_4}\right) + b_g^M\left(\frac{\delta \mathcal{H}}{\delta q_3}\right) - b_g^M\left(\frac{\delta \mathcal{H}}{\delta q_4}\right).$$

On substituting the variational derivatives into the above equation, we have

$$\partial_t(m_\ell v_\ell) = -m_\ell \partial_x\left(-\frac{1}{2}v_\ell^2 + c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right) + c_\ell^2\right) - \partial_x(m_\ell v_\ell^2) - m_\ell v_\ell \partial_x v_\ell + b_g^M(v_g - v_\ell).$$

Simplifying the above equation leads to:

$$\partial_t(m_\ell v_\ell) = -\partial_x(m_\ell v_\ell^2) - m_\ell \partial_x\left(c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right)\right) + b_g^M(v_g - v_\ell).\qquad(3.33)$$

The claim of the theorem follows by observing that (3.30), (3.31), (3.32) and (3.33) are identical to (3.9a), (3.9b), (3.9c) and (3.9d), respectively.

Alternatively, we can also argue that the TFM exhibits similarities in structure with the model presented in [173], where the Hamiltonian structure was discussed for single-phase dynamics. The TFM with $b_g^M = 0$ can be viewed as two separately existing phases. The contributions due to the non-zero b_g^M enter into the dissipation matrix $\mathscr{R}_T$.

The proof of the symmetric and positive semi-definite nature of $\mathscr{R}_T$ is straightforward. The operator $\mathscr{J}_T$ is formally skew-adjoint (with respect to the $\mathscr{L}^2$ inner product). To prove formal skew-adjointness of $\mathscr{J}_T$, we check whether $\langle \mathbf{e^1}, \mathscr{J}_T \mathbf{e^2} \rangle_{\mathscr{L}^2(\Omega)} + \langle \mathscr{J}_T \mathbf{e^1}, \mathbf{e^2} \rangle_{\mathscr{L}^2(\Omega)} = 0$ holds for smooth $\mathbf{e^1}, \mathbf{e^2}$ which are zero at the boundary, where we define $\mathbf{e^i} = (e_1^i, e_2^i, e_3^i, e_4^i)^T$. Here, the variable e_j^i refers to the j-th element of $\mathbf{e^i}$. $\mathscr{J}_T$ is formally skew-adjoint with respect to the $\mathscr{L}^2$ inner

product as

$$
-\langle \mathbf{e^1}, \mathcal{J}_T \mathbf{e^2} \rangle_{\mathscr{L}^2(\Omega)} - \langle \mathcal{J}_T \mathbf{e^1}, \mathbf{e^2} \rangle_{\mathscr{L}^2(\Omega)} =
$$

$$
\int_{\Omega} \quad e_1^1 \partial_x(q_1 e_3^2) + q_1 e_3^2 \partial_x e_1^1 + e_1^2 \partial_x(q_1 e_3^1) + q_1 e_3^1 \partial_x e_1^2 +
$$
$$
e_2^1 \partial_x(q_2 e_4^2) + q_2 e_4^2 \partial_x e_2^1 + e_2^2 \partial_x(q_2 e_4^1) + q_2 e_4^1 \partial_x e_2^2 +
$$
$$
e_3^1 \left[\partial_x(q_3 e_3^2) + q_3 \partial_x e_3^2 \right] + e_3^2 \left[\partial_x(q_3 e_3^1) + q_3 \partial_x e_3^1 \right] +
$$
$$
e_4^1 \left[\partial_x(q_4 e_4^2) + q_4 \partial_x e_4^2 \right] + e_4^2 \left[\partial_x(q_4 e_4^1) + q_4 \partial_x e_4^1 \right] dx =
$$

$$
\left(\begin{bmatrix} e_1^1 & e_2^1 & e_3^1 & e_4^1 \end{bmatrix} \underbrace{\begin{bmatrix} 0 & 0 & q_1 & 0 \\ 0 & 0 & 0 & q_2 \\ q_1 & 0 & 2q_3 & 0 \\ 0 & q_2 & 0 & 2q_4 \end{bmatrix}}_{Q_T} \begin{bmatrix} e_1^2 \\ e_2^2 \\ e_3^2 \\ e_4^2 \end{bmatrix} \right) \Big|_a^b,
$$

which vanishes under our assumptions on the boundary conditions.   ∎

Remark 3.3.8 *The matrix $\mathscr{R}_T$ and the skew-adjoint operator $\mathcal{J}_T(q)$ are well defined. However, the computation of variational derivatives is not defined for vanishing phases. Accordingly, q_1 and q_3 should not become zero, and this is guaranteed by Assumption 1.*

Remark 3.3.9 *On the one hand, the formally skew-adjoint operator in Theorem 3.3.3 is state independent. On the other hand, the formally skew-adjoint operator in Theorem 3.3.6 is state dependent. Conversely, $\tilde{\mathscr{R}}_T$ is a state-dependent matrix in Theorem 3.3.3 and $\mathscr{R}_T$ is a state-independent matrix in Theorem 3.3.6. In the absence of dissipation terms, a constant Stokes-Dirac structure [116] can be defined for the formulation in Theorem 3.3.3 (unlike Theorem 3.3.6, where the Stokes-Dirac structure is non-constant).*

3.3.2   Dissipative Hamiltonian Formulation for the DFM

So far, we focused on the dissipative Hamiltonian representations for the TFM under two different variants of generalized state coordinates. We will now deal with the DFM under gravitational and frictional effects, and present a corresponding dissipative Hamiltonian formulation. For the DFM, we focus only on

a case in which there is no slip between the phases, i.e., $v := v_g = v_\ell$ (the reason for adopting this no-slip assumption is provided in Section 3.5). Since gravitation is considered, the gravitational potential energy needs to be added to the Hamiltonian. The Hamiltonian now takes the following form:

$$\mathcal{H}_D(m_g, m_\ell, v) = \int_\Omega m_g \frac{v^2}{2} + m_\ell \frac{v^2}{2} + m_\ell c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right) + m_g c_g^2 \ln\left(\frac{p}{c_g^2}\right) +$$

$$\alpha_\ell \beta + (m_g + m_\ell)\left(\int_a^x g\sin(\theta(\xi))d\xi\right)dx. \quad (3.34)$$

Using the above candidate Hamiltonian function $\mathcal{H}_D$, a dissipative Hamiltonian representation of a special case of the DFM is shown below.

Theorem 3.3.10 *The governing equations (3.15) together with $v := v_g = v_\ell$ (case of no slip), the closure equations (3.2a), (3.2e), (3.2f) and (3.16) can be written in dissipative Hamiltonian form as follows:*

$$\partial_t z_D = \left(\mathcal{J}_D(z_D) - \mathcal{R}_D(z_D)\right)\delta_{z_D}\mathcal{H}_D(z_D) \quad (3.35)$$

with $z_D := [m_g, m_\ell, v]^T$, the Hamiltonian functional $\mathcal{H}_D$ as in (3.34), and where

$$\mathcal{J}_D(z_D) = -\begin{bmatrix} 0 & 0 & \partial_x\left(\frac{m_g}{m_g+m_\ell}\cdot\right) \\ 0 & 0 & \partial_x\left(\frac{m_\ell}{m_g+m_\ell}\cdot\right) \\ \frac{m_g}{m_g+m_\ell}\partial_x(\cdot) & \frac{m_\ell}{m_g+m_\ell}\partial_x(\cdot) & 0 \end{bmatrix}$$

is a formally skew-adjoint operator with respect to the $\mathcal{L}^2$ inner product, and

$$\mathcal{R}_D(z_D) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{32\mu_m}{d^2(m_g+m_\ell)^2} \end{bmatrix}$$

is a symmetric and positive semi-definite matrix.

Proof: First note that, using (3.15a) and (3.15b), the left-hand side of equation (3.15c) can be rewritten as

$$(m_g + m_\ell)\partial_t v + v\partial_t(m_g + m_\ell) + \partial_x\left((m_g + m_\ell)v^2\right)$$
$$= (m_g + m_\ell)\left(\partial_t v + \partial_x\left(\frac{v^2}{2}\right)\right).$$

Thus, instead of (3.15c) we can also consider

$$\partial_t v + \partial_x \left(\frac{v^2}{2} \right) = \frac{1}{m_g + m_\ell} \left(-\partial_x p + Q_g + Q_v \right). \tag{3.36}$$

The variational derivatives of $\mathcal{H}_D$ are given by:

$$\frac{\delta \mathcal{H}_D}{\delta m_g} = c_g^2 \ln \left(\frac{p}{c_g^2} \right) + \frac{v^2}{2} + c_g^2 + \int_a^x g \sin(\theta(\xi)) d\xi,$$

$$\frac{\delta \mathcal{H}_D}{\delta m_\ell} = c_\ell^2 \ln \left(\frac{p + \beta}{c_\ell^2} \right) + \frac{v^2}{2} + c_\ell^2 + \int_a^x g \sin(\theta(\xi)) d\xi,$$

$$\frac{\delta \mathcal{H}_D}{\delta v} = \left(m_g + m_\ell \right) v.$$

Next, we prove the claim equation by equation. The first line of (3.35) reads

$$\partial_t m_g = -\partial_x \left(\frac{m_g}{m_g + m_\ell} \left(m_g + m_\ell \right) v \right) = -\partial_x \left(m_g v \right). \tag{3.37}$$

Similarly, the second line is

$$\partial_t m_\ell = -\partial_x \left(\frac{m_\ell}{m_g + m_\ell} \left(m_g + m_\ell \right) v \right) = -\partial_x \left(m_\ell v \right). \tag{3.38}$$

Let us introduce a short-hand notation $G = \int_a^x g \sin(\theta(\xi)) d\xi$. Then, the third line yields

$$\begin{aligned}
\partial_t v &= -\frac{m_g}{m_g + m_\ell} \partial_x \left(c_g^2 \ln \left(\frac{p}{c_g^2} \right) + \frac{v^2}{2} + c_g^2 + G \right) \\
&\quad - \frac{m_\ell}{m_g + m_\ell} \partial_x \left(c_\ell^2 \ln \left(\frac{p + \beta}{c_\ell^2} \right) + \frac{v^2}{2} + c_\ell^2 + G \right) \\
&\quad - \frac{32 \mu_m}{d^2 \left(m_g + m_\ell \right)^2} \left(m_g + m_\ell \right) v \\
&= -\partial_x \left(\frac{v^2}{2} \right) - \frac{1}{\left(m_g + m_\ell \right)} \left(\partial_x p + Q_g + Q_v \right).
\end{aligned} \tag{3.39}$$

The claim of the theorem follows by observing that (3.37), (3.38), and (3.39) are identical to (3.15a), (3.15b), and (3.36), respectively.

The symmetric and positive semi-definite nature of $\mathcal{R}_D$ follows immediately from the positivity of μ_m. The formal skew-adjointness of $\mathcal{J}_D$ essentially follows from integration by parts and neglecting the boundary conditions. The operator $\mathcal{J}_D$ contains terms similar to the skew-adjoint operator $\mathcal{J}_T$, the formal skew-adjointness of which was discussed extensively in the proof of Theorem 3.3.6. For the sake of brevity, we refer the reader to follow similar lines of reasoning to show the formal skew-adjointness of $\mathcal{J}_D$, and just present the final result. $\mathcal{J}_D$ is formally skew-adjoint with respect to the $\mathcal{L}^2$ inner product because

$$-\langle \mathbf{e^1}, \mathcal{J}_D \mathbf{e^2} \rangle_{\mathcal{L}^2(\Omega)} - \langle \mathcal{J}_D \mathbf{e^1}, \mathbf{e^2} \rangle_{\mathcal{L}^2(\Omega)} = \left(\begin{bmatrix} e_1^1 & e_2^1 & e_3^1 \end{bmatrix} \underbrace{\begin{bmatrix} 0 & 0 & \frac{m_g}{m_g + m_\ell} \\ 0 & 0 & \frac{m_\ell}{m_g + m_\ell} \\ \frac{m_g}{m_g + m_\ell} & \frac{m_\ell}{m_g + m_\ell} & 0 \end{bmatrix}}_{Q_D} \begin{bmatrix} e_1^2 \\ e_2^2 \\ e_3^2 \end{bmatrix} \right) \Big|_a^b,$$

(3.40)

which vanishes under smooth $\mathbf{e^1} = [e_1^1,\ e_2^1,\ e_3^1]^T, \mathbf{e^2} = [e_1^2,\ e_2^2,\ e_3^2]^T$ that are zero at the boundary. ∎

Remark 3.3.11 *Similar to different port-Hamiltonian realizations of the TFM under different choices of generalized state-variables, the DFM (without slip) can also be written in other port-Hamiltonian formulation(s) depending on the choice of state-variables.*

3.4   Geometrical properties of the system

We now define a geometric structure, a generalization of symplectic and Poisson structures, called a Dirac structure.

Definition 3.4.1 *(Dirac Structures) [53, 101] Consider $\mathscr{F}$ and $\mathscr{E}$ as real Hilbert spaces which are isometrically isomorphic. The subspace $\mathscr{D} \subset \mathscr{F} \times \mathscr{E}$ is a Dirac structure if $\mathscr{D} = \mathscr{D}^\perp$, where $\mathscr{D}^\perp$ denotes the orthogonal complement which is defined as*

$$\mathscr{D}^\perp := \{ (\tilde{\mathbf{f}}, \tilde{\mathbf{e}}) \in \mathscr{F} \times \mathscr{E} \mid \ll (\tilde{\mathbf{f}}, \tilde{\mathbf{e}}), (\mathbf{f}, \mathbf{e}) \gg = 0 \quad \forall (\mathbf{f}, \mathbf{e}) \in \mathscr{D} \}. \tag{3.41}$$

Here, $\ll (\tilde{\mathbf{f}}, \tilde{\mathbf{e}}), (\mathbf{f}, \mathbf{e}) \gg$ is defined as follows:

$$\ll (\tilde{\mathbf{f}}, \tilde{\mathbf{e}}), (\mathbf{f}, \mathbf{e}) \gg := \langle \tilde{\mathbf{f}} \mid \mathbf{e} \rangle + \langle \mathbf{f} \mid \tilde{\mathbf{e}} \rangle, \tag{3.42}$$

where the notation $\langle \mathbf{f} \mid \mathbf{e} \rangle$ indicates a non-degenerate bilinear form defined on the bond space $\mathscr{B} = \mathscr{F} \times \mathscr{E}$.

This geometric structure relates the composing elements of a system in a power-preserving manner [140]. Such geometric structures often have a compositionality property [6, 99, 140].

A Stokes-Dirac structure is a special case of a Dirac structure, where also boundary ports appear and Stokes' Theorem plays a role, see [53, 101]. This geometric structure is defined next.

Definition 3.4.2 *(Stokes-Dirac structures) [53, 101] Let $\mathscr{F}_\Omega$ and $\mathscr{E}_\Omega$, respectively, denote the subspace of flow variables and effort variables defined on the domain Ω. Also, Let $\mathscr{F}_\partial$ and $\mathscr{E}_\partial$, respectively, denote the subspace of flow variables and effort variables defined on the boundary $\partial\Omega$. Furthermore, consider the real Hilbert space of flow variables $\mathscr{F}_S$ and of effort variables $\mathscr{E}_S$, where $\mathscr{F}_S = \mathscr{F}_\Omega \times \mathscr{F}_\partial$, and $\mathscr{E}_S = \mathscr{E}_\Omega \times \mathscr{E}_\partial$, respectively. Moreover, consider that the Hilbert space $\mathscr{E}_S$ is the dual of $\mathscr{F}_S$ [53, 194] or consider that the Hilbert spaces $\mathscr{E}_S$ and $\mathscr{F}_S$ are isometrically isomorphic to each other [101]. Finally, we endow the bond space $\mathscr{B}_S = \mathscr{F}_S \times \mathscr{E}_S$ with the following pairing:*

$$\ll (f, f_\partial, e, e_\partial), (\tilde{f}, \tilde{f}_\partial, \tilde{e}, \tilde{e}_\partial) \gg = \langle f \mid \tilde{e} \rangle + \langle e \mid \tilde{f} \rangle \pm \langle f_\partial \mid \tilde{e}_\partial \rangle \pm \langle e_\partial \mid \tilde{f}_\partial \rangle. \qquad (3.43)$$

Then, the subset $\mathscr{D}_S$ of $\mathscr{B}_S$ is a Stokes-Dirac structure with respect to the non-degenerate bilinear form (3.43) if $\mathscr{D}_S = \mathscr{D}_S^\perp$, where $\mathscr{D}_S^\perp$ denotes the orthogonal complement of $\mathscr{D}_S$ and is defined like (3.41).

Remark 3.4.3 *The usage of + or − before the boundary terms in the above equation depends on the adopted sign convention.*

For a (f, e) element of a Dirac structure (or Stokes-Dirac structure), it is easy to see that $\langle f \mid e \rangle = 0$, and thus there is a close relation to (formally) skew-adjoint operators, see also (3.19). However, if $f = \mathscr{J}e$ for all $(f, e) \in D$, and J is formally skew-adjoint, then $\mathscr{D} \subset \mathscr{D}^\perp$ (or $\mathscr{D}_S \subset \mathscr{D}_S^\perp$). To make such a $\mathscr{D}$ into a Dirac/Stokes-Dirac structure, it is required that $\mathscr{D} = \mathscr{D}^\perp$ (or $\mathscr{D}_S \subset \mathscr{D}_S^\perp$) holds. The formally skew-adjoint part of a port-Hamiltonian system will form the foundation of the associated Dirac or Stokes-Dirac structure, as we will show as well.

The non-linearity encoded within the Hamiltonian along with a linear Dirac or Stokes-Dirac structure constitutes a favorable representation of PDEs. Such a structure facilitates the analysis of non-linear systems as the linearity of the Dirac or Stokes-Dirac structure can be exploited to assess system behaviour. These geometric structures can also be used to formulate boundary control systems [101].

Before delving into the analysis of the underlying geometrical properties in the scope of two-phase flow models of interest, we recall a well-known result

in the context of defining Stokes-Dirac structures for linear infinite-dimensional port-Hamiltonian realizations, and parametrizing the boundary port variables.

Definition 3.4.4 *[101] Consider the formally skew-symmetric differential operator $\mathscr{J} = P_1 \partial_x e$ such that $P_1 = P_1^T$. Also, consider the boundary trace operator $\tau : H^1(\Omega)^n \to \mathbb{R}^{2n}$ defined in the following way:*

$$\tau(e) = \begin{pmatrix} e(b) \\ e(a) \end{pmatrix}.$$

Next, choose the flow space $\mathscr{F}$ and effort space $\mathscr{E}$ endowed with the natural inner product in the following manner:

$$\mathscr{F} = \mathscr{E} = \mathscr{L}^2(\Omega)^n \times \mathbb{R}^n,$$

and consider the bond space $\mathscr{B}$ as $\mathscr{F} \times \mathscr{E}$. Furthermore, endow $\mathscr{B}$ with the following canonical symmetric pairing:

$$\ll (f, f_\partial, e, e_\partial), (\tilde{f}, \tilde{f}_\partial, \tilde{e}, \tilde{e}_\partial) \gg = \langle f, \tilde{e} \rangle_{\mathscr{L}^2(\Omega)} + \langle e, \tilde{f} \rangle_{\mathscr{L}^2(\Omega)} - \langle f_\partial, \tilde{e}_\partial \rangle_{\mathbb{R}} - \langle e_\partial, \tilde{f}_\partial \rangle_{\mathbb{R}}. \quad (3.44)$$

Then, the subspace $\mathscr{D}$ of $\mathscr{B}$ defined by

$$\mathscr{D} = \left\{ \begin{bmatrix} f \\ f_\partial \\ e \\ e_\partial \end{bmatrix} \in \mathscr{B} \mid e \in H^1(\Omega)^n, \mathscr{J}e = f, \begin{bmatrix} f_\partial \\ e_\partial \end{bmatrix} = R_{ext}\tau(e) \right\}, \quad (3.45)$$

with the matrix $R_{ext} \in \mathbb{R}^{2n \times 2n}$ given by

$$R_{ext} = \frac{1}{\sqrt{2}} \begin{bmatrix} Q & -Q \\ I & I \end{bmatrix},$$

is a Dirac structure with respect to the bilinear form (3.44). The Dirac structure $\mathscr{D}$ is defined by the fact that $\mathscr{D} = \mathscr{D}^\perp$, where $\mathscr{D}^\perp$ denotes the orthogonal complement

$$\mathscr{D}^\perp := \{(\tilde{\mathbf{f}}, \tilde{\mathbf{e}}) \in \mathscr{F} \times \mathscr{E} \mid \ll (\tilde{\mathbf{f}}, \tilde{\mathbf{e}}), (\mathbf{f}, \mathbf{e}) \gg = 0 \quad \forall (\mathbf{f}, \mathbf{e}) \in \mathscr{D}\}. \quad (3.46)$$

Remark 3.4.5 *The above definition also holds for higher-order (state-independent) differential operators with constant coefficients; see [101, 194] for further details.*

In the existing results [53,101,194], the effort variables **e** belong to the function class $H^1(\Omega)$ since they dealt with *constant* matrix differential operators of first order (see Definition 3.4.4). Given the state-dependent nature of the first-order skew-adjoint operators in (3.28) and (3.35) (unlike in [101]), a combination of the states and the effort variables have to belong to the function class $H^1(\Omega)$ or suitable conditions have to be imposed on the state variables in order to have effort variables belonging to the function class $H^1(\Omega)$ (see Theorems 3.4.9 and 3.4.11). Boundary port-variables have been parametrized in [101] using the trace operators; see $\tau(e)$ in Definition 3.4.4. However, in [101], such a parametrization is limited to the case of a non-singular matrix Q (synonymous to Q_T in Theorem 3.3.6) arising in linear problems with *state-independent* operators. To the best of our knowledge, the work [194] is the only work in the scope of parametrization of boundary port-variables for a singular matrix Q (synonymous to Q_D in Theorem 3.3.10), thereby enlarging the class of systems that can be dealt with. Villegas in [194] demonstrated the approach to define the non-degenerate bilinear form under singular Q and consequently modified the definition of the boundary port-variables. However, [194] is limited to the setting of *state-independent* Stokes-Dirac structures. In this work, we strive to extend the definition of boundary port-variables to eventually obtain *state-dependent* Stokes-Dirac structures with boundary ports for non-linear problems with non-quadratic Hamiltonian functional. It should be mentioned that the authors in [184] have also considered state-dependent Stokes-Dirac structures for problems (for instance, ideal isentropic fluid) with non-quadratic Hamiltonian functional by using a differential geometric viewpoint. We, however, use the operator-theoretic viewpoint in the consideration of such geometric structures in the scope of the two-phase flow models.

Remark 3.4.6 *The parametrization of boundary port variables is non-trivial for underlying state-dependent Stokes-Dirac structure representations of two-phase flow models. This indicates that fundamental modifications to the approach outlined in [101, 194] are sought. Hence, such a parametrization for non-linear distributed parameter port-Hamiltonian systems will not be the subject of this chapter. Instead, we will identify a particular (non-unique) definition of boundary port-variables to define geometrical structure for the TFM and the DFM.*

Remark 3.4.7 *Boundary port-variables, in our setting, will remain unchanged in the presence of dissipation. This is only true since our resistive operator ($\mathscr{R}$) does not include any differential operator. In general, the boundary ports could also*

include contributions from the resistive part. In this work, we only consider Stokes-Dirac structures without accounting for resistive ports (for the above mentioned reason) and finally arrive at a definition of the boundary port-variables, which is practical for port-Hamiltonian representations.

We recall the following fundamental lemma of calculus of variations.

Lemma 3.4.8 *If the pair $(h, m) \in \mathscr{L}^2(\Omega)^2$ satisfies*

$$\int_a^b [h(x)\partial_x f(x) + m(x)f(x)]\mathrm{d}x = 0, \tag{3.47}$$

for all $f \in H_0^1(\Omega)$, then

$$h \in H^1(\Omega), \quad and \quad \partial_x h = m(x). \tag{3.48}$$

Lemma 3.4.8 will be extensively used in order to prove that a certain structure is a Stokes-Dirac structure.

Using the above mathematical preliminaries, we first propose a Stokes-Dirac structure for the TFM and present a corresponding proof, and then we propose it for the DFM without slip.

3.4.1 Stokes-Dirac structure representation for the TFM

We first introduce the notations

$$\mathbf{f}_t = \begin{bmatrix} f_{m_\mathrm{g}} & f_{m_\ell} & f_{I_\mathrm{g}} & f_{I_\ell} & f_{a,t}^B & f_{b,t}^B \end{bmatrix}^T, \tag{3.49a}$$

$$\mathbf{e}_t = \begin{bmatrix} e_{m_\mathrm{g}} & e_{m_\ell} & e_{I_\mathrm{g}} & e_{I_\ell} & e_{a,t}^B & e_{b,t}^B \end{bmatrix}^T, \tag{3.49b}$$

$$\mathbf{f}_{tr} = \begin{bmatrix} f_{m_\mathrm{g}} & f_{m_\ell} & f_{I_\mathrm{g}} & f_{I_\ell} \end{bmatrix}^T, \tag{3.49c}$$

$$\mathbf{e}_{tr} = \begin{bmatrix} e_{m_\mathrm{g}} & e_{m_\ell} & e_{I_\mathrm{g}} & e_{I_\ell} \end{bmatrix}^T \tag{3.49d}$$

with $\mathbf{f}_t \in \mathscr{F}_t$, $\mathbf{e}_t \in \mathscr{E}_t$ where $\mathscr{F}_t = \mathscr{E}_t = \mathscr{L}^2(\Omega)^4 \times \mathbb{R}^2 \times \mathbb{R}^2$. On $\mathscr{F}_t \times \mathscr{E}_t$ the following non-degenerate bilinear form is defined:

$$\langle \mathbf{f}_t \mid \mathbf{e}_t \rangle = \int_\Omega (f_{m_\mathrm{g}} e_{m_\mathrm{g}} + f_{m_\ell} e_{m_\ell} + f_{I_\mathrm{g}} e_{I_\mathrm{g}} + f_{I_\ell} e_{I_\ell})\mathrm{d}x + (f_{b,t}^B)^T e_{b,t}^B + (f_{a,t}^B)^T e_{a,t}^B. \tag{3.50}$$

Using these notations, the Stokes-Dirac structure corresponding to the dissipative Hamiltonian representation of the TFM can be expressed as follows.

Theorem 3.4.9 *Consider $\mathscr{F}_t$ and $\mathscr{E}_t$ as introduced above. Moreover, assume that $m_g, m_\ell, I_g, I_\ell =: q_1, q_2, q_3, q_4 \in H^1(\Omega)$. We also assume that $q_1, q_2 > 0$ on Ω. Then, the linear subset $\mathscr{D}_t \subset \mathscr{F}_t \times \mathscr{E}_t$ defined as follows:*

$$\mathscr{D}_t = \Big\{ (\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{F}_t \times \mathscr{E}_t \mid \mathbf{e}_{tr} \in H^1(\Omega)^4, \mathbf{f}_{tr} = \mathscr{J}_t(q)\mathbf{e}_{tr}, $$

$$\begin{pmatrix} f_{b,t}^B \\ e_{b,t}^B \end{pmatrix} = \begin{pmatrix} f_{b1,t}^B \\ f_{b2,t}^B \\ e_{b1,t}^B \\ e_{b2,t}^B \end{pmatrix} = \left(\begin{pmatrix} q_1 & 0 & q_3 & 0 \\ 0 & q_2 & 0 & q_4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_{m_g} \\ e_{m_\ell} \\ e_{I_g} \\ e_{I_\ell} \end{pmatrix} \right) (b), \tag{3.51}$$

$$\begin{pmatrix} f_{a,t}^B \\ e_{a,t}^B \end{pmatrix} = \begin{pmatrix} f_{a1,t}^B \\ f_{a2,t}^B \\ e_{a1,t}^B \\ e_{a2,t}^B \end{pmatrix} = \left(\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ q_1 & 0 & q_3 & 0 \\ 0 & q_2 & 0 & q_4 \end{pmatrix} \begin{pmatrix} e_{m_g} \\ e_{m_\ell} \\ e_{I_g} \\ e_{I_\ell} \end{pmatrix} \right) (a) \Big\},$$

where

$$\mathscr{J}_t(q) = - \begin{bmatrix} 0 & 0 & \partial_x(m_g \cdot) & 0 \\ 0 & 0 & 0 & \partial_x(m_\ell \cdot) \\ m_g \partial_x(\cdot) & 0 & \partial_x(I_g \cdot) + I_g \partial_x(\cdot) & 0 \\ 0 & m_\ell \partial_x(\cdot) & 0 & \partial_x(I_\ell \cdot) + I_\ell \partial_x(\cdot) \end{bmatrix} \tag{3.52}$$

is a Stokes-Dirac structure with respect to the symmetric pairing given by

$$\ll (\mathbf{f}_t, \mathbf{e}_t), (\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \gg = \langle \mathbf{f}_t \mid \tilde{\mathbf{e}}_t \rangle + \langle \tilde{\mathbf{f}}_t \mid \mathbf{e}_t \rangle, (\mathbf{f}_t, \mathbf{e}_t), (\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \in \mathscr{F}_t \times \mathscr{E}_t, \tag{3.53}$$

where the pairing $\langle \cdot \mid \cdot \rangle$ is given in (3.50).

Proof: The proof is divided into two parts. We first prove that $\mathscr{D}_t \subset \mathscr{D}_t^{\perp}$.

We consider two pairs of flow and effort variables belonging to the Stokes-Dirac structure, i.e., $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ and $(\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \in \mathscr{D}_t$. Using the earlier introduced notations, we obtain:

$$\ll (\mathbf{f}_t, \mathbf{e}_t), (\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \gg = \int_\Omega (f_{m_g} \tilde{e}_{m_g} + f_{m_\ell} \tilde{e}_{m_\ell} + f_{I_g} \tilde{e}_{I_g} + f_{I_\ell} \tilde{e}_{I_\ell}) \mathrm{d}x + \tag{3.54}$$

$$\int_\Omega (\tilde{f}_{m_g} e_{m_g} + \tilde{f}_{m_\ell} e_{m_\ell} + \tilde{f}_{I_g} e_{I_g} + \tilde{f}_{I_\ell} e_{I_\ell}) \mathrm{d}x +$$

$$(f_{a,t}^B)^T \tilde{e}_{a,t}^B + (f_{b,t}^B)^T \tilde{e}_{b,t}^B + (\tilde{f}_{a,t}^B)^T e_{a,t}^B + (\tilde{f}_{b,t}^B)^T e_{b,t}^B.$$

Substituting the mappings between the flow and the effort variables, as in (3.51), it is straightforward to see that equation (3.54) equals zero and so $\mathscr{D}_t \subset \mathscr{D}_t^\perp$. This concludes the first part of the proof.

We now prove the converse part, i.e., $\mathscr{D}_t^\perp \subset \mathscr{D}_t$. For this, we follow the steps similar to Proposition 4.1 in [53]. The proof consists of several repeated steps, which are summarized below. We take $(\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \in \mathscr{D}_t^\perp$ i.e., $(\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \in \mathscr{F}_t \times \mathscr{E}_t$ such that $\ll (\mathbf{f}_t, \mathbf{e}_t), (\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \gg = 0 \quad \forall (\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$. To this end, we use the freedom in the choice of the effort variables and exploit Lemma 3.4.8.

Step 1: Let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_\ell}, e_{I_g}, e_{I_\ell} = 0$ and $e_{m_g}(a) = e_{m_g}(b) = 0$. Using (3.54), we find that

$$\int_\Omega -(q_1 \partial_x e_{m_g}) \tilde{e}_{I_g} + \tilde{f}_{m_g} e_{m_g} \mathrm{d}x = 0 \quad \forall e_{m_g} \in H_0^1(\Omega). \tag{3.55}$$

Lemma 3.4.8 gives

$$q_1 \tilde{e}_{I_g} \in H^1(\Omega) \quad \text{and} \quad \tilde{f}_{m_g} = -\partial_x(q_1 \tilde{e}_{I_g}). \tag{3.56}$$

Using $q_1 \in H^1(\Omega)$ along with $q_1 > 0$ on Ω, we obtain that $\tilde{e}_{I_g} \in H^1(\Omega)$.

Step 2: Considering $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_g}, e_{I_g}, e_{I_\ell} = 0$ and $e_{m_\ell} \in H_0^1(\Omega)$, and following the procedure as in the previous step, we obtain:

$$q_2 \tilde{e}_{I_\ell} \in H^1(\Omega) \quad \text{and} \quad \tilde{f}_{m_\ell} = -\partial_x(q_2 \tilde{e}_{I_\ell}). \tag{3.57}$$

As before, using $q_2 \in H^1(\Omega)$ along with $q_2 > 0$ on Ω, we have that $\tilde{e}_{I_\ell} \in H^1(\Omega)$.

Step 3: For $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_g}, e_{m_\ell}, e_{I_\ell} = 0$ and $e_{I_g} \in H_0^1(\Omega)$, we obtain:

$$\int_\Omega -\partial_x(q_1 e_{I_g}) \tilde{e}_{m_g} - \partial_x(q_3 e_{I_g}) \tilde{e}_{I_g} - (q_3 \partial_x e_{I_g}) \tilde{e}_{I_g} + \tilde{f}_{I_g} e_{I_g} \mathrm{d}x = 0 \quad \forall e_{I_g} \in H_0^1(\Omega).$$

We rewrite the above equation as follows:

$$\int_\Omega -(\partial_x q_1)(e_{I_g} \tilde{e}_{m_g}) - (\partial_x q_3)(e_{I_g} \tilde{e}_{I_g}) - (\partial_x e_{I_g}) \cdot$$
$$\left(q_1 \tilde{e}_{m_g} + 2 q_3 \tilde{e}_{I_g} \right) + \tilde{f}_{I_g} e_{I_g} \mathrm{d}x = 0 \quad \forall e_{I_g} \in H_0^1(\Omega).$$

As a result of Lemma 3.4.8, we have that $q_1 \tilde{e}_{m_g} + 2 q_3 \tilde{e}_{I_g} \in H^1(\Omega)$. Moreover, we obtain the following identity:

$$\tilde{f}_{I_g} = -\partial_x(q_1 \tilde{e}_{m_g} + 2 q_3 \tilde{e}_{I_g}) + \tilde{e}_{m_g} \partial_x q_1 + \tilde{e}_{I_g} \partial_x q_3. \tag{3.58}$$

Using $q_1, q_3, \tilde{e}_{I_g} \in H^1(\Omega)$ and that $q_1 > 0$, it can easily be deduced that $\tilde{e}_{m_g} \in H^1(\Omega)$, and so (3.58) can be written as

$$\tilde{f}_{I_g} = -q_1 \partial_x \tilde{e}_{m_g} - \partial_x(q_3 \tilde{e}_{I_g}) - q_3 \partial_x \tilde{e}_{I_g}. \tag{3.59}$$

Step 4: Considering $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_g}, e_{m_\ell}, e_{I_g} = 0$ and $e_{I_\ell} \in H^1_0(\Omega)$, we obtain:

$$\int_\Omega -\partial_x(q_2 e_{I_\ell}) \tilde{e}_{m_\ell} - \partial_x(q_4 e_{I_\ell}) \tilde{e}_{I_\ell} - (q_4 \partial_x e_{I_\ell}) \tilde{e}_{I_\ell} + \tilde{f}_{I_\ell} e_{I_\ell} \mathrm{d}x = 0 \quad \forall e_{I_\ell} \in H^1_0(\Omega).$$

Re-writing the above equation as in the previous step and using Lemma 3.4.8, we have that $q_2 \tilde{e}_{m_\ell} + 2q_4 \tilde{e}_{I_\ell} \in H^1(\Omega)$ and also obtain:

$$\tilde{f}_{I_\ell} = -\partial_x(q_2 \tilde{e}_{m_\ell} + 2q_4 \tilde{e}_{I_\ell}) + \tilde{e}_{m_\ell} \partial_x q_2 + \tilde{e}_{I_\ell} \partial_x q_4. \tag{3.60}$$

Using $q_2, q_4, \tilde{e}_{I_\ell} \in H^1(\Omega)$ and that $q_2 > 0$, it can easily be deduced that $\tilde{e}_{m_\ell} \in H^1(\Omega)$ and so

$$\tilde{f}_{I_\ell} = -q_2 \partial_x \tilde{e}_{m_\ell} - \partial_x(q_4 \tilde{e}_{I_\ell}) - q_4 \partial_x \tilde{e}_{I_\ell}. \tag{3.61}$$

Step 5: Let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_\ell} = e_{I_g} = e_{I_\ell} = 0$ and $e_{m_g}(a) = 0, e_{m_g}(b) \neq 0$. Using the procedure outlined above, we obtain the following identity: $\tilde{e}^B_{b1,t} = \tilde{e}_{I_g}|_b$.

Step 6: Let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_g} = e_{I_g} = e_{I_\ell} = 0$ and $e_{m_\ell}(a) = 0, e_{m_\ell}(b) \neq 0$. We now observe that $\tilde{e}^B_{b2,t} = \tilde{e}_{I_\ell}|_b$ holds.

Step 7: Let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_g} = e_{m_\ell} = e_{I_\ell} = 0$ and $e_{I_g}(a) = 0, e_{I_g}(b) \neq 0$. Using the outlined procedure, we now obtain:

$$-(q_1 \tilde{e}_{m_g} e_{I_g})|_b - (q_3 \tilde{e}_{I_g} e_{I_g})|_b + \tilde{f}^B_{b1,t} e_{I_g}|_b = 0. \tag{3.62}$$

Finally, we obtain the following identity:

$$\tilde{f}^B_{b1,t} = \left(q_1 \tilde{e}_{m_g} + q_3 \tilde{e}_{I_g}\right)|_b. \tag{3.63}$$

Step 8: Let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_{m_g} = e_{m_\ell} = e_{I_g} = 0$ and $e_{I_\ell}(a) = 0, e_{I_\ell}(b) \neq 0$. Using the outlined procedure, we now obtain the following identity:

$$\tilde{f}^B_{b2,t} = \left(q_2 \tilde{e}_{m_\ell} + q_4 \tilde{e}_{I_\ell}\right)|_b. \tag{3.64}$$

The boundary port-variables $f^B_{a1.t}, f^B_{a2,t}, e^B_{a1,t}$ and $e^B_{a2,t}$ can be obtained in a manner similar to the one outlined for computing the boundary port-variables at the right boundary of the spatial domain Ω.

Thus, in summary we have shown $\mathscr{D}^\perp_t \subset \mathscr{D}_t$ and, hence, $\mathscr{D}_t$ is a Stokes-Dirac structure. ∎

Remark 3.4.10 *We have demonstrated the Stokes-Dirac structure representation of the TFM for one particular (conservative) choice of state coordinates. The Stokes-Dirac structure representation can be demonstrated analogously for the choice of the state vector as in Theorem 3.3.3, which is trivial due to the state-independent nature of the (formally) skew-adjoint operator $\tilde{\mathcal{J}}_T$.*

We now discuss the representation of the Stokes-Dirac structure corresponding to the skew-adjoint operator $\mathcal{J}_D$ in the scope of the DFM without slip.

3.4.2　Stokes-Dirac structure representation for the DFM

We introduce the notations

$$\mathbf{f}_d = \begin{bmatrix} f_{m_{\mathrm{g}},d} & f_{m_\ell,d} & f_{v,d} & f_{a,d}^B & f_{b,d}^B \end{bmatrix}^T, \tag{3.65a}$$

$$\mathbf{e}_d = \begin{bmatrix} e_{m_{\mathrm{g}},d} & e_{m_\ell,d} & e_{v,d} & e_{a,d}^B & e_{b,d}^B \end{bmatrix}^T, \tag{3.65b}$$

$$\mathbf{f}_{dr} = \begin{bmatrix} f_{m_{\mathrm{g}},d} & f_{m_\ell,d} & f_{v,d} \end{bmatrix}^T, \tag{3.65c}$$

$$\mathbf{e}_{dr} = \begin{bmatrix} e_{m_{\mathrm{g}},d} & e_{m_\ell,d} & e_{v,d} \end{bmatrix}^T. \tag{3.65d}$$

A Stokes-Dirac structure for the dissipative Hamiltonian representation of the DFM can be expressed as follows.

Theorem 3.4.11 *Consider $\mathcal{F}_d = \mathcal{E}_d = \mathcal{L}^2(\Omega)^3 \times \mathbb{R}^2$. We assume that $A_g := \frac{m_{\mathrm{g}}}{m_{\mathrm{g}}+m_\ell}$, $A_\ell := \frac{m_\ell}{m_{\mathrm{g}}+m_\ell} \in H^1(\Omega)$. We also consider that the non-degenerate bilinear form on $\mathcal{F}_d \times \mathcal{E}_d$ is defined in the following way:*

$$\langle \mathbf{f}_d \mid \mathbf{e}_d \rangle = \int_\Omega (f_{m_{\mathrm{g}},d}e_{m_{\mathrm{g}},d} + f_{m_\ell,d}e_{m_\ell,d} + f_{v,d}e_{v,d})\mathrm{d}x + f_{b,d}^B e_{b,d}^B + f_{a,d}^B e_{a,d}^B. \tag{3.66}$$

Then, the linear subset $\mathcal{D}_d \subset \mathcal{F}_d \times \mathcal{E}_d$ given by

$$\mathcal{D}_d = \left\{ (\mathbf{f}_d, \mathbf{e}_d) \in \mathcal{F}_d \times \mathcal{E}_d, \begin{pmatrix} A_g e_{m_{\mathrm{g}},d} + A_\ell e_{m_\ell,d} \\ e_{v,d} \end{pmatrix} \in H^1(\Omega)^2, \mathbf{f}_{dr} = \mathcal{J}_d(z_D)\mathbf{e}_{dr}, \right.$$

$$\left. \begin{pmatrix} f_{a,d}^B \\ e_{a,d}^B \end{pmatrix} = \left(\begin{pmatrix} -A_g & -A_\ell & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_{m_{\mathrm{g}},d} \\ e_{m_\ell,d} \\ e_{v,d} \end{pmatrix} \right)(a), \begin{pmatrix} f_{b,d}^B \\ e_{b,d}^B \end{pmatrix} = \left(\begin{pmatrix} A_g & A_\ell & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_{m_{\mathrm{g}},d} \\ e_{m_\ell,d} \\ e_{v,d} \end{pmatrix} \right)(b) \right\}, \tag{3.67}$$

where

$$\mathcal{J}_d(z_D) = \begin{pmatrix} 0 & 0 & -\partial_x(A_g\cdot) \\ 0 & 0 & -\partial_x(A_\ell\cdot) \\ -D(A_g\cdot)\&D(A_\ell\cdot) & & 0 \end{pmatrix}$$

is a Stokes-Dirac structure with respect to the symmetric pairing given by, see (3.66):

$$\ll (\mathbf{f}_d, \mathbf{e}_d), (\tilde{\mathbf{f}}_d, \tilde{\mathbf{e}}_d) \gg = \langle \mathbf{f}_d \mid \tilde{\mathbf{e}}_d \rangle + \langle \tilde{\mathbf{f}}_d \mid \mathbf{e}_d \rangle, (\mathbf{f}_d, \mathbf{e}_d), (\tilde{\mathbf{f}}_d, \tilde{\mathbf{e}}_d) \in \mathscr{F}_d \times \mathscr{E}_d. \quad (3.68)$$

$D(A_g \cdot) \& D(A_\ell \cdot)$ *is the operator with domain all* $e_{m_g, d}, e_{m_\ell, d} \in \mathscr{L}^2(\Omega)$ *such that* $A_g e_{m_g, d} + A_\ell e_{m_\ell, d} \in H^1(\Omega)$, *and its action is given by*

$$D(A_g \cdot) \& D(A_\ell \cdot) \begin{pmatrix} e_{m_g, d} \\ e_{m_\ell, d} \end{pmatrix} = \partial_x (A_g e_{m_g, d} + A_\ell e_{m_\ell, d}) - e_{m_g, d} \partial_x A_g - e_{m_\ell, d} \partial_x A_\ell. \quad (3.69)$$

Remark 3.4.12 *This can be considered as a special case of the extended structure shown in Proposition 4.2.9, Chapter 4 in the context of spatially-varying cross-section. We skip the proof of Theorem 3.4.11 and instead refer to Proposition 4.2.9, Chapter 4 for its derivation.*

Remark 3.4.13 *The formally skew-adjoint operator* $\mathscr{J}_T(q)$ *in Theorem 3.3.6 is equal to the skew-adjoint operator* $\mathscr{J}_t(q)$ *associated to the Stokes-Dirac structure representation in Theorem 3.4.9. These operators are found to be equal only because of the assumptions on the state variables q; see Theorem 3.4.9. In general, the formally skew-adjoint operator and the skew-adjoint operator associated to the Stokes-Dirac structure representation need not be the same, see Theorem 3.4.11.*

We have shown Stokes-Dirac structure representations for both dissipative Hamiltonian formulations of the mathematical models under consideration.

3.5 Special case considerations for the DFM

In this section, we disqualify the DFM with the Zuber-Findlay slip conditions as an energy consistent model for two-phase flow, and, thus, motivate the reasons behind considering the DFM without slip.

  We recall the dissipation inequality obeyed by the TFM (see Theorem 3.3.6). Under the imposition of periodic boundary conditions, the time derivative of the Hamiltonian (3.25) can be expressed using (3.2d) as follows:

$$\begin{aligned}
\frac{\mathrm{d}\mathscr{H}}{\mathrm{d}t} &= -\int_\Omega (\delta_q \mathscr{H}(q))^T (\mathscr{R}_T) \delta_q \mathscr{H}(q) \, \mathrm{d}x, \\
&= -\int_\Omega b_g^M (v_g - v_\ell)^2 \, \mathrm{d}x, \qquad (3.70) \\
&= -\int_\Omega M_{ig}(v_\ell - v_g) \, \mathrm{d}x.
\end{aligned}$$

The equivalence between the TFM and DFM, discussed in Section 3.2, gives a better understanding of the DFM, especially when comparing the energy considerations between these two models since the only difference is how the term M_{ig} is chosen. In the TFM, it is chosen to be proportional to the slip velocity $v_\ell - v_g$ with a non-negative coefficient of proportionality b_g^M. This linear relationship has been chosen to enforce an entropy inequality [87] and it is the basic ingredient to show that the Hamiltonian is non-increasing along solutions, see Theorem 3.3.6. However, to imitate the behaviour of the DFM from the TFM, the expression for M_{ig} in (3.17) is much more complex and it is challenging to analytically investigate the sign of the term $\int_\Omega M_{ig}(v_\ell - v_g)\, dx$ that appears in (3.70).

If the term $\int_\Omega M_{ig}(v_\ell - v_g)\, dx$ is always positive, it can be claimed that the dissipation inequality $d\mathcal{H}/dt \le 0$ also holds for the (general) DFM (using (3.70)). It is worth recalling that the dissipation inequality $d\mathcal{H}_D/dt \le 0$ holds for the DFM under zero slip considerations (see Theorem 3.3.10).

The theoretical assessment of the term $\int_\Omega M_{ig}(v_\ell - v_g)\, dx$ for the model with non-zero slip is not straightforward due to its complex non-linear nature. Hence, we investigate its behaviour numerically. In order to calculate μ_g, μ_ℓ and ζ as in Theorem 3.2.5, the same expressions as computed in [56] are used. The Rusanov scheme [156] together with Zuber-Findlay slip (with $K = 1.07$ and $S = 0.216$ m/s cf. (3.14)) is used to solve the DFM numerically in a horizontal 1000 m-long spatial domain with the spatial and temporal step size of 0.5 m and 0.0005 s, $p_{\ell 0} = 1$ bar, $\rho_{\ell 0} = 1000$ kg/m^3, $c_\ell = 1000$ m/s, and $c_g = 316$ m/s. We consider periodic boundary conditions with the initial condition as shown in Figure 3.1. We use this test case to draw a concrete conclusion on the sign of $\int_\Omega M_{ig}(v_\ell - v_g)\, dx$. As obvious from Figure 3.1, we have found a counter example for which this integral is negative for all time instants.

The numerical results indicate that the proposed Hamiltonian $\mathcal{H}_D$ with periodic boundary considerations does not guarantee the non-increasing behaviour of the Hamiltonian functional along solutions of the DFM. A possible underlying reason for this effect could be that the Hamiltonian (5.8) (under zero gravitational contribution) is not suitable for the DFM with the Zuber-Findlay slip. However, the Hamiltonian $\mathcal{H}_D$ has the interpretation of the energy. The increment in this energy along the solutions in principle disqualifies the DFM for such slip conditions as an energy-consistent model for two-phase flow. Hence, we do not consider the general case of the DFM and only focus on a special case of the model, i.e., the model without slip.
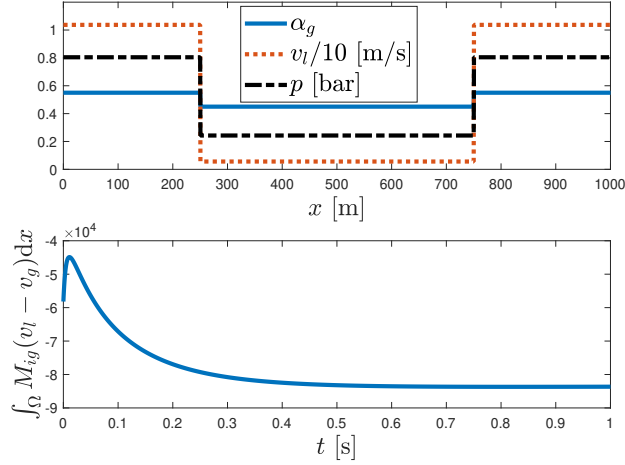
Figure 3.1: Initial condition (top) and the temporal evolution of $\int_\Omega M_{ig}(v_\ell - v_g)\mathrm{d}x$ for the DFM with periodic boundary conditions (bottom) .

3.6 Conclusions

We introduced a dissipative Hamiltonian formulation for two variants of multi-phase flow models, i.e., the Two-Fluid Model (TFM) and the no-slip Drift Flux Model (DFM) across a constant cross-section. Moreover, we presented *state-dependent* Stokes-Dirac structure representations corresponding to the skew-adjoint operators obtained for both the TFM and for the DFM without slip (under certain choice of state-variables) along with the proof of corresponding representation for the TFM. Port-Hamiltonian representations for the multi-phase models are implicitly represented in terms of the Stokes-Dirac structures. Additionally, we numerically reasoned, by exploiting a connection to the TFM, to support the consideration of the DFM without slip.

We have shown that complex non-linear partial differential equations, such as the ones governing two-phase flow, can also be cast in port-Hamiltonian formulations via underlying Stokes-Dirac structures. We stress that the developed port-Hamiltonian realizations are not able to deal with vanishing phases, and this calls for coming up with an appropriate mitigating measure in order to have port-Hamiltonian formulations across whole range of operating condi-

tions. Developing Stokes-Dirac structure representations requires generalizations/extensions to the matrix/operator theory in the scope of non-linear distributed parameter port-Hamiltonian systems; and a large part of this chapter is dedicated to address this aspect. The parametrization of boundary port variables in the scope of obtained port-Hamiltonian realizations turns out to be a non-trivial task, and deserves further attention for developing a sound mathematical framework.

A result presented in Section 3.5 implies that the DFM with the Zuber-Findlay slip conditions is not an energy consistent model for two-phase flow and, thus, motivates the need for identifying energy-consistent slip laws. In the past, the DFM with the Zuber-Findlay slip conditions has been observed to yield numerically unstable results (for some operating conditions) when finite-volume based discretization is applied on such model. However, these observations lack an underlying explanation in the sense that the response due to the nature of the mathematical model and the nature of the employed numerical method for discretization has not been decoupled; and, generally, the numerical method has been held responsible for observed issues. In view of the non-existing rigorous theoretical/mathematical analysis of two-phase flow models, the result obtained in Section 3.5 holds a significance as it indicates that the mathematical nature of the model itself is susceptible to observed behaviour; and that numerical solvers may not be the 'right' tooling to be held accountable for the observed numerical behaviour. In hindsight, we also believe that the lack of an energy-consistent model under Zuber-Findlay slip conditions is attributed to the conditional hyperbolicity of the DFM. However, these aspects do not form the focus of this dissertation, and definitely deserve further focus in the future.

Chapter 4

Compositional port-Hamiltonian modelling for MPD systems

Many single- and multi-phase fluid dynamical systems, as, e.g., encountered in Managed Pressure Drilling (MPD), are governed by non-linear evolutionary partial differential equations. A key aspect of these systems is that the fluid typically flows across spatially and temporally varying cross-sections. We, first, show that not any choice of state-variables may be apt for obtaining a port-Hamiltonian realization under spatially varying cross-sections, which are encountered in MPD. We propose a suitable modified choice of the state-variables and then represent fluid dynamical systems of interest in port-Hamiltonian representations. We define these port-Hamiltonian representations under spatial variation in the cross-section with respect to a new proposed state-dependent and extended Stokes-Dirac structure. Moreover, we also account for temporal variations in the cross-section and obtain a suitable structure that respects key properties, such as, for instance, the property of dissipation inequality. Furthermore, a single-phase flow model, a dissipative mathematical component, for instance, a drill bit, and a single-phase/two-phase flow model are interconnected to predict hydraulics for MPD applications. These subsystems are interconnected in a power-preserving manner to build an aggregated port-Hamiltonian system for real-life MPD scenarios. We derive the condition for the passivity of the composed system. We also prove that the interconnection

junction connecting the single- and two-phase flow models possesses a conditional dissipativity property, which is closely related to the passivity of the composed MPD system.

4.1   Introduction

In the previous chapter, we focused on developing port-Hamiltonian representations for two-phase flow models (across a unit and *constant* cross-section) useful in the scope of modelling multi-phase dynamics in a wellbore under gas influx from the reservoir. The focus of this chapter is on developing a (composed) port-Hamiltonian representation of an automated Managed Pressure Drilling (MPD) system, introduced earlier in Chapter 2, in order to model a vast range of real-life drilling operations. To this end, we recall that hydraulics in the MPD system can be characterized by interconnection of subsystems governed by single- and two-phase flow models, and mathematical models governed by non-linear ordinary differential or algebraic equations (for instance, a hydraulic model across a bit). Furthermore, as mentioned in Section 2.3.1, Chapter 2, the fluid flow path comprises of different geometrical specifications. This motivates the need to develop port-Hamiltonian representations for (single or multi-phase) fluid dynamical systems admitting flows across variable cross-section before interconnecting different subsystems to obtain a (composed) port-Hamiltonian representation of an automated MPD system. Moreover, the geometrical cross-section across which the fluid flows can vary over time during some drilling operations. For instance, during tripping, the drill string moves at a certain speed, and this results in temporally varying flow cross-section across different parts of the annulus. In view of the aforementioned discussion, the focus of Section 4.2 and Section 4.3 is on developing port-Hamiltonian representations for fluid dynamical systems with spatially and temporally varying cross-sections. In Section 4.4 - Section 4.6, we will focus on interconnecting subsystems to form an aggregated port-Hamiltonian representation for the MPD system.

The modelling effort to develop port-Hamiltonian realizations for fluid dynamics across varying cross-sections is a stepping stone towards simulation and control for MPD systems as introduced in [130]. In this scope, we provide an overview of the state of the art. A port-Hamiltonian formulation for single-phase models for flows across constant cross-section already exists with several different choices of the equation of state for the phase of interest; see [48]. Moreover, we recall that port-Hamiltonian formulations have been presented in Chapter 3 (or [19]), for two-phase models with fluid(s)/phases(s) flowing across *con-*

stant (and unit) cross-section. However, to the best of our knowledge, no works have considered a port-Hamiltonian representation of single- and two-phase flow models across spatially and temporally varying cross-sections. Moreover, we recall that most of the research in the field of port-Hamiltonian systems has been done for quadratic Hamiltonian functionals. It is also well known that single-phase or multi-phase flow models possess non-quadratic Hamiltonian functionals; see Chapter 3 and [95]. It is of great interest to investigate whether accounting for varying geoemetrical cross-sections in the port-Hamiltonian framework requires fundamental mathematical modifications to the existing theory of port-Hamiltonian systems.

The benefits of port-Hamiltonian framework have been already outlined in Chapter 3. Here, we motivate the advantage of such a framework in the scope of compositional system modelling. The port-Hamiltonian framework has gained a lot of momentum recently [13, 16, 34, 43, 62, 97, 116–119, 178, 184, 185, 207]. This is mainly owed to the favourable properties of such a formalism for both finite- and infinite-dimensional dynamical systems that are characterized by differential, algebraic or mixture of differential and algebraic equations; refer to [185, 188, 189]. Furthermore, this paradigm has helped to integrate finite- and infinite-dimensional components and preserve key system-theoretic properties due to an attractive property called compositionality [139, 185]. In other words, a key property of port-Hamiltonian systems is that the interconnection of several port-Hamiltonian systems still preserve the port-Hamiltonian structure (in an aggregated sense) if the interconnection is performed in a power-preserving manner [42, 53].

A lot of work has already been done in the past in the scope of power-preserving interconnection of finite-dimensional and infinite-dimensional components [42, 53]. The fundamental tool to compose subsystems in a power-preserving manner is an underlying geometric structure called Dirac structure. The special class of Dirac structures, known as Stokes-Dirac structures, were extensively discussed in Chapter 3. Such geometric structures are also studied in this chapter. In the scope of power-preserving interconnection, it is already well known that the composition of Dirac structures in the finite-dimensional setting is known to always yield a (composed) Dirac structure [120, 121, 186]. However, the composition of Dirac structures in an infinite-dimensional setting may not always be a (composed) Stokes-Dirac structure. In other words, the compositionality holds only conditionally in an infinite-dimensional setting. For instance, necessary and sufficient conditions for the composition of two infinite-dimensional Dirac structures to be a Dirac structure were given in [99]. In view of such conditional composition in the presence of infinite-dimensional systems,

the compositional port-Hamiltonian modelling of an MPD system is not only a significant research topic for application purposes, but also from a mathematical standpoint.

To the best of our knowledge, compositional port-Hamiltonian modelling for MPD has not been taken up yet and, hence, this forms one of the focal objectives in this chapter. We employ the existing theory to interconnect (individual) mathematical subsystem models: *(i)* a single-phase flow port-Hamiltonian realization corresponding to the fluid dynamics in the drill string, *(ii)* hydraulics model across drill bit, and *(iii)* a single- or two-phase flow port-Hamiltonian realization corresponding to the fluid dynamics in the annulus under gas/liquid influx from the reservoir. We connect these models in a power-preserving manner and eventually construct an aggregated port-Hamiltonian model for MPD. A compositional port-Hamiltonian representation of the MPD model is useful when it needs to be connected to other systems such as a reservoir model, where each system is characterized by a particular energy property. The compositional structure, presented in this work, can be viewed as a stepping stone towards a holistic paradigm for predicting MPD-controlled drilling dynamics. Most controllers for MPD are designed based on lumped-parameter models approximating the hydraulics and, thereby, ignore the fast pressure dynamics [2]. The framework introduced in this chapter enables an energy-based controller design while taking all (infinite-dimensional) dynamics into account.

The main contributions of this chapter are as follows: *(i)* we develop (dissipative) Hamiltonian representations of single-phase and two-phase flow models for flows across spatially and temporally varying cross-sections, *(ii)* we propose a novel state-dependent and extended Stokes-Dirac structure to account for dissipation effects and the exchange of energy via the boundaries, and define port-Hamiltonian representations with respect to these Stokes-Dirac structures, *(iii)* we perform a power-preserving interconnection of fluid dynamics model for the drill string, hydraulics model across bit, and fluid dynamics model for the annulus, and *(iv)* we derive a conditional power-preserving region at the junction (i.e., across the bit) connecting the drill string and the annulus.

The structure of this chapter is as follows. We consider (only) spatially varying geometry and present (dissipative) Hamiltonian representations, propose state-dependent and extended Stokes-Dirac structures, and define port-Hamiltonian representations with respect to these structures in Section 4.2 for mathematical models of interest. Section 4.3 discusses the corresponding port-Hamiltonian structure under both spatial and temporal variations in the area of cross-section. In Section 4.4, single- and two-phase flow port-Hamiltonian model realizations (accounting for effects of area) are restated in the sense of

considered global coordinate system to define a composed MPD system. In Section 4.5, the liquid or gas influx from the reservoir is considered for the composition of an MPD system, and the interconnection of the single- and two-phase flow models together with the conditional power preservation at the interconnection junction is discussed. In Section 4.6, the power-preserving interconnection of MPD systems in a real-life drilling scenario is investigated. We finally end the chapter with the conclusions.

4.2 Port-Hamiltonian modelling - spatial area variations

In this section, we consider a single-phase flow to be modelled by Isothermal Euler equations, and a two-phase flow to be modelled by the Drift Flux Model, as introduced in Chapter 2. We refer the reader to Section 2.3.1, Chapter 2 for mathematical details of these models governing fluid dynamical behaviour under variations in the geometrical cross-sections. We will consider zero slippage between the two phases while employing the Drift Flux Model in the wake of observations provided in Section 3.5, Chapter 3. In this section, we will also consider reference pressure to be zero (i.e., $p_{\ell 0} = 0$) for the sake of simplicity.

Remark 4.2.1 *In light of Lemma 3.2.2 and Lemma 3.2.3, it is worth mentioning that, using elimination of variables, the system in* (2.1) *can be rewritten in terms of two partial differential equations in two unknowns. Similarly, the set of equations* (2.2) *and* (2.3) - (2.6) *along with the assumption of zero slippage, i.e., $v_g = v_\ell$ can be expressed in terms of three partial differential equations in three unknowns. We omit this model reformulation in this chapter and instead refer to Chapter 3 for further insights on similar models.*

We focus on accounting for only smooth spatial cross-section variations and developing corresponding port-Hamiltonian model representations in this section. Adopting the same recipe as in Chapter 3, we first introduce (dissipative) Hamiltonian representations, i.e., without boundary effects/under the assumption of zero boundary conditions for the mathematical models under consideration. The resulting formally skew-adjoint operator(s) and the resistive matrix are used as a tool to define a candidate geometrical structure, which is later shown to be a non-canonical/extended Stokes-Dirac structure. This geometric structure yields a way to describe the boundary port variables ultimately lead-

ing to port-Hamiltonian representations of the models of interest. These port-Hamiltonian representations inherit properties from the Stokes-Dirac structures.

4.2.1 Dissipative Hamiltonian Representation

We consider the single-phase model governed by (2.1). Considering the total energy of the system, the Hamiltonian functional, consisting of kinetic, internal and potential energy, is given by:

$$\mathcal{H}_s = \int_\Omega A(\rho \frac{v^2}{2} + \rho c_\ell^2 \ln \rho + c_\ell^2 \rho_{\ell 0} + \rho g x \sin\theta) dx, \tag{4.1}$$

where $\Omega = [a, b]$ refers to the spatial domain.

Remark 4.2.2 *The above functional is similar to the functional used in [48]. However, here $\mathcal{H}_s$ is distinct as it accounts for the effects of area (A). Moreover, the equation of state (an algebraic relation relating density and pressure) is also different.*

We first choose a state coordinate vector comprised of non-conservative variables, i.e., ρ and v, and aim to develop a port-Hamiltonian representation for Isothermal Euler equations (governed by the set of equations (2.1)) across a variable cross-section. This case is used as a test-bed to emphasize that not any choice of state-variables may be apt to obtain a structure with the required properties.

The Isothermal Euler equations in (2.1) can be re-written as follows:

$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} \partial_t \rho \\ \partial_t v \end{pmatrix} = \left[\underbrace{\begin{pmatrix} 0 & -\partial_x(\cdot) \\ -\partial_x(\cdot) + \left(\frac{1}{A}\partial_x A\right) \times & 0 \end{pmatrix}}_{M} + \begin{pmatrix} 0 & 0 \\ 0 & -\frac{32\mu}{\rho^2 d^2} \end{pmatrix} \right] \begin{pmatrix} \frac{\delta \mathcal{H}_s}{\delta \rho} \\ \frac{\delta \mathcal{H}_s}{\delta v} \end{pmatrix}. \tag{4.2}$$

Here, $k \times l$ indicates standard multiplication, i.e., $\left(\frac{1}{A}\partial_x A\right) \times \frac{\delta \mathcal{H}_s}{\delta \rho}$ is a standard product between $\left(\frac{1}{A}\partial_x A\right)$ and $\frac{\delta \mathcal{H}_s}{\delta \rho}$. Other notations carry the same meaning as introduced earlier in Chapter 3.

We omit the derivation as the above formulation can be obtained in a straight-forward manner.

We decompose the operator M, introduced in (4.2), as follows:

$$M := \begin{pmatrix} 0 & -\partial_x(\cdot) \\ -\partial_x(\cdot) & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \left(\frac{1}{A}\partial_x A\right) & 0 \end{pmatrix}. \tag{4.3}$$

It is trivial to see that the first term in the right-hand side of (4.3) is formally skew-adjoint. However, the second term in the above equation is not formally skew-adjoint (under spatial cross-section variation). As a result, the operator M is not formally skew-adjoint. It is also clearly observable that the second term in the right-hand side of (4.3) would have been the zero matrix (which is trivially formally skew-adjoint) under constant cross-section. Hence, the operator M would be formally skew-adjoint in that case. The above observations illustrate that the primitive state variables may not always have the desired properties attributed to general port-Hamiltonian representations. However, we stress that the system written in terms of non-conservative state variables can be formulated in a port-Hamiltonian representation with special care, as discussed next.

We first recall the definition of port-Hamiltonian descriptor systems introduced in [123] in a slightly adapted manner.

Definition 4.2.3 *A port-Hamiltonian descriptor system is a system of differential(-algebraic) equations of the form*

$$E(t,x)\dot{x} + r(t,x) = (J(t,x) - R(t,x))z(t,x) + B(t,x)u,$$
$$y = B(t,x)^T z(t,x),$$
(4.4)

with state $x(t) \in \mathbb{R}^n$, input $u(t) \in \mathbb{R}^m$, output $y(t) \in \mathbb{R}^m$, the flow matrix $E(t,x) \in \mathbb{R}^{l \times n}$, time-flow function $r(t,x) \in \mathbb{R}^l$, effort function $z(t,x) \in \mathbb{R}^l$, structure matrix $J = -J^T \in \mathbb{R}^{l \times l}$, dissipation matrix $R = R^T \geq 0$, port matrix $B(t,x) \in \mathbb{R}^{l \times m}$, and the gradient of the Hamiltonian $\mathcal{H}$ satisfies $\partial_x \mathcal{H} = E^T z$, and $\partial_t \mathcal{H} = z^T r$ pointwise. From this structure the dissipation inequality immediately directly follows, since

$$\frac{d\mathcal{H}}{dt} = (\partial_x \mathcal{H})^T \dot{x} + \partial_t \mathcal{H} = z^T(E\dot{x} + r) = z^T((J-R)z + Bu) = -z^T Rz + z^T Bu \leq y^T u.$$
(4.5)

In [123], the framework of port-Hamiltonian descriptor system has been used to describe differential algebraic equations. It has also been mentioned that this framework can also be used to describe partial differential algebraic equations. We next show that the above described single-phase flow model written in terms of non-conservative variables can be formulated in an adapted version of the framework introduced in Definition 4.2.3.

Remark 4.2.4 *We ignore the resistive effects for the sake of simplicity in the port-Hamiltonian descriptor realization mentioned below. Moreover, we also consider that $\rho_{\ell 0} = 0$. These effects do not play an essential role in showing that the model written in terms of non-conservative variables can be formulated as a port-Hamiltonian descriptor realization.*

Theorem 4.2.5 *Considering the governing equations* (2.1), *the associated dissipative Hamiltonian representation in a port-Hamiltonian descriptor formulation is given by*

$$E_s \partial_t \bar{q} = (\bar{\mathcal{J}}_s) \bar{z}_s(t, \bar{q}) \tag{4.6}$$

with state variables $\bar{q} := [\bar{q}_1, \bar{q}_2]^T = [\rho, v]^T$, *the Hamiltonian functional in* $\mathcal{H}_s$ (4.1), *where*

$$E_s = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}, \tag{4.7}$$

$$\bar{\mathcal{J}}_s = \begin{pmatrix} 0 & -\partial_x(A \cdot) \\ -A\partial_x(\cdot) & 0 \end{pmatrix} \tag{4.8}$$

is a formally skew-adjoint operator with respect to the $\mathscr{L}^2$ *inner product, and*

$$\bar{z}_s(t, \bar{q}) = (E_s^T)^{-1} \delta_{\bar{q}} \mathcal{H}_s = \begin{pmatrix} \frac{v^2}{2} + c_\ell^2 \ln \rho + c_\ell^2 \\ \rho v \end{pmatrix}. \tag{4.9}$$

Proof: The model (2.1) can be straightforwardly written only in terms of $\bar{q}$ as follows:

$$A\partial_t \rho = -\partial_x(A\rho v), \tag{4.10a}$$

$$A\partial_t v = -Av\partial_x v - \frac{Ac_\ell^2}{\rho}\partial_x \rho. \tag{4.10b}$$

The variational derivatives of $\mathcal{H}_s$ with respect to state variables $\bar{q}$ are given by:

$$\delta_{\bar{q}_1} \mathcal{H}_s = A(\frac{v^2}{2} + c_\ell^2 \ln \rho + c_\ell^2), \quad \delta_{\bar{q}_2} \mathcal{H}_s = A\rho v. \tag{4.11}$$

Using E_s as in (4.7) along with the variational derivatives computed in (4.11), it can be straightforwardly shown that (4.9) holds.

Next, we prove the claim equation by equation. The first line of (4.6) reads

$$A\partial_t \rho = -\partial_x \Big(A(\rho v) \Big). \tag{4.12}$$

The second line of (4.6) reads

$$A\partial_t v = -A\partial_x \Big(\frac{v^2}{2} + c_\ell^2 \ln \rho + c_\ell^2 \Big) = -Av\partial_x v - \frac{Ac_\ell^2}{\rho}\partial_x \rho. \tag{4.13}$$

The claim follows by observing that (4.12) together with (4.13) is identical to (4.10).

Furthermore, the (formal) skew-adjointness of $\bar{\mathcal{J}}_s$ (with respect to the $\mathscr{L}^2$ inner product) can be straightforwardly shown by following the lines of reasoning used in Section 3.3, Chapter 3.

This completes the proof. ∎

One of the key hindrances in obtaining a port-Hamiltonian formulation of a model written in terms of non-conservative variables could be due to the use of the standard $\mathscr{L}^2$ inner product. For instance, consider the model arising in musical acoustics as in [122]. The model is equivalent to a wave equation, and accounts for model coefficients varying in space. Such a spatial variation of model coefficients/parameters bears similarities to the spatially varying geometrical cross-section in the scope of this chapter. The role of inner product is clearly evident from this model example, as shown in [122]. In the context of musical acoustic model, the formal skew-adjointness of an operator holds if it is defined with respect to a weighted inner product. It is, hence, tempting to check if the operator M in (4.3), which is not formally skew-adjoint with respect to the $\mathscr{L}^2$ inner product, can be shown to be formally skew-adjoint with respect to a weighted inner product. However, it is observed that the operator M is not formally skew-adjoint even with respect to a weighted inner product. We care to mention that the operator resulting in our setting and the one arising in the context of musical acoustics differ and, hence, weighted inner product may not always resolve the issues in deriving a port-Hamiltonian formulation for a model written in terms of non-conservative variables.

The above discussion shows that a special treatment is required in deriving port-Hamiltonian (descriptor) formulations with respect to non-conservative state variables. Although we could formulate a port-Hamiltonian descriptor formulation using non-conservative state variables, in the rest of this section, we will derive port-Hamiltonian formulations with respect to conservative state variables. Besides the fact that the conservative state variables (generally) yield relevant structural properties, relying on conservative state variables is also motivated from the fact that the non-conservative numerical schemes may not converge to the correct solution if a shock wave is present in the flow.

We now define the state vector in terms of conservative state variables. In addition, we extend the reduced version of (2.1) (obtained upon elimination of variable p) by an extra equation $\partial_t A = 0$, which means that only spatial variations of A are allowed. Such an additional equation governing spatial or temporal change of area makes sense since temporal variations in geometrical cross-section will basically increase the number of differential equations by one. The latter aspect is treated in detail in Section 4.3. The proposed methodology, even for accounting only spatial variations in the cross-sectional area, can be

viewed as a step towards a unified framework under both spatial and temporal variations in the cross-section. Finally, by invoking the aforementioned modifications, we demonstrate the dissipative Hamiltonian representation for the single-phase flow model while accounting for (smooth) spatial cross-sectional area variations.

We re-write the Hamiltonian functional in terms of the chosen set of state variables $q = [q_1, q_2, q_3]^T := [A, A\rho, A\rho v]^T$. This yields

$$\mathcal{H}_s = \int_\Omega \frac{q_3^2}{2q_2} + q_2 c_\ell^2 \ln(\frac{q_2}{q_1}) + q_1 c_\ell^2 \rho_{\ell 0} + q_2 g x \sin\theta \, \mathrm{d}x. \tag{4.14}$$

We now introduce the dissipative Hamiltonian representation for the single-phase model.

Theorem 4.2.6 *Considering the reformulated form of governing equations* (2.1), *the associated dissipative Hamiltonian representation is given by*

$$\partial_t q = (\mathcal{J}_s(q) - \mathcal{R}_s(q))\delta_q \mathcal{H}_s(q) \tag{4.15}$$

with the Hamiltonian functional (4.14), *where*

$$\mathcal{J}_s = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\partial_x(q_2\cdot) \\ 0 & -q_2\partial_x(\cdot) & -q_3\partial_x(\cdot) - \partial_x(q_3\cdot) \end{pmatrix} \tag{4.16}$$

is a formally skew-adjoint operator with respect to the $\mathcal{L}^2$ *inner product, and*

$$\mathcal{R}_s = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & q_1 \frac{32\mu}{d^2} \end{pmatrix} \tag{4.17}$$

is symmetric and positive semi-definite matrix.

Proof: We evaluate the variational derivatives with respect to the state variables q. These are given by:

$$\frac{\delta \mathcal{H}_s}{\delta q_1} = -\frac{q_2}{q_1} c_\ell^2 + \rho_{\ell 0} c_\ell^2, \tag{4.18a}$$

$$\frac{\delta \mathcal{H}_s}{\delta q_2} = -\frac{q_3^2}{2q_2^2} + c_\ell^2 \ln(\frac{q_2}{q_1}) + c_\ell^2 + g x \sin\theta, \tag{4.18b}$$

$$\frac{\delta \mathcal{H}_s}{\delta q_3} = \frac{q_3}{q_2}. \tag{4.18c}$$

Using these variational derivatives, the claim that (4.15) is equivalent to a reformulated version of (2.1) (with additional $\partial_t A = 0$) follows in a manner similar to the derivation discussed in-depth in Theorem 4.2.7. Hence, we omit the derivation here.

The positive semi-definiteness and symmetric nature of $\mathscr{R}_s$ follows immediately from the positivity of q_1, μ and d, and the structure of the matrix. The formal skew-adjointness of $\mathscr{J}_s$ essentially follows from integration by parts and neglecting the boundary conditions. The operator $\mathscr{J}_s$ has terms similar to the skew-adjoint operators in Chapter 3 (or [19]). For the sake of brevity, we omit the proof and instead refer to Chapter 3 (or [19]) for a similar derivation.

Using the properties of $\mathscr{J}_s$ and $\mathscr{R}_s$, the following dissipation inequality holds:

$$
\begin{aligned}
\frac{d\mathscr{H}_s}{dt} &= \int_\Omega (\delta_q \mathscr{H}_s(q))^T \partial_t q \, dx \\
&= \int_\Omega (\delta_q \mathscr{H}_s(q))^T \Big((\mathscr{J}_s(q) - \mathscr{R}_s(q)) \delta_q \mathscr{H}_s(q) \Big) \, dx \qquad (4.19) \\
&= \int_\Omega (\delta_q \mathscr{H}_s(q))^T (-\mathscr{R}_s(q)) \delta_q \mathscr{H}_s(q) \, dx \leq 0.
\end{aligned}
$$

This completes the proof. ∎

We now consider a two-phase Drift Flux Model without slip, i.e., (2.2) and (2.3) - (2.6) along with the assumption of zero slippage, i.e., $v_g = v_\ell = v$, and show the corresponding dissipative Hamiltonian representation under the choice of conservative state variables. Following the choice of candidate Hamiltonian functional in [19], we now choose the Hamiltonian functional in the following manner:

$$
\mathscr{H}_t = \int_\Omega A(m_g \frac{v^2}{2} + m_\ell \frac{v^2}{2} + m_g c_g^2 \ln \rho_g + m_\ell c_\ell^2 \ln \rho_\ell +
$$
$$
(1-\alpha_g)\beta + (m_g + m_\ell) g x \sin\theta) dx,
$$

where $\beta = \rho_{\ell 0} c_\ell^2$. The above functional can be expressed in terms of the following choice of state-variables $\tilde{q} = [\tilde{q}_1, \tilde{q}_2, \tilde{q}_3, \tilde{q}_4]^T := [A, Am_g, Am_\ell, A(m_g + m_\ell)v]^T$ as follows:

$$
\mathscr{H}_t = \int_\Omega \Big(\tilde{q}_1 (\frac{\tilde{q}_2}{2\tilde{q}_1} v^2 + \frac{\tilde{q}_3}{2\tilde{q}_1} v^2) + \tilde{q}_2 c_g^2 \ln(\frac{p}{c_g^2}) +
$$
$$
\tilde{q}_3 c_\ell^2 \ln(\frac{p+\beta}{c_\ell^2}) + \tilde{q}_1 (1-\alpha_g)\beta + (\tilde{q}_2 + \tilde{q}_3) g x \sin\theta) dx, \quad (4.20)
$$

where v can be expressed in terms of the chosen state-variables by a relation $v = \frac{\tilde{q}_4}{\tilde{q}_2 + \tilde{q}_3}$. Moreover, we use the relations in [5] to obtain the gas void fraction α_g from the mass variables, which is given by:

$$\alpha_g = -\frac{\tilde{q}_2}{\tilde{q}_1}\frac{c_g^2}{2\beta} - \frac{\tilde{q}_3}{\tilde{q}_1}\frac{c_\ell^2}{2\beta} + \frac{1}{2} + \sqrt{\Delta}, \tag{4.21}$$

where

$$\Delta = \left(\left(\frac{\tilde{q}_2}{\tilde{q}_1}\frac{c_g^2}{2\beta} + \frac{\tilde{q}_3}{\tilde{q}_1}\frac{c_\ell^2}{2\beta} - \frac{1}{2} \right)^2 + \frac{\tilde{q}_2}{\tilde{q}_1}\frac{c_g^2}{\beta} \right). \tag{4.22}$$

The pressure p can be computed in the following way:

$$p = \frac{\tilde{q}_2}{\tilde{q}_1}c_g^2 + \frac{\tilde{q}_3}{\tilde{q}_1}c_\ell^2 - \beta(1 - \alpha_g). \tag{4.23}$$

Next, we discuss the dissipative Hamiltonian representation for the two-phase model. We consider a model reformulation of the governing equations (2.2) along with the closure equations (2.3) - (2.6) under the assumption of zero slippage, i.e., $v_g = v_\ell = v$, and, express these as a system composed of three equations in three unknowns (state-variables). Moreover, as before, we consider an additional equation $\partial_t A = 0$. We refer to the resulting model as Σ in the sequel.

Theorem 4.2.7 *The dissipative Hamiltonian representation of the reformed model Σ in the scope of two-phase flow models takes the following form:*

$$\partial_t \tilde{q} = (\mathscr{J}_t(\tilde{q}) - \mathscr{R}_t(\tilde{q}))\delta_{\tilde{q}}\mathscr{H}_t(\tilde{q}) \tag{4.24}$$

with state variables $\tilde{q} = [\tilde{q}_1, \tilde{q}_2, \tilde{q}_3, \tilde{q}_4]^T = [A, Am_g, Am_\ell, A(m_g + m_\ell)v]^T$, the Hamiltonian functional (4.20), and where

$$\mathscr{J}_t = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\partial_x(\tilde{q}_2 \cdot) \\ 0 & 0 & 0 & -\partial_x(\tilde{q}_3 \cdot) \\ 0 & -\tilde{q}_2 \partial_x(\cdot) & -\tilde{q}_3 \partial_x(\cdot) & -\partial_x(\tilde{q}_4 \cdot) - \tilde{q}_4 \partial_x(\cdot) \end{pmatrix} \tag{4.25}$$

is a formally skew-adjoint operator with respect to the $\mathscr{L}^2$ inner product, and

$$\mathscr{R}_t = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{q}_1 \frac{32\mu_m}{d^2} \end{pmatrix} \tag{4.26}$$

is a symmetric and positive semi-definite matrix.

Proof: We first compute the variational derivatives, given by:

$$\frac{\delta \mathcal{H}_t}{\delta \tilde{q}_2} = -\frac{\tilde{q}_4^2}{2(\tilde{q}_2 + \tilde{q}_3)^2} + c_g^2 \ln\left(\frac{p}{c_g^2}\right) + c_g^2 + gx\sin\theta, \qquad (4.27a)$$

$$\frac{\delta \mathcal{H}_t}{\delta \tilde{q}_3} = -\frac{\tilde{q}_4^2}{2(\tilde{q}_2 + \tilde{q}_3)^2} + c_\ell^2 \ln\left(\frac{p + \beta}{c_\ell^2}\right) + c_\ell^2 + gx\sin\theta, \qquad (4.27b)$$

$$\frac{\delta \mathcal{H}_t}{\delta \tilde{q}_4} = \frac{\tilde{q}_4}{\tilde{q}_2 + \tilde{q}_3} = v. \qquad (4.27c)$$

The variational derivative with respect to $\tilde{q}_1$ can also be computed. However, we omit its computation as the corresponding elements in the operator $\mathcal{J}_t$ and the matrix $\mathcal{R}_t$ are zero.

We now prove the claim equation by equation. The first line holds trivially as we assume that the cross-sectional area only varies spatially. The second line reads

$$\partial_t(Am_g) = -\partial_x(\tilde{q}_2 \frac{\delta \mathcal{H}_t}{\delta \tilde{q}_4}) = -\partial_x(Am_g v). \qquad (4.28)$$

Similarly, the third line results in

$$\partial_t(Am_\ell) = -\partial_x(\tilde{q}_3 \frac{\delta \mathcal{H}_t}{\delta \tilde{q}_4}) = -\partial_x(Am_\ell v). \qquad (4.29)$$

Finally, the fourth line yields

$$\partial_t(\tilde{q}_4) = -\tilde{q}_2 \partial_x(\frac{\delta \mathcal{H}_t}{\delta \tilde{q}_2}) - \tilde{q}_3 \partial_x(\frac{\delta \mathcal{H}_t}{\delta \tilde{q}_3}) - \partial_x(\tilde{q}_4 \frac{\delta \mathcal{H}_t}{\delta \tilde{q}_4}) - \tilde{q}_4 \partial_x(\frac{\delta \mathcal{H}_t}{\delta \tilde{q}_4}) - \tilde{q}_1 \frac{32\mu_m}{d^2} \frac{\delta \mathcal{H}_t}{\delta \tilde{q}_4}.$$

Substituting the variational derivatives, we have

$$\partial_t(\tilde{q}_4) = -Am_g \partial_x\left(-\frac{v^2}{2} + c_g^2 \ln\left(\frac{p}{c_g^2}\right) + c_g^2\right) - Am_\ell \partial_x\left(-\frac{v^2}{2} + c_\ell^2 \ln\left(\frac{p + \beta}{c_\ell^2}\right) + c_\ell^2\right)$$

$$- \partial_x(A(m_g + m_\ell)v^2) - A(m_g + m_\ell)v\partial_x v - A(m_g + m_\ell)g\sin\theta - A\frac{32\mu_m v}{d^2}.$$

This simplifies to:

$$\partial_t(A(m_g + m_\ell)v) = -\partial_x(A(m_g + m_\ell)v^2) - A\partial_x p -$$

$$A(m_g + m_\ell)g\sin\theta - A\frac{32\mu_m v}{d^2}, \qquad (4.30)$$

where we have used the identity

$$-Am_g c_g^2 \partial_x (\ln \frac{p}{c_g^2}) - Am_\ell c_\ell^2 \partial_x (\ln \frac{p+\beta}{c_\ell^2}) =: -A\partial_x p.$$

The formal skew-adjointness of $\mathcal{J}_t$ with respect to the $\mathcal{L}^2$ inner product and the symmetric positive semi-definiteness of $\mathcal{R}_t$ can directly be recognized in (4.25), (8.21) by following the line of reasoning as outlined in earlier proofs. This completes the proof.   ∎

Remark 4.2.8 *We have only used constant pipe-inclination θ in this work. However, it is straightforward to account for spatially varying pipe inclinations; see Chapter 3 or [19].*

4.2.2   Stokes-Dirac Structures

Infinite-dimensional port-Hamiltonian systems are described through a geometric structure known as Stokes-Dirac structure; refer to [53, 101] and Chapter 3. We recall that this geometric object helps to gain insight in describing the consistent boundary port variables, and that the properties of the Stokes-Dirac structure can be exploited in the development of energy-based boundary control laws for distributed port-Hamiltonian systems. Such a geometric structure has been associated to canonical skew-symmetric differential operators in [101]. Moreover, in [101], the notion of Stokes-Dirac structures has been extended to skew-symmetric differential operators of any order. Existing works have focused on the *state-independent* operators and have also considered an extended Stokes-Dirac structure to account for dissipative effects (which may include differential terms), while mostly dealing with quadratic Hamiltonian functionals. In the previous chapter, i.e., Chapter 3, we considered *state-dependent* Stokes-Dirac structures for (multi-phase) flows across a unit and *constant* cross-section. These defined structures were, however, not extended to account for resistive effects. In this section, we will propose *state-dependent* Stokes-Dirac structures for single- and two-phase fluid dynamics across a spatially varying geoemetrical cross-section. We will also extend these structures to account for the resistive effects in the model description of (multi-phase) hydraulics.

We refer to [53, 101] and Chapter 3 of this dissertation for the formal definition of infinite-dimensional (extended) Stokes-Dirac structures.

Next, we propose two variants of extended Stokes-Dirac structures. The port-Hamiltonian representation for the two-phase model (see Section 4.2.3)

will be defined with respect to the structure in Proposition 4.2.9. Secondly, the Stokes-Dirac structure in Proposition 4.2.10 will be used to define port-Hamiltonian representation for the single-phase model.

We first show the extended Stokes-Dirac structure representation that will be useful in the scope of the Drift Flux Model without slip. Hereto, we introduce the following notations

$$\mathbf{f}_t = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 & f_R & f_a^B & f_b^B \end{bmatrix}^T,$$

$$\mathbf{e}_t = \begin{bmatrix} e_1 & e_2 & e_3 & e_4 & e_R & e_a^B & e_b^B \end{bmatrix}^T,$$

$$\mathbf{f}_{tr} = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 & f_R \end{bmatrix}^T,$$

and,

$$\mathbf{e}_{tr} = \begin{bmatrix} e_1 & e_2 & e_3 & e_4 & e_R \end{bmatrix}^T,$$

and define the space of flow variables and effort variables in the following manner:

$$\mathscr{F}_t = \mathscr{E}_t = \mathscr{L}^2(\Omega)^5 \times \mathbb{R}^2. \tag{4.31}$$

The non-degenerated bilinear product on $\mathscr{F}_t \times \mathscr{E}_t$ is defined in the following way:

$$< \mathbf{f}_t \mid \mathbf{e}_t >= \int_\Omega (f_1 e_1 + f_2 e_2 + f_3 e_3 + f_4 e_4 + f_R e_R)\mathrm{d}x + f_b^B e_b^B + f_a^B e_a^B. \tag{4.32}$$

Proposition 4.2.9 *We consider $\mathscr{F}_t$ and $\mathscr{E}_t$ as given in (4.31). We assume that $\tilde{q}_1, \tilde{q}_2, \tilde{q}_3, \tilde{q}_4 \in H^1(\Omega)$ and that $\tilde{q}_2 + \tilde{q}_3 > 0$ on Ω. Then, the linear subset $\mathscr{D}_t \subset \mathscr{F}_t \times \mathscr{E}_t$ given by:*

$$\mathscr{D}_t = \left\{ (\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{F}_t \times \mathscr{E}_t \mid \begin{pmatrix} \tilde{q}_2 e_2 + \tilde{q}_3 e_3 \\ \tilde{q}_2 e_4 \\ e_4 \end{pmatrix} \in H^1(\Omega)^3, \mathbf{f}_{tr} = \mathscr{J}_{ext}\mathbf{e}_{tr}, \right.$$

$$\left. \begin{pmatrix} f_a^B \\ e_a^B \end{pmatrix} = \begin{pmatrix} -\tilde{q}_2 & -\tilde{q}_3 & -\tilde{q}_4 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_2 \\ e_3 \\ e_4 \end{pmatrix} |_a, \quad \begin{pmatrix} f_b^B \\ e_b^B \end{pmatrix} = \begin{pmatrix} \tilde{q}_2 & \tilde{q}_3 & \tilde{q}_4 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_2 \\ e_3 \\ e_4 \end{pmatrix} |_b \right\}, \tag{4.33}$$

with

$$\mathscr{J}_{ext} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\partial_x(\tilde{q}_2 \cdot) & 0 \\ 0 & 0 & 0 & -\partial_x(\tilde{q}_3 \cdot) & 0 \\ 0 & -D(\tilde{q}_2 \cdot)\&D(\tilde{q}_3 \cdot) & -\partial_x(\tilde{q}_4 \cdot) - \tilde{q}_4 \partial_x & -I \\ 0 & 0 & 0 & I & 0 \end{pmatrix}, \tag{4.34}$$

is an extended Stokes-Dirac structure with respect to the symmetric pairing given by:

$$
\ll \begin{bmatrix} \mathbf{f}_t \\ \mathbf{e}_t \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{f}}_t \\ \tilde{\mathbf{e}}_t \end{bmatrix} \gg = < \mathbf{f}_t \mid \tilde{\mathbf{e}}_t > + < \tilde{\mathbf{f}}_t \mid \mathbf{e}_t >, \quad \begin{bmatrix} \mathbf{f}_t \\ \mathbf{e}_t \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{f}}_t \\ \tilde{\mathbf{e}}_t \end{bmatrix} \in \mathscr{F}_t \times \mathscr{E}_t, \quad (4.35)
$$

where the pairing $< \cdot \mid \cdot >$ is given in (4.32). Furthermore, the notation $(\cdot) \mid_a$ (similarly for $(\cdot) \mid_b$) refers to the function value evaluated at the boundary $x = a$ (similarly for $x = b$). Moreover, $D(\tilde{q}_2 \cdot)\&D(\tilde{q}_3 \cdot)$ is the operator with domain all $e_2, e_3 \in \mathscr{L}^2(\Omega)$ such that $\tilde{q}_2 e_2 + \tilde{q}_3 e_3 \in H^1(\Omega)$ and the action of this operator is

$$
D(\tilde{q}_2 e_2)\&D(\tilde{q}_3 e_3) = \partial_x(\tilde{q}_2 e_2 + \tilde{q}_3 e_3) - e_2 \partial_x \tilde{q}_2 - e_3 \partial_x \tilde{q}_3. \qquad (4.36)
$$

The above action is an extension of the normal action of the operator, which for all $e_2, e_3 \in H^1(\Omega)$ will take the following form:

$$
D(\tilde{q}_2 e_2)\&D(\tilde{q}_3 e_3) = \tilde{q}_2 \partial_x e_2 + \tilde{q}_3 \partial_x e_3.
$$

Proof: The proof consists of two parts. The first part comprises of the proof of $\mathscr{D}_t \subset \mathscr{D}_t^\perp$. And, the second part comprises of the proof of $\mathscr{D}_t^\perp \subset \mathscr{D}_t$. For the first part of the proof, we begin with considering two pairs of flow and effort variables belonging to the Dirac structure, i.e., $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ and $(\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \in \mathscr{D}_t$. Using the earlier introduced notations, the pairing (4.35) takes the following form:

$$
\ll \begin{bmatrix} \mathbf{f}_t \\ \mathbf{e}_t \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{f}}_t \\ \tilde{\mathbf{e}}_t \end{bmatrix} \gg = \int_\Omega (f_1 \tilde{e}_1 + f_2 \tilde{e}_2 + f_3 \tilde{e}_3 + f_4 \tilde{e}_4 + f_R \tilde{e}_R) \mathrm{d}x + \int_\Omega (\tilde{f}_1 e_1 + \tilde{f}_2 e_2 +
$$
$$
\tilde{f}_3 e_3 + \tilde{f}_4 e_4 + \tilde{f}_R e_R) \mathrm{d}x + f_a^B \tilde{e}_a^B + f_b^B \tilde{e}_b^B + \tilde{f}_a^B e_a^B + \tilde{f}_b^B e_b^B. \quad (4.37)
$$

Using (4.33), (4.34) and (4.36) in (4.37), we obtain:

$$
\ll \begin{bmatrix} \mathbf{f}_t \\ \mathbf{e}_t \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{f}}_t \\ \tilde{\mathbf{e}}_t \end{bmatrix} \gg = \int_\Omega \Big((-\partial_x \tilde{q}_2 e_4) \tilde{e}_2 + (-\partial_x \tilde{q}_3 e_4) \tilde{e}_3 + \Big(-\partial_x(\tilde{q}_2 e_2 + \tilde{q}_3 e_3) +
$$
$$
e_2 \partial_x \tilde{q}_2 + e_3 \partial_x \tilde{q}_3 \Big) \tilde{e}_4 - \partial_x(\tilde{q}_4 e_4) \tilde{e}_4 - \tilde{q}_4 (\partial_x e_4) \tilde{e}_4 - e_R \tilde{e}_4 +
$$
$$
e_4 \tilde{e}_R \Big) \mathrm{d}x + \int_\Omega \Big((-\partial_x \tilde{q}_2 \tilde{e}_4) e_2 + (-\partial_x \tilde{q}_3 \tilde{e}_4) e_3 + \Big(-\partial_x(\tilde{q}_2 \tilde{e}_2 +
$$
$$
\tilde{q}_3 \tilde{e}_3) + \tilde{e}_2 \partial_x \tilde{q}_2 + \tilde{e}_3 \partial_x \tilde{q}_3 \Big) e_4 - \partial_x(\tilde{q}_4 \tilde{e}_4) e_4 - \tilde{q}_4 (\partial_x \tilde{e}_4) e_4 - \tilde{e}_R e_4 +
$$
$$
\tilde{e}_4 e_R \Big) \mathrm{d}x + f_a^B \tilde{e}_a^B + f_b^B \tilde{e}_b^B + \tilde{f}_a^B e_a^B + \tilde{f}_b^B e_b^B. \quad (4.38)
$$

Performing integration by parts on few terms in the above equation, (4.38) can be easily shown to be zero and hence, $\mathscr{D}_t \subset \mathscr{D}_t^\perp$. This concludes the first part of the proof, which carries the symbolism of power-conserving structure.

We now prove the converse part: $\mathscr{D}_t^\perp \subset \mathscr{D}_t$. The proof consists of several small but repeated steps. Hence, we summarize the key steps for proving the converse part. We take $(\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \in \mathscr{D}_t^\perp$, i.e., $(\tilde{\mathbf{f}}_t, \tilde{\mathbf{e}}_t) \in \mathscr{F}_t \times \mathscr{E}_t$ such that $\ll \begin{bmatrix} \mathbf{f}_t \\ \mathbf{e}_t \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{f}}_t \\ \tilde{\mathbf{e}}_t \end{bmatrix} \gg\, = 0$ for all $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$. Furthermore, we make a certain choice on the effort variables (which can be freely chosen as per the definition of the Stokes-Dirac structure) in each step. We also exploit the fundamental lemma of calculus of variations (i.e., Lemma 3.4.8) to obtain several identities. Each step (and the associated choices) is described below:

Step 1: Let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_2, e_3, e_4, e_R = 0$ and $e_1(a) = e_1(b) = 0$. Following the procedure leads to:

$$\int_\Omega \tilde{f}_1 e_1 \mathrm{d}x = 0 \quad \forall e_1 \in \mathscr{L}^2(\Omega). \tag{4.39}$$

Thus $\tilde{f}_1 = 0$.

Step 2: We now consider $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_1, e_3, e_4, e_R = 0$ and $e_2(a) = e_2(b) = 0$. Plugging the flow-effort relations (4.33) in (4.37) under the aforementioned considerations gives:

$$\int_\Omega \Big((-\tilde{q}_2 \partial_x e_2) \tilde{e}_4 + \tilde{f}_2 e_2 \Big) \mathrm{d}x = 0 \quad \forall e_2 \in H_0^1(\Omega). \tag{4.40}$$

Using the fundamental lemma of calculus of variations gives

$$\tilde{q}_2 \tilde{e}_4 \in H^1(\Omega) \quad \text{and} \quad \tilde{f}_2 = -\partial_x(\tilde{q}_2 \tilde{e}_4). \tag{4.41}$$

Considering $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with e_1, e_2, e_4 and $e_R = 0$ along with $e_3(a) = e_3(b) = 0$ gives by a similar argument that

$$\tilde{q}_3 \tilde{e}_4 \in H^1(\Omega) \quad \text{and} \quad \tilde{f}_3 = -\partial_x(\tilde{q}_3 \tilde{e}_4). \tag{4.42}$$

Now using $\tilde{q}_2 \tilde{e}_4 \in H^1(\Omega)$ and $\tilde{q}_3 \tilde{e}_4 \in H^1(\Omega)$, we have that $(\tilde{q}_2 + \tilde{q}_3) \tilde{e}_4 \in H^1(\Omega)$. Furthermore, using $\tilde{q}_2, \tilde{q}_3 \in H^1(\Omega)$ along with $\tilde{q}_2 + \tilde{q}_3 > 0$ on Ω, we have that $\tilde{e}_4 \in H^1(\Omega)$.

Step 3: Now choosing $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_1, e_2, e_3, e_R = 0$ and $e_4 \in H_0^1(\Omega)$ gives

$$\int_\Omega \Big(-\partial_x(\tilde{q}_2 e_4) \tilde{e}_2 - \partial_x(\tilde{q}_3 e_4) \tilde{e}_3 - \partial_x(\tilde{q}_4 e_4) \tilde{e}_4 -$$
$$(\tilde{q}_4 \partial_x e_4) \tilde{e}_4 + e_4 \tilde{e}_R + \tilde{f}_4 e_4 \Big) \mathrm{d}x = 0. \tag{4.43}$$

We rewrite (4.43) as

$$
\int_{\Omega} \Big(-e_4 \tilde{e}_2 \partial_x \tilde{q}_2 - e_4 \tilde{e}_3 \partial_x \tilde{q}_3 - (\tilde{q}_2 \tilde{e}_2 + \tilde{q}_3 \tilde{e}_3) \partial_x e_4 - \partial_x (\tilde{q}_4 e_4) \tilde{e}_4 - (\tilde{q}_4 \partial_x e_4) \tilde{e}_4 +
$$
$$
e_4 \tilde{e}_R + \tilde{f}_4 e_4 \Big) \mathrm{d}x = 0 \quad \forall e_4 \in H_0^1(\Omega). \quad (4.44)
$$

As a result of the fundamental lemma of calculus of variations, we obtain the following identity:

$$
(\tilde{q}_2 \tilde{e}_2 + \tilde{q}_3 \tilde{e}_3) \in H^1(\Omega) \quad \text{and}
$$
$$
\tilde{f}_4 = -\partial_x (\tilde{q}_2 \tilde{e}_2 + \tilde{q}_3 \tilde{e}_3) + \tilde{e}_2 \partial_x \tilde{q}_2 + \tilde{e}_3 \partial_x \tilde{q}_3 - \partial_x (\tilde{q}_4 \tilde{e}_4) - \tilde{q}_4 \partial_x \tilde{e}_4 - \tilde{e}_R. \quad (4.45)
$$

Step 4: Let us consider $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_1, e_2, e_3, e_4 = 0$. The identity that follows under these considerations is:

$$
-e_R \tilde{e}_4 + \tilde{f}_R e_R = 0 \implies \tilde{f}_R = \tilde{e}_4. \quad (4.46)
$$

Step 5: Let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_1, e_3, e_4, e_R = 0$ and $e_2(a) = 0$ and $e_2(b) \neq 0$. Performing similar steps now gives:

$$
-\tilde{q}_2 e_2 \tilde{e}_4 \mid_b + \tilde{e}_b^B (\tilde{q}_2 e_2) \mid_b = 0. \quad (4.47)
$$

The identity that follows is:
$$
\tilde{e}_b^B = \tilde{e}_4 \mid_b. \quad (4.48)
$$

Step 6: We now let $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_1, e_3, e_4, e_R = 0$ and $e_2(b) = 0$ and $e_2(a) \neq 0$. We follow the procedure similar to Step 5 and obtain the following identity:

$$
\tilde{e}_a^B = \tilde{e}_4 \mid_a. \quad (4.49)
$$

Step 7: Consider $(\mathbf{f}_t, \mathbf{e}_t) \in \mathscr{D}_t$ with $e_1, e_2, e_3, e_R = 0$ and $e_4(a) = 0$ and $e_4(b) \neq 0$. Following the outlined procedure and using $\tilde{e}_b^B$ from (4.48), we have:

$$
\tilde{f}_b^B e_4 \mid_b + \tilde{e}_4 \mid_b (\tilde{q}_4 e_4) \mid_b - \tilde{q}_2 e_4 \tilde{e}_2 \mid_b -
$$
$$
\tilde{q}_3 e_4 \tilde{e}_3 \mid_b - \tilde{q}_4 e_4 \tilde{e}_4 \mid_b - \tilde{q}_4 e_4 \tilde{e}_4 \mid_b = 0. \quad (4.50)
$$

This results in the following identity:

$$
\tilde{f}_b^B = \Big(\tilde{q}_2 \tilde{e}_2 + \tilde{q}_3 \tilde{e}_3 \Big) \mid_b + \Big(\tilde{q}_4 \tilde{e}_4 \Big) \mid_b. \quad (4.51)
$$

Step 8: We now consider $(\mathbf{f}_t, \mathbf{e}_t) \in \mathcal{D}_t$ with $e_1, e_2, e_3, e_R = 0$ and $e_4(b) = 0$ and $e_4(a) \neq 0$. Under these considerations, we follow the procedure similar to Step 7 and also use $\tilde{e}_a^B$ from (4.49) to obtain the following identity:

$$\tilde{f}_a^B = -\left(q_2 \tilde{e}_2 + q_3 \tilde{e}_3\right)|_a - \left(q_4 \tilde{e}_4\right)|_a. \tag{4.52}$$

Thus, we have shown that $\mathcal{D}_t^\perp \subset \mathcal{D}_t$ and, hence, $\mathcal{D}_t$ is an extended Stokes-Dirac structure. This completes the proof. ∎

We now propose the extended Stokes-Dirac structure representation that will be useful in the scope of the single-phase model. We introduce

$$\mathbf{f}_s = \begin{bmatrix} f_1 & f_2 & f_3 & f_R & f_a^B & f_b^B \end{bmatrix}^T,$$

and

$$\mathbf{e}_s = \begin{bmatrix} e_1 & e_2 & e_3 & e_R & e_a^B & e_b^B \end{bmatrix}^T.$$

Using these notations, we define the space of flow and effort variables as follows:

$$\mathcal{F}_s = \mathcal{E}_s = \mathcal{L}^2(\Omega)^4 \times \mathbb{R}^2. \tag{4.53}$$

The non-degenerated bilinear product on $\mathcal{F}_s \times \mathcal{E}_s$ is defined as:

$$< \mathbf{f}_s \mid \mathbf{e}_s > = \int_\Omega (f_1 e_1 + f_2 e_2 + f_3 e_3 + f_R e_R)\mathrm{d}x + f_b^B e_b^B + f_a^B e_a^B. \tag{4.54}$$

Proposition 4.2.10 *Consider $\mathcal{F}_s$ and $\mathcal{E}_s$ as given in (4.53). Additionally, we consider $q_1, q_2, q_3 \in H^1(\Omega)$ and q_2 (or $A\rho$) is invertible. The linear subset $\mathcal{D}_s \subset \mathcal{F}_s \times \mathcal{E}_s$ defined as:*

$$\mathcal{D}_s = \left\{ (\mathbf{f}_s, \mathbf{e}_s) \in \mathcal{F}_s \times \mathcal{E}_s \mid \begin{pmatrix} q_2 e_2 \\ e_3 \end{pmatrix} \in H^1(\Omega)^2, \right.$$

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_R \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\partial_x(q_2 \cdot) & 0 \\ 0 & -D(q_2 \cdot) & -\partial_x(q_3 \cdot) - q_3 \partial_x & -I \\ 0 & 0 & I & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_R \end{pmatrix},$$

$$\left. \begin{pmatrix} f_a^B \\ e_a^B \end{pmatrix} = \begin{pmatrix} -q_2 & -q_3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e_2 \\ e_3 \end{pmatrix}\Big|_a, \quad \begin{pmatrix} f_b^B \\ e_b^B \end{pmatrix} = \begin{pmatrix} q_2 & q_3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e_2 \\ e_3 \end{pmatrix}\Big|_b \right\}, \tag{4.55}$$

is an extended Stokes-Dirac structure with respect to the symmetric pairing given by:

$$\ll \begin{bmatrix} \mathbf{f}_s \\ \mathbf{e}_s \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{f}}_s \\ \tilde{\mathbf{e}}_s \end{bmatrix} \gg = < \mathbf{f}_s \mid \tilde{\mathbf{e}}_s > + < \tilde{\mathbf{f}}_s \mid \mathbf{e}_s >, \quad \begin{bmatrix} \mathbf{f}_s \\ \mathbf{e}_s \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{f}}_s \\ \tilde{\mathbf{e}}_s \end{bmatrix} \in \mathscr{F}_s \times \mathscr{E}_s, \quad (4.56)$$

where the pairing $< \cdot \mid \cdot >$ is given in (4.54). Moreover, $D(q_2 \cdot)$ is the operator with domain all $e_2 \in \mathscr{L}^2(\Omega)$ such that $q_2 e_2 \in H^1(\Omega)$ and the extended action of the operator is

$$D(q_2 e_2) = \partial_x(q_2 e_2) - e_2 \partial_x q_2.$$

Remark 4.2.11 *We do not prove the Proposition 4.2.10. The proof of the corresponding extended Stokes-Dirac structure can be easily demonstrated by following the similar lines of reasoning as in the proof of the Proposition 4.2.9.*

Remark 4.2.12 *We have associated a particular choice of boundary port variables with an extended Stokes-Dirac structure. In principle, it would be ideal to derive an admissible set of boundary conditions in a parametrized way similar to [101], where a parametrization has been derived for a canonical skew-symmetric differential operator. However, the structures derived in this chapter/dissertation are non-canonical and hinder the elegant parametrization for the class of systems under discussion.*

4.2.3 Port-Hamiltonian representation

The preceding discussion showed that interconnection relations associated with the conservation laws define a so-called extended Stokes-Dirac structure. The interconnection structure, given by an extended Stokes-Dirac structure, together with the Hamiltonian functional representing the total energy of the system constitute a port-Hamiltonian representation. Following [53], we define the resulting port-Hamiltonian system as follows.

Definition 4.2.13 *A port-Hamiltonian system for the Drift Flux Model (without slip) with state-variables*

$$\tilde{q}(t) = \begin{bmatrix} \tilde{q}_1(t), & \tilde{q}_2(t), & \tilde{q}_3(t), & \tilde{q}_4(t) \end{bmatrix}^T, \qquad (4.57)$$

and port-variables

$$\begin{bmatrix} f_a^B(t), & f_b^B(t), & e_a^B(t), & e_b^B(t) \end{bmatrix}^T \in \mathscr{L}^2(\partial\Omega)^4, \qquad (4.58)$$

generated by the Hamiltonian functional (with smooth integrand) (4.20), and, with respect to the extended Stokes-Dirac structure (4.33) is defined by:

$$
\left(
\begin{bmatrix}
\partial_t \tilde{q}_1(t) \\
\partial_t \tilde{q}_2(t) \\
\partial_t \tilde{q}_3(t) \\
\partial_t \tilde{q}_4(t) \\
f_a^B(t) \\
f_b^B(t) \\
f_R
\end{bmatrix}
,
\begin{bmatrix}
\delta_{\tilde{q}_1} \mathcal{H}_t \\
\delta_{\tilde{q}_2} \mathcal{H}_t \\
\delta_{\tilde{q}_3} \mathcal{H}_t \\
\delta_{\tilde{q}_4} \mathcal{H}_t \\
e_a^B(t) \\
e_b^B(t) \\
\frac{32\mu_m \tilde{q}_1}{d^2} f_R
\end{bmatrix}
\right)
\in \mathcal{D}_t.
\tag{4.59}
$$

Remark 4.2.14 *Similarly, we can also define a port-Hamiltonian system for the single-phase model with respect to the extended Stokes-Dirac structure (4.55). We omit it for the sake of brevity.*

4.3 Port-Hamiltonian Modelling - temporal area variations

We now briefly consider temporal variations in the geometrical cross-section, i.e., $\partial_t A \neq 0$. We consider that the evolution of the area is described as:

$$
\partial_t A = r_1(t, \tilde{z}),
\tag{4.60}
$$

where r_1 is a function, which say is known a-priori or can be determined via some control law, and $\tilde{z}$ represents the vector of state variable.

Allowing for temporal variations in area can be viewed as the structure (with state variable $\tilde{z}$), which contains additional terms (relative to the structure with only spatial variations) that can be perceived as state and time-dependent control inputs. See the following theorem.

Theorem 4.3.1 *Consider the system Σ_t[1] governed by the combination of (2.2), (2.3) - (2.6) under the assumption of zero slippage, i.e., $v_g = v_\ell = v$, and (4.60). Then, it can be formulated in the dissipative Hamiltonian representation of the following form:*

$$
\partial_t \tilde{z} = (\mathcal{J}(\tilde{z}) - \mathcal{R}(\tilde{z}))\delta_{\tilde{z}} \mathcal{H}(\tilde{z}) - r(t, \tilde{z}).
\tag{4.61}
$$

---

[1]*The first equation of the composed system Σ_t is (4.60). The rest of the equations in the composed system are the mass and the momentum conservation laws.*

Remark 4.3.2 *In the scope of two-phase flow models, $\tilde{z} = \tilde{q}$, $\mathscr{J} = \mathscr{J}_t$, $\mathscr{R} = \mathscr{R}_t$ and $\mathscr{H} = \mathscr{H}_t$. Equivalently, the structure holds in the scope of single-phase models with corresponding state variables, interconnection (formal skew-adjoint) operator, resistive matrix and the Hamiltonian functional.*

The above structure can be viewed as a special case of the representation in [123]. If we ignore the boundary ports in the pHDAE definition of [123] and use slightly different notations for the sake of consistency in this chapter, then we obtain:

$$\mathscr{E}\dot{\tilde{z}} = (\mathscr{J}(\tilde{z}) - \mathscr{R}(\tilde{z}))s - r(t, \tilde{z}), \tag{4.62}$$

where $r(t, \tilde{z})$ is of the form: $[r_1(t, \tilde{z}) \quad 0 \quad 0 \quad 0]^T$. The reasoning behind the choice of this form is apparent[1].

We consider the mapping $\mathscr{E} = I$, $\partial_{\tilde{z}}\mathscr{H} = s$ and $\partial_t \mathscr{H} = s^T r$ and follow [123] to obtain the dissipation inequality.

$$\begin{aligned}
\frac{d\mathscr{H}}{dt} &= (\partial_z \mathscr{H})^T \dot{\tilde{z}} + \partial_t \mathscr{H} \\
&= s^T \Big((\mathscr{J} - \mathscr{R})s - r \Big) + s^T r \\
&= -s^T \mathscr{R} s \leq 0.
\end{aligned} \tag{4.63}$$

The structural representation as in (4.62) has already been shown to be a Dirac structure in [123]. Hence, we refer the reader to [123] for further details.

Remark 4.3.3 *Structure (4.61) or (4.62) has been presented in rather general sense. It is worth mentioning that a desirable structure is realizable for the models governing the single-phase and two-phase fluid flow across variable geometrical cross-section by using specific choice of state variables and the associated interconnection operator and the dissipation matrix.*

Having discussed the methodology to obtain a port-Hamiltonian formalism that accounts for variations in geometrical cross-section, we now turn the focus towards composing subsystems in an MPD configuration, i.e., the drill string, the drill bit, and the annulus, in a power-preserving manner.

4.4 Revisiting Port-Hamiltonian Models

We first define a global/local coordinate system that will be used in the scope of composing MPD subsystems. Figure 4.1 provides a simplified schematic of

an MPD system; refer to Chapter 2 for a detailed description. Here, x_d and x_a indicate the spatial location inside the drill string and the annulus, respectively. The (local) axes, corresponding to the drill string and the annulus are oriented as shown in red and blue, respectively, in Figure 4.1. The circulation path of the fluid(s) can be observed by following the green arrows in Figure 4.1. We recall that the flow inside the drill string is primarily of single-phase nature. The fluid exits the drill string and flows across the drill bit before passing into the annulus section of the wellbore. If there is no influx from the reservoir into the annulus, the flow inside the annulus also constitutes only a single-phase nature. A liquid and/or gaseous influx into the annulus changes the flow type from single-phase flow to multi-phase flow inside this part of the wellbore. In case of liquid influx, we will consider that the reservoir produces the same fluid as the one pumped inside the drill string via the pump shown in Figure 4.1 and, hence, only a single-phase flow model will be considered inside the annulus. The aforementioned assumption sounds restrictive in the scope of drilling applications. However, we stress that a liquid influx with potentially different properties than the circulating fluid can be dealt with in the same manner as the scenario of gas influx. In case of gas influx, different properties of the in-flowing gas and the circulating fluid are considered and, hence, the two-phase flow model is employed for studying the behaviour inside the annulus. A liquid and gas influx can also occur simultaneously. Such a situation demands employing three-phase or more complicated multi-phase flow models. However, such scenarios fall outside the scope of this dissertation and, hence, the focus will be on single- or two-phase flow models only.

Next, we briefly introduce the Isothermal Euler equations for the drill string and the Two-Fluid Model for the annulus along with their corresponding port-Hamiltonian formulations in Section 4.4.1 and Section 4.4.2, respectively.

4.4.1   Isothermal Euler Equations for the drill string

For a drill string with constant cross-sectional area A_d and a constant inclination θ (see Figure 4.1), the Isothermal Euler equations, under the assumption of laminar flow, read as follows:

$$\partial_t \rho + \partial_{x_d} \left(\rho v \right) = 0, \tag{4.64a}$$

$$\partial_t \left(\rho v \right) + \partial_{x_d} \left(\rho v^2 + p \right) = -32 \frac{\mu_\ell v}{d_d^2} + \rho g \sin\theta, \tag{4.64b}$$
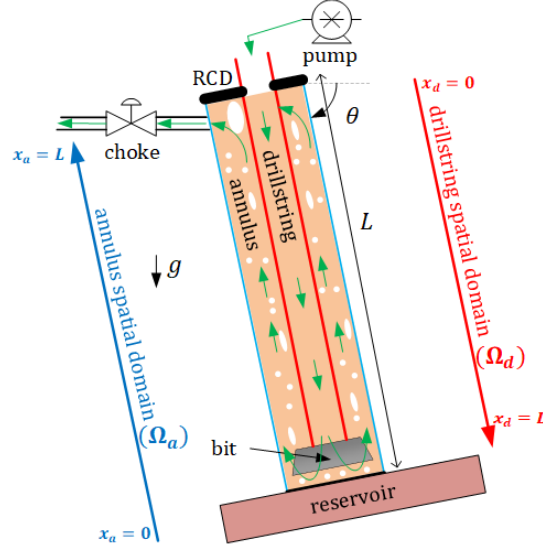
Figure 4.1: A simplified schematic of an MPD system.

where $t \in \mathbb{R}_{\geq 0}$ and x_d carries the meaning as mentioned earlier. Variables $\rho(t, x_d)$, $v(t, x_d)$ and $p(t, x_d)$ refer to density, velocity and pressure of the fluid inside the drill string, respectively. Moreover, μ_ℓ and d_d denote viscosity of the fluid and hydraulic diameter of the drill string, respectively, and g is the gravitational acceleration. As before, to complete the set of Isothermal Euler equations, an Equation of State (EOS) is provided as $p = (\rho - \rho_{\ell 0})c_\ell^2 + p_{\ell 0}$ with constants $\rho_{\ell 0}$ and $p_{\ell 0}$, and c_ℓ being the speed of sound in the fluid.

Remark 4.4.1 *The signs of the source terms in the right-hand side of* (4.64b) *are in accordance with the coordinate system for the drill string defined in Figure 4.1. It should be stressed that in Chapter 3 and earlier part of this chapter, gravitational source terms were associated with the "-" sign. All previous results still hold if "+" sign needs to be considered. Moreover, we emphasize that the gravitational source terms are automatically accounted for in the port-Hamiltonian representations and, hence, do not impact the structure of the (formally) skew-adjoint operator and the resistive matrix; see Theorem 3.3.10, Theorem 4.2.6 and Theorem 4.2.7 and the corresponding Hamiltonian functionals which included gravitational energy component.*

The Hamiltonian functional for (4.64) in terms of state variables $z := [\rho, v]^T$ is

$$\mathcal{H}(z) := A_d \int_{\Omega_d} \left(\rho \frac{v^2}{2} + \rho c_\ell^2 \ln \rho + (c_\ell^2 \rho_{\ell 0} - p_{\ell 0}) - \rho g x \sin \theta \right) dx. \qquad (4.65)$$

We now recall the port-Hamiltonian formulation for the above described model.

Theorem 4.4.2 *The governing equations (4.64) together with the EOS $p = (\rho - \rho_{\ell 0}) c_\ell^2 + p_{\ell 0}$ can be written in the following dissipative port-Hamiltonian formulation:*

$$\partial_t z = \left(\mathcal{J}_d - \mathcal{R}_d(z) \right) \delta_z \mathcal{H}(z), \qquad (4.66a)$$

where $z := [\rho, v]^T$, the Hamiltonian functional $\mathcal{H}(z)$ is given by (4.65),

$$\mathcal{J}_d = -\frac{1}{A_d} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{\partial}{\partial x_d}, \quad \mathcal{R}_d(z) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{32 \mu_\ell}{A_d d_d^2 \rho^2} \end{bmatrix}, \qquad (4.66b)$$

and the power conjugated input $\mathbf{u} := [u_{0,d}^B, u_{L,d}^B]^T$ and output $\mathbf{y} := [y_{0,d}^B, y_{L,d}^B]^T$ variables at the boundaries are as follows:

$$\begin{pmatrix} y_{L,d}^B \\ u_{L,d}^B \end{pmatrix} = \begin{pmatrix} \frac{1}{A_d} \delta_\rho \mathcal{H}(z) \\ \delta_v \mathcal{H}(z) \end{pmatrix} \Big|_{L,d}, \quad \begin{pmatrix} y_{0,d}^B \\ u_{0,d}^B \end{pmatrix} = \begin{pmatrix} -\frac{1}{A_d} \delta_\rho \mathcal{H}(z) \\ \delta_v \mathcal{H}(z) \end{pmatrix} \Big|_{0,d}. \qquad (4.67)$$

Here, $u_{0,d}^B$ and $u_{L,d}^B$ indicate the input variable at $x_d = 0$ and $x_d = L$, respectively. Similarly, $y_{0,d}^B$ and $y_{L,d}^B$ indicate the output variable at $x_d = 0$ and $x_d = L$, respectively. Also, $(\cdot)|_{L,d}$ and $(\cdot)|_{0,d}$ refer to evaluating the quantity within the argument at $x_d = L$ and $x_d = 0$, respectively.

Remark 4.4.3 *The proof to Theorem 4.4.2 can be obtained by following the lines of reasoning as that provided earlier for the derivation of Theorem 4.2.6 and Proposition 4.2.10.*

The physical interpretation of $\delta_v \mathcal{H}$ is the mass flow rate of the mud/fluid passing across the spatial location x_d at time t. The energy per unit mass multiplied by the cross section A_d can be inferred from $\delta_\rho \mathcal{H}$.

The power balance equation reads as follows:

$$\dot{\mathcal{H}} + \mathscr{P}^B = -\int_{\Omega_d} e_d^T \mathscr{R}_d \ e_d \ \mathrm{d}x + \left(\frac{1}{A_d}\delta_\rho\mathscr{H}(z)\delta_v\mathscr{H}(z)\right)|_{0,d} -$$
$$\left(\frac{1}{A_d}\delta_\rho\mathscr{H}(z)\delta_v\mathscr{H}(z)\right)|_{L,d} + \underbrace{y_{0,d}^B u_{0,d}^B + y_{L,d}^B u_{L,d}^B}_{\mathscr{P}^B}, \qquad (4.68)$$
$$= -\int_{\Omega_d} e_d^T \mathscr{R}_d \ e_d \ \mathrm{d}x \le 0,$$

where the last inequality is due to the symmetric and positive semi-definite nature of $\mathscr{R}_d(z)$. Here, $\dot{\mathcal{H}}$ refers to the time derivative of the Hamiltonian functional $\mathcal{H}$ along the solutions of the single-phase flow model and $\mathscr{P}^B :=$ $y_{0,d}^B u_{0,d}^B + y_{L,d}^B u_{L,d}^B$ indicates the power exchange through the boundaries of the drill string.

4.4.2   Two-Fluid Model for the annulus

In accordance with the introduced coordinate system for the annulus (referring to Figure 4.1), the TFM can be defined for each phase in the annulus with constant cross-sectional area A_a, constant hydraulic diameter d_a and constant inclination θ as follows [176]:

$$\partial_t \left(\alpha_i \rho_i\right) + \partial_{x_a}\left(\alpha_i \rho_i v_i\right) = 0, \qquad (4.69a)$$
$$\partial_t \left(\alpha_i \rho_i v_i\right) + \partial_{x_a}\left(\alpha_i \rho_i v_i^2\right) = -\partial_{x_a}\left(\alpha_i p\right) + M_i - \alpha_i \rho_i g \sin\theta - 32\frac{\mu_g \alpha_i v_i}{d_a^2}, \qquad (4.69b)$$

where $i \in \{\ell, g\}$ and x_a carries the meaning as introduced earlier. Also, the sense of gravity has been accounted for in the model description (4.69b). We refer the reader to Section 2.3.1 for detailed description of the TFM. To complete the model, we recall a set of widely applied closure laws as in [56]:

$$\alpha_g + \alpha_\ell = 1, \qquad (4.70a)$$
$$M_g + M_\ell = 0, M_g = p\partial_{x_a}\alpha_g + b_g^M(v_\ell - v_g), \qquad (4.70b)$$
$$\rho_g = \frac{p}{c_g^2}, \rho_\ell = \rho_{\ell 0} + \frac{p - p_{\ell 0}}{c_\ell^2}. \qquad (4.70c)$$

We also refer the reader to Section 2.3.1 for the physical interpretation of the algebraic closure laws (4.70).

The Hamiltonian for the flow inside the annulus with $z_a := [m_g, m_\ell, v_g, v_\ell]^T$ takes the following form:

$$\mathcal{H}_a(z_a) := A_a \int_{\Omega_a} \left(m_g \frac{v_g^2}{2} + m_\ell \frac{v_\ell^2}{2} + m_g c_g^2 \ln \rho_g + m_\ell c_\ell^2 \ln \rho_\ell + \right.$$

$$\left. \alpha_\ell (c_\ell^2 \rho_{\ell 0} - p_{\ell 0}) - (m_\ell + m_g) g(L-x) \sin\theta \right) dx. \quad (4.71)$$

Notably, variables $\rho_\ell, \rho_g, \alpha_\ell$ can be written in terms of m_g and m_ℓ, see [19]. In the following theorem, the port-Hamiltonian formulation corresponding to (4.69) and (4.70) is presented.

Theorem 4.4.4 *The governing equations* (4.69) *associated with the closure equations* (4.70) *can be written in the dissipative port-Hamiltonian formulation as follows:*

$$\partial_t z_a = \left(\mathcal{J}_a - \mathcal{R}_a(z_a) \right) \delta_{z_a} \mathcal{H}_a(z_a), \quad (4.72a)$$

where $z_a := [m_g, m_\ell, v_g, v_\ell]^T$ *and the Hamiltonian functional* $\mathcal{H}_a$ *is given by* (4.71),

$$\mathcal{J}_a = -\frac{1}{A_a} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \frac{\partial}{\partial x_a}, \quad \mathcal{R}_a(z_a) = \frac{1}{A_a} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{b_g^M}{m_g^2} + \frac{32\mu_g \alpha_g}{m_g^2 d^2} & -\frac{b_g^M}{m_g m_\ell} \\ 0 & 0 & -\frac{b_g^M}{m_g m_\ell} & \frac{b_g^M}{m_\ell^2} + \frac{32\mu_l \alpha_\ell}{m_l^2 d^2} \end{bmatrix},$$

$$(4.72b)$$

and the power conjugated inputs $\mathbf{u}_a := [u_{0,a}^{B,1}, u_{0,a}^{B,2}, u_{L,a}^{B,1}, u_{L,a}^{B,2}]^T$ *and power conjugated outputs* $\mathbf{y}_a := [y_{0,a}^{B,1}, y_{0,a}^{B,2}, y_{L,a}^{B,1}, y_{L,a}^{B,2}]^T$ *at the boundaries of the spatial domain* Ω_a *are as follows:*

$$\begin{pmatrix} y_{0,a}^{B,1} \\ y_{0,a}^{B,2} \\ u_{0,a}^{B,1} \\ u_{0,a}^{B,2} \end{pmatrix} = \begin{pmatrix} -\frac{1}{A_a} \delta_{m_g} \mathcal{H}_a \\ -\frac{1}{A_a} \delta_{m_\ell} \mathcal{H}_a \\ \delta_{v_g} \mathcal{H}_a \\ \delta_{v_\ell} \mathcal{H}_a \end{pmatrix} |_{0,a}, \quad \begin{pmatrix} y_{L,a}^{B,1} \\ y_{L,a}^{B,2} \\ u_{L,a}^{B,1} \\ u_{L,a}^{B,2} \end{pmatrix} = \begin{pmatrix} \frac{1}{A_a} \delta_{m_g} \mathcal{H}_a \\ \frac{1}{A_a} \delta_{m_\ell} \mathcal{H}_a \\ \delta_{v_g} \mathcal{H}_a \\ \delta_{v_\ell} \mathcal{H}_a \end{pmatrix} |_{L,a}. \quad (4.73)$$

Here, $(\cdot)_{0,a}^{B,i}$ *denotes the i-th component of the decomposed variable* $(\cdot)$ *at the boundary* $x_a = 0$. *Similarly,* $(\cdot)_{L,a}^{B,i}$ *denotes the i-th component of the decomposed variable*

(·) *at the boundary* $x_a = L$. *Also,* $(\cdot)|_{L,a}$ *and* $(\cdot)|_{0,a}$ *refer to evaluating the quantity within the argument at* $x_a = L$ *and* $x_a = 0$, *respectively.*

Remark 4.4.5 *The proof to Theorem 4.4.4 can be obtained by following the lines of reasoning as that provided for the derivation of Theorems 3.3.3 and 3.4.9.*

The variable $\delta_{v_i}\mathscr{H}_a$ represents the mass flow rate of the phase i. Similarly, the energy per unit mass of phase i multiplied by the cross-section A_a can be inferred from $\delta_{m_i}\mathscr{H}_a$.

The power balance equation reads as follows:

$$\dot{\mathscr{H}}_a + \mathscr{P}_a^B = -\int_{\Omega_a} e_a^T\, \mathscr{R}_a\, e_a\, \mathrm{d}x + \left(\frac{1}{A_a}\delta_{m_g}\mathscr{H}_a\delta_{v_g}\mathscr{H}_a\right)|_{0,a}+$$

$$\left(\frac{1}{A_a}\delta_{m_\ell}\mathscr{H}_a\delta_{v_\ell}\mathscr{H}_a\right)|_{0,a} - \left(\frac{1}{A_a}\delta_{m_g}\mathscr{H}_a\delta_{v_g}\mathscr{H}_a\right)|_{L,a}-$$

$$\left(\frac{1}{A_a}\delta_{m_\ell}\mathscr{H}_a\delta_{v_\ell}\mathscr{H}_a\right)|_{L,a} + \underbrace{y_{0,a}^{B,1}u_{0,a}^{B,1} + y_{0,a}^{B,2}u_{0,a}^{B,2} + y_{L,a}^{B,1}u_{L,a}^{B,1} + y_{L,a}^{B,2}u_{L,a}^{B,2}}_{\mathscr{P}_a^B}, \qquad (4.74)$$

$$= -\int_{\Omega_a} e_a^T\, \mathscr{R}_a\, e_a\, \mathrm{d}x \leq 0,$$

where the last inequality is due to the symmetric and positive semi-definite nature of $\mathscr{R}_a(z)$. Here, $\dot{\mathscr{H}}_a$ refers to the time derivative of the Hamiltonian functional $\mathscr{H}_a$ along the solutions of the TFM and $\mathscr{P}_a^B := y_{0,a}^{B,1}u_{0,a}^{B,1} + y_{0,a}^{B,2}u_{0,a}^{B,2} + y_{L,a}^{B,1}u_{L,a}^{B,1} + y_{L,a}^{B,2}u_{L,a}^{B,2}$ indicates the power exchange through the boundaries of the annulus.

4.5   Power-Preserving Interconnection under influx

In this section, the drill string, the drill bit and the annulus are connected in a power-preserving manner. First, the boundary conditions are introduced. Then, the dissipation of energy across the bit and the passivity of the composed system is studied.

4.5.1   Boundary conditions of the single- and two-phase flow models in MPD

The governing equations of the pump, bit and choke are summarized in Table 4.1. For the pump, $\dot{m}_p, A_p, \rho_p$ and v_p represent the mass flow rate, the cross-

sectional area, the density and the velocity of the liquid through the pump, respectively. For the bit, $\Delta p_b, \rho_b, \dot{m}_b, A_N$ and C_D denote the pressure drop over the bit, density at the drill string side of the bit, the mass flow rate through the bit, the total area of the nozzles of the bit and the nozzle coefficient, respectively. For the choke, $\dot{m}_c, \rho_c, p_c$ and p_0 are the mass flow rate, the density, the pressure at the choke inlet and atmospheric pressure, respectively. Finally, K_c and z_c, are the choke constant and the choke opening, respectively.

The boundary conditions governing the fluid flow in the drill string and the annulus follow from Figure 4.1. First, we define the boundary conditions for the drill string to be used in (4.67).

Pump: At the pump location, we have $A_p = A_d, v_p = v|_{0,d}$ and $\rho_p = \rho|_{0,d}$. Hence

$$\delta_v \mathcal{H}|_{0,d} = (\rho A_d v)|_{0,d} = \dot{m}_p(t). \tag{4.75}$$

The boundary condition at the left side of the spatial domain Ω_d is assigned. Input and output variables at this boundary can be defined with (4.75) and (4.67).

Bit and reservoir: Through the bit, the mass conservation holds and the pressure drop is governed by the bit equation. Moreover, the flow that passed through the bit is then mixed with the known liquid mass flow rate $\dot{m}_\ell(t)$ and the gaseous mass flow rate $\dot{m}_g(t)$ coming out of the reservoir. Then, the mixture enters the annulus. Considering $\Delta p_b = p|_{L,d} - p|_{0,a}, \dot{m}_b = (\rho A_d v)|_{L,d}, \rho_b = \rho|_{L,d}$, we have

$$\begin{cases} p|_{L,d} - p|_{0,a} = \dfrac{1}{2\rho|_{L,d}} \left(\dfrac{\delta_v \mathcal{H}|_{L,d}}{A_N C_D} \right)^2, \\ \delta_v \mathcal{H}|_{L,d} + \dot{m}_\ell(t) = \delta_{v_\ell} \mathcal{H}_a|_{0,a}, \\ \delta_{v_g} \mathcal{H}_a|_{0,a} = \dot{m}_g(t). \end{cases} \tag{4.76}$$

Remark 4.5.1 *To solve the TFM, typically either 2 boundary conditions are specified at the left side of the domain and 2 at the right side or 3 boundary conditions*

Table 4.1: Governing equations of the pump, bit and choke.

Equipment	Governing equation
pump	$\dot{m}_p = A_p \rho_p v_p$
bit	$\Delta p_b = \dfrac{1}{2\rho_b} \left(\dfrac{\dot{m}_b}{A_N C_D} \right)^2$
choke	$\dot{m}_c = K_c z_c \sqrt{2\rho_c (p_c - p_0)}$

at the left side and one at the right [65]. In this chapter, we consider the 2-2 case. In (4.76), one equation corresponds to the right boundary for the isothermal Euler equations in the drill string and two equations correspond to the left boundary of the TFM in the annulus.

Remark 4.5.2 *Above the bit and inside the drill string, a non-return valve is installed to restrict the flow in one direction only from the drill string to the annulus. When this valve is open, the pressure drop over the bit is governed by the bit equation in Table 4.1. When this valve is closed, the drill string and the annulus become disconnected.*

Choke: For the TFM at the choke, we have two boundary conditions, i.e., the explicit value of gas void fraction over time and the nonlinear choke equation. Following the same procedure in [130], we rewrite $\dot{m}_c$ as the mass flow rate of the mixture, $\dot{m}_c = (\rho_\ell A_a v_\ell)|_{L,a} + (\rho_g A_a v_g)|_{L,a}$. We also consider ρ_c as the mixture density, which is governed by: $\rho_c = \alpha_\ell \rho_\ell + \alpha_g \rho_g$. Therefore, we have:

$$(\delta_{v_\ell}\mathcal{H}_a + \delta_{v_g}\mathcal{H}_a)|_{L,a} = K_c z_c \sqrt{2(m_g + m_\ell)|_{L,a}(p_a|_{L,a} - p_0)},$$
$$\alpha_g|_{L,a} = g(t), \tag{4.77}$$

where $g(t)$ is a function of time, explicitly specifying the gas void fraction at the choke.

Remark 4.5.3 *Notably, boundary conditions (4.76) and (4.77) form an implicit function of variational derivative of Hamiltonian (4.65) and (4.71) with respect to z and z_a, respectively.*

4.5.2   Dissipativity of power through the bit

We have shown the power-preserving structure of the single-phase flow model and the TFM in Section 4.4 via the corresponding dissipative port-Hamiltonian formulation (4.66)-(4.67) and (4.72)-(4.73), respectively.

In this section, we consider the models for the drill string, drill bit and annulus to develop an aggregated port-Hamiltonian model for MPD systems. We also account for the interaction of the wellbore with the (liquid and gaseous) reservoir. Refer to Figure 4.2, which represents a schematic of the power-preserving interconnection of various subsystems of a drilling well. In this figure, the storage, resistive and boundary (external) ports corresponding to the drill string and the annulus are shown within the purple and red dashed
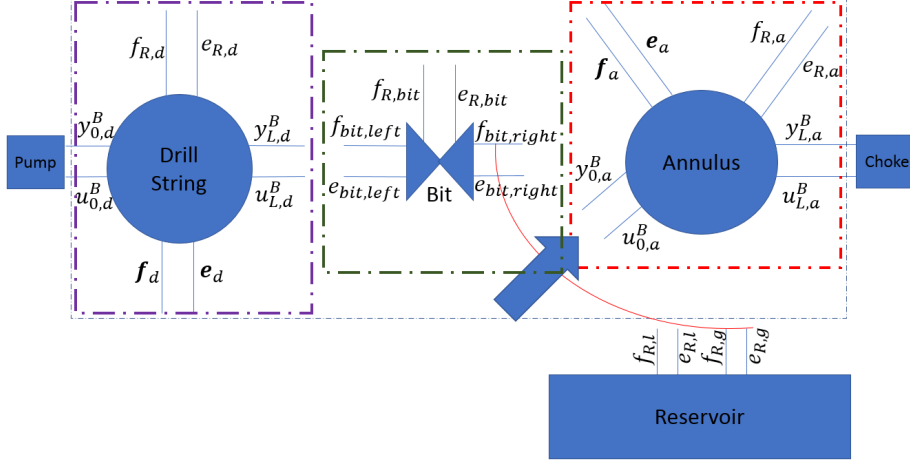
Figure 4.2: The power-preserving interconnection of different components of a drilling well.

rectangular box, respectively. Similarly, the resistive and boundary (external) ports corresponding to the bit are shown within the green dashed box. The composed MPD system is represented by the light blue rectangular box. We can also clearly see that the pump and the choke form the left and the right end of the composed system, respectively. The variables $y_{0,d}^B, y_{L,d}^B, u_{L,0}^B, u_{L,d}^B$ have already been introduced in Section 4.4. The notations $y_{0,a}^B, y_{L,a}^B, u_{0,a}^B, u_{L,a}^B$ stand for $y_{0,a}^B = [y_{0,a}^{B,1}, \ y_{0,a}^{B,2}]^T, y_{L,a}^B = [y_{L,a}^{B,1}, \ y_{L,a}^{B,2}]^T, u_{L,a}^B = [u_{0,a}^{B,1}, \ u_{0,a}^{B,2}]^T, u_{L,a}^B = [u_{L,a}^{B,1}, \ u_{L,a}^{B,2}]^T$. The variables $y_{0,a}^{B,1}, y_{0,a}^{B,2}, y_{L,a}^{B,1}, y_{L,a}^{B,2}, u_{0,a}^{B,1}, u_{0,a}^{B,2}, u_{L,a}^{B,1}, u_{L,a}^{B,2}$ have also been introduced in Section 4.4. The sub-systems (here, the drill string and annulus) interact with their environment via the external ports $(y_{0,d}^B, u_{0,d}^B)$, $(y_{0,a}^B, u_{0,a}^B)$, $(y_{L,d}^B, u_{L,d}^B)$ and $(y_{L,a}^B, u_{L,a}^B)$. The ports $(\mathbf{f}_d, \mathbf{e}_d)$ and $(\mathbf{f}_a, \mathbf{e}_a)$ are linked to the energy storage within the spatial domain of the drill string and the annulus, respectively, due to distributed effects. The ports $(f_{R,d}, e_{R,d})$ and $(f_{R,a}, e_{R,a})$ correspond to energy dissipation within the drill string and the annulus, respectively. The ports $(f_{R,l}, e_{R,l})$ and $(f_{R,g}, e_{R,g})$ correspond to the power transfer from the reservoir to the annular part of the wellbore due to the liquid and the gaseous influx, respectively. The scalar quantity $(y_{L,d}^B)^T u_{L,d}^B$ represents the (outgoing) instantaneous power at the right-hand side of the drill string. This power will now enter at

the left end of the drill bit. There is a natural relation[2] between $(y_{L,d}^{B})^{T} u_{L,d}^{B}$ and $(f_{bit,left})^{T} e_{bit,left}$, where the latter means the (incoming) power at the left end side of the drill bit. The scalar quantity $(f_{bit,right})^{T} e_{bit,right}$ denotes the (outgoing) instantaneous power from the right end of the bit, and eventually enters the annulus. And, the port $(f_{R,bit}, e_{R,bit})$ corresponds to the dissipated energy due to the dissipative nature of the drilling bit. It can also be observed that the instantaneous power entering the annulus is coming from various sources, i.e., from across the bit and the reservoir. This aspect is indicated via an arrow across a red circular section in Figure 4.2.

Invoking the notion of power-preserving interconnection for the composition, the following relations naturally hold:

$$(y_{L,d}^{B})^{T} u_{L,d}^{B} + (f_{bit,left})^{T} e_{bit,left} = 0, \qquad (4.78)$$

and

$$(f_{bit,right})^{T} e_{bit,right} + (f_{R,l})^{T} e_{R,l} + (f_{R,g})^{T} e_{R,g} + (y_{0,a}^{B})^{T} u_{0,a}^{B} = 0. \qquad (4.79)$$

Furthermore, for power-preserving nature across the bit, the following relation:

$$(f_{bit,left})^{T} e_{bit,left} + (f_{R,bit})^{T} e_{R,bit} + (f_{bit,right})^{T} e_{bit,right} = 0, \qquad (4.80)$$

must hold.

Using (4.79) and (4.73), and exploiting the scalar nature of some variables, we have:

$$f_{bit,right} e_{bit,right} + f_{R,l} e_{R,l} + f_{R,g} e_{R,g} = \left(\frac{(\delta_{m_g} \mathcal{H}_a) \delta_{v_g} \mathcal{H}_a}{A_a} + \frac{(\delta_{m_\ell} \mathcal{H}_a) \delta_{v_\ell} \mathcal{H}_a}{A_a} \right) |_{x_a=0}. \qquad (4.81)$$

Similarly, using (4.78) and (4.67), and exploiting the scalar nature of some variables, we have:

$$f_{bit,left} e_{bit,left} = - \frac{(\delta_{\rho} \mathcal{H}) \delta_{v} \mathcal{H}}{A_d} |_{x_d=L}. \qquad (4.82)$$

Using (4.80), (4.81) and (4.82), we have:

$$- \frac{(\delta_{\rho} \mathcal{H}) \delta_{v} \mathcal{H}}{A_d} |_{x_d=L} + f_{R,bit} e_{R,bit} + \left(\frac{(\delta_{m_g} \mathcal{H}_a) \delta_{v_g} \mathcal{H}_a}{A_a} + \frac{(\delta_{m_\ell} \mathcal{H}_a) \delta_{v_\ell} \mathcal{H}_a}{A_a} \right) |_{x_a=0}$$
$$- f_{R,l} e_{R,l} - f_{R,g} e_{R,g} = 0. \qquad (4.83)$$

---

[2] The proper relation will be made further apparent technically in the subsequent discussion.

To ensure the dissipative property across the bit, $f_{R,bit}e_{R,bit} \geq 0$ must hold. As a result of this property and by using (4.83), we have:

$$\frac{(\delta_\rho \mathcal{H})\delta_v \mathcal{H}}{A_d}|_{x_d=L} - \left(\frac{(\delta_{m_g} \mathcal{H}_a)\delta_{v_g} \mathcal{H}_a}{A_a} + \frac{(\delta_{m_\ell} \mathcal{H}_a)\delta_{v_\ell} \mathcal{H}_a}{A_a}\right)|_{x_a=0} + f_{R,l}e_{R,l} + f_{R,g}e_{R,g} \geq 0.$$
$$(4.84)$$

While $\delta_\rho \mathcal{H}, \delta_v \mathcal{H}, \delta_{m_g} \mathcal{H}_a, \delta_{v_g} \mathcal{H}_a, \delta_{m_\ell} \mathcal{H}_a, \delta_{v_\ell} \mathcal{H}_a$ are known in terms of physical quantities, the power exchange between the reservoir and the wellbore, $f_{R,l}e_{R,l} + f_{R,g}e_{R,g}$, is not yet explicitly known in terms of the physical variables. It should be mentioned that, in this dissertation, we have not undertaken the aspect of developing a port-Hamiltonian model of the reservoir. Hence, we do not have a-priori knowledge on the boundary ports $(f_{R,l}, e_{R,l})$ and $(f_{R,g}, e_{R,g})$. In order to assess whether (4.84) holds universally or conditionally, we impose some simplifications. Motivated from the form of the power conjugated inputs and outputs in (4.73), we assume that:

$$f_{R,g}e_{R,g} = \frac{(\delta_{m_g} \mathcal{H}_a)\delta_{v_g} \mathcal{H}_a}{A_a}|_{x_a=0}. \qquad (4.85)$$

Under the above assumption, (4.84) simplifies to:

$$\frac{(\delta_\rho \mathcal{H})\delta_v \mathcal{H}}{A_d}|_{x_d=L} - \left(\frac{(\delta_{m_g} \mathcal{H}_a)\delta_{v_g} \mathcal{H}_a}{A_a} + \frac{(\delta_{m_\ell} \mathcal{H}_a)\delta_{v_\ell} \mathcal{H}_a}{A_a}\right)|_{x_a=0} + f_{R,l}e_{R,l} +$$
$$\frac{(\delta_{m_g} \mathcal{H}_a)\delta_{v_g} \mathcal{H}_a}{A_a}|_{x_a=0} \geq 0, \quad (4.86)$$

which when further simplified yields:

$$\frac{(\delta_\rho \mathcal{H})\delta_v \mathcal{H}}{A_d}|_{x_d=L} - \left(\frac{(\delta_{m_\ell} \mathcal{H}_a)\delta_{v_\ell} \mathcal{H}_a}{A_a}\right)|_{x_a=0} + f_{R,l}e_{R,l} \geq 0. \qquad (4.87)$$

If $f_{R,l}e_{R,l} = \frac{\delta_{m_\ell} \mathcal{H}_a}{A_a}\dot{m}_\ell|_{x_a=0}$, then the relation (4.87) takes the following form:

$$\frac{(\delta_\rho \mathcal{H})\delta_v \mathcal{H}}{A_d}|_{x_d=L} - \left(\frac{(\delta_{m_\ell} \mathcal{H}_a)\delta_{v_\ell} \mathcal{H}_a}{A_a}\right)|_{x_a=0} + \frac{\delta_{m_\ell} \mathcal{H}_a}{A_a}\dot{m}_\ell|_{x_a=0} \geq 0. \qquad (4.88)$$

We can express $\dot{m}_\ell$ in terms of known physical quantities in the following manner:

$$\dot{m}_\ell = (\delta_{v_\ell} \mathcal{H}_a|_{x_a=0} - \delta_v \mathcal{H}|_{x_d=L}). \qquad (4.89)$$

Using (4.89) in (4.88), we have:

$$\frac{(\delta_\rho\mathcal{H})\delta_v\mathcal{H}}{A_d}|_{x_d=L} - \left(\frac{(\delta_{m_\ell}\mathcal{H}_a)\delta_{v_\ell}\mathcal{H}_a}{A_a}\right)|_{x_a=0} + \frac{\delta_{m_\ell}\mathcal{H}_a}{A_a}(\delta_{v_\ell}\mathcal{H}_a|_{x_a=0} - \delta_v\mathcal{H}|_{x_d=L}) \geq 0,$$

(4.90)

which under simplification becomes:

$$\delta_v\mathcal{H}|_{x_d=L}\left(\frac{\delta_\rho\mathcal{H}}{A_d}|_{x_d=L} - \frac{\delta_{m_\ell}\mathcal{H}_a}{A_a}|_{x_a=0}\right) \geq 0.$$

(4.91)

As $\delta_v\mathcal{H}|_{x=L} \geq 0$, we need to ensure the following condition for the power-preserving interconnection:

$$\left(\frac{\delta_\rho\mathcal{H}}{A_d}|_{x_d=L} - \frac{\delta_{m_\ell}\mathcal{H}_a}{A_a}|_{x_a=0}\right) \geq 0.$$

(4.92)

Using the definition of the variational derivative of the Hamiltonian functionals with respect to (desired) states, (4.92) reduces to:

$$(\frac{v^2}{2} + c_\ell^2\ln\rho - gL\sin\theta)|_{x_d=L} - (\frac{v_\ell^2}{2} + c_\ell^2\ln\rho_\ell - gL\sin\theta)|_{x_a=0} \geq 0,$$

(4.93)

which can be further simplified to

$$\frac{1}{2}M_d^2(-\frac{M_a^2}{M_d^2} + 1) + \ln\frac{\rho|_{x_d=L}}{\rho_\ell|_{x_a=0}} \geq 0,$$

(4.94)

where $M_d = \frac{v|_{x_d=L}}{c_\ell}, M_a = \frac{v_\ell|_{x_a=0}}{c_\ell}$ are the Mach numbers of the flow at the outlet of the drill string and at the inlet of the annulus near the bit. To further simplify the relation (4.94), we use the bit equation,

$$\Delta p_b = (\rho|_{x_d=L} - \rho_\ell|_{x_a=0})c_\ell^2 = \frac{1}{2}\rho|_{x_d=L}(\frac{A_d}{A_N C_D})^2 v^2|_{x_d=L} \rightarrow$$

(4.95a)

$$\frac{\rho_\ell|_{x_a=0}}{\rho|_{x_d=L}} = 1 - \frac{1}{2}(\frac{A_d}{A_N C_D})^2 M_d^2,$$

(4.95b)

The mass conservation across the bit (4.76) can also be simplified to

$$(A_d\rho v)|_{x_d=L} + \dot{m}_\ell(t) = (A_a\rho_\ell v_\ell)|_{x_a=0} \rightarrow A_d M_d + \frac{\dot{m}_\ell(t)}{\rho|_{x_d=L}c_\ell} = A_a\frac{\rho_\ell|_{x_a=0}}{\rho|_{x_d=L}}M_a.$$

(4.96)

Using (4.95), we can further simplify (4.96) and obtain:

$$M_a = \frac{A_d}{A_a} \frac{M_d}{1 - \frac{1}{2}(\frac{A_d}{A_N C_D})^2 M_d^2} + M_{\ell_r}, \tag{4.97}$$

where $M_{\ell_r} := \frac{\dot{m}_\ell(t)}{A_a c_\ell \rho_\ell|_{x_a=0}}$ is the Mach number at the interface of the reservoir and annulus. Finally by using (4.97), the inequality (4.94) simplifies to

$$\mathscr{M} := \frac{1}{2} M_d^2 ((\frac{A_d}{A_a} \frac{1}{1 - \frac{1}{2}(\frac{A_d}{A_N C_D})^2 M_d^2} + \frac{M_{\ell_r}}{M_d})^2 - 1) + \ln(1 - \frac{1}{2}\frac{A_d}{A_N C_D})^2 M_d^2) \leq 0. \tag{4.98}$$

For the bit, connecting the drill string and the annulus, to be power-preserving and physically dissipative, the inequality (4.98) should hold.

We now discuss the power balance of the system composed of the drill string, bit and the annulus. We recall that the power balance across the drill string is given by:

$$\dot{\mathscr{H}} + \mathscr{P}^B + \int_{\Omega_d} e_d^T \mathscr{R}_d e_d \, dx = 0, \tag{4.99}$$

also see (4.68). Similarly, as in (4.74), the power balance across the annulus is given by:

$$\dot{\mathscr{H}}_a + \mathscr{P}_a^B + \int_{\Omega_a} e_a^T \mathscr{R}_a e_a \, dx = 0. \tag{4.100}$$

The power balance across the composed system is the summation of the power balance equation across each individual sub-systems. Using (4.99), (4.100), and (4.80), we have:

$$\mathscr{P}_{tot} := \dot{\mathscr{H}} + \mathscr{P}^B + \int_{\Omega_d} e_d^T \mathscr{R}_d e_d \, dx + \dot{\mathscr{H}}_a + \mathscr{P}_a^B + \int_{\Omega_a} e_a^T \mathscr{R}_a e_a \, dx +$$
$$(f_{bit,left})^T e_{bit,left} + (f_{R,bit})^T e_{R,bit} + (f_{bit,right})^T e_{bit,right}, \quad (4.101)$$

which, upon using (4.81) and (4.82), can be written as:

$$\mathscr{P}_{tot} = \dot{\mathscr{H}} + y_{0,d}^B u_{0,d}^B + y_{L,d}^B u_{L,d}^B + \int_{\Omega_d} e_d^T \mathscr{R}_d e_d \, dx + \dot{\mathscr{H}}_a + y_{0,a}^{B,1} u_{0,a}^{B,1} + y_{0,a}^{B,2} u_{0,a}^{B,2} +$$

$$y_{L,a}^{B,1} u_{L,a}^{B,1} + y_{L,a}^{B,2} u_{L,a}^{B,2} + \int_{\Omega_a} e_a^T \mathscr{R}_a e_a \, dx - \frac{(\delta_\rho \mathscr{H}) \delta_v \mathscr{H}}{A_d}|_{x_d=L} + (f_{R,bit})^T e_{R,bit} +$$

$$\left(\frac{(\delta_{m_g} \mathscr{H}_a) \delta_{v_g} \mathscr{H}_a}{A_a} + \frac{(\delta_{m_\ell} \mathscr{H}_a) \delta_{v_\ell} \mathscr{H}_a}{A_a}\right)|_{x_a=0} - (f_{R,l}^T e_{R,l} + f_{R_g}^T e_{R_g}). \quad (4.102)$$

Furthermore using (4.68) and (4.74), (4.102) simplifies to

$$\mathscr{P}_{tot} = \dot{\mathscr{H}} + y_{0,d}^{B} u_{0,d}^{B} + \int_{\Omega_d} e_d^T \mathscr{R}_d e_d \ \mathrm{d}x + \dot{\mathscr{H}}_a + y_{L,a}^{B,1} u_{L,a}^{B,1} + y_{L,a}^{B,2} u_{L,a}^{B,2} +$$

$$\int_{\Omega_a} e_a^T \mathscr{R}_a e_a \ \mathrm{d}x + (f_{R,bit})^T e_{R,bit} - (f_{R,l}^T e_{R,l} + f_{R_g}^T e_{R_g}). \quad (4.103)$$

Defining the power exchange at the boundaries of the composed system by $\mathscr{P}_T^B$, we can say that:

$$\mathscr{P}_T^B = y_{0,d}^{B} u_{0,d}^{B} + y_{L,a}^{B,1} u_{L,a}^{B,1} + y_{L,a}^{B,2} u_{L,a}^{B,2} - (f_{R,l}^T e_{R,l} + f_{R_g}^T e_{R_g}). \quad (4.104)$$

Using (4.104), we can write (4.103) as:

$$\mathscr{P}_{tot} = \dot{\mathscr{H}} + \dot{\mathscr{H}}_a + \mathscr{P}_T^B + \underbrace{\int_{\Omega_d} e_d^T \mathscr{R}_d e_d \ \mathrm{d}x + \int_{\Omega_a} e_a^T \mathscr{R}_a e_a \ \mathrm{d}x + (f_{R,bit})^T e_{R,bit}}_{\mathscr{R}_{tot}}. \quad (4.105)$$

The right-hand side of (4.105) is required to be zero as it is actually the sum of the right-hand sides of (4.99), (4.100), and (4.80) and, hence, we have

$$\dot{\mathscr{H}} + \dot{\mathscr{H}}_a + \mathscr{P}_T^B + \underbrace{\int_{\Omega_d} e_d^T \mathscr{R}_d e_d \ \mathrm{d}x + \int_{\Omega_a} e_a^T \mathscr{R}_a e_a \ \mathrm{d}x + (f_{R,bit})^T e_{R,bit}}_{\mathscr{R}_{tot}} = 0. \quad (4.106)$$

The composed system will be passive if $\mathscr{R}_{tot} \geq 0$. Noting that $\int_{\Omega_d} e_d^T \mathscr{R}_d e_d \ \mathrm{d}x \geq 0$ (see (4.68)) and $\int_{\Omega_a} e_a^T \mathscr{R}_a e_a \ \mathrm{d}x \geq 0$ (see (4.74)), we have:

$$\int_{\Omega_d} e_d^T \mathscr{R}_d e_d \ \mathrm{d}x + \int_{\Omega_a} e_a^T \mathscr{R}_a e_a \ \mathrm{d}x \geq 0. \quad (4.107)$$

As a result of the nature of $\int_{\Omega_d} e_d^T \mathscr{R}_d e_d \ \mathrm{d}x + \int_{\Omega_a} e_a^T \mathscr{R}_a e_a \ \mathrm{d}x$, we easily observe that $\mathscr{R}_{tot}$ can be greater than or equal to zero even if $(f_{R,bit})^T e_{R,bit}$ is negative. However, in the absence of the dissipation across the drill string and the annulus, clearly, the passivity of the composed system reduces to the condition of the dissipativity across the drill bit. As a consequence, the passivity of the system (in the absence of dissipation) holds if (4.98) is satisfied. It can, hence, be concluded that satisfying (4.98) provides the conservative bound on the passivity of the composed system. However, the nature of $\mathscr{R}_{tot}$ should be assessed for the (true) passivity of the composed MPD system. Due to the complicated nature of

$\mathcal{R}_{tot}$, a proper (structure-preserving) numerical tooling is required to assess the its positivity. In the rest of this chapter, we will base the satisfaction of (4.98) on a practical example of a drilling well.

Remark 4.5.4 *It could be compelling to assess whether the conditional dissipativity across the drill bit is related with multi-phase flows. It is easy to repeat the steps in the scope of the single-phase flow model as well. It is observed that the dissipativity property across the drill bit again holds conditionally.*

4.6 Numerical examples

In this section, a real drilling well is considered and the region where the inequality (4.98) holds is investigated. The corresponding geometry and bit property are studied to define the region across which the dissipativity property holds.

Remark 4.6.1 *In drilling operations, the velocity inside the drill string is typically around 1 m/s while the speed of sound in the mud is around 1000 m/s. Therefore, for drilling applications, $M_d \approx 0.001$.*

The geometry and equipment properties of the drilling platform are given by

$$d_d = 76.2 \text{ mm}, d_{od} = 241.3 \text{ mm}, d_w = 444.5 \text{ mm},$$
$$A_N = 1418.7 \text{ mm}^2, C_D = 0.8, \tag{4.108}$$

where d_{od} and d_w are, respectively, the outer diameter of the drill string and the diameter of the wellbore.

For this well and this drill bit, to render the argument inside the logarithmic function in (4.98) to be positive, we observe that $M_d < 0.35$ should hold. As shown in the top part of Figure 4.3, in this restricted region for M_d with no flow from the reservoir (which is true in the normal drilling scenarios), (4.98) always holds for $\dot{m}_\ell = 0$ for the drilling well under consideration and the bit model is indeed dissipative (power-preserving). This might be the experimental condition under which the model for the bit was derived.

In the case of contingencies, where the fluid of the reservoir flows into the annulus, condition (4.98) is not always satisfied in the restricted region for M_d, as shown in the bottom part of Figure 4.3. When the reservoir also contains liquid, the velocity of this flow should be less than the velocity of the flow coming
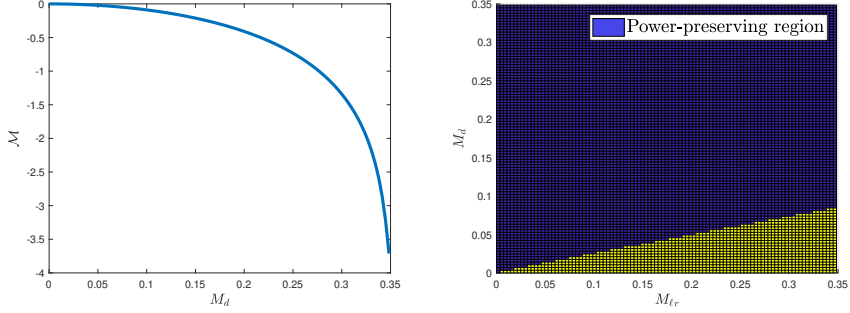
Figure 4.3: Left: The value of the function $\mathscr{M}$ in (4.98) for different admissible Mach number with $\dot{m}_\ell = 0$, Right: Power-preserving region for different Mach numbers of M_d and $M_{\ell r}$ satisfying inequality (4.98).

through the bit. This situation most probably occurs when the drilling process and the mud injection are halted ($M_d = 0$) and a new pipe section is added to the drill string to increase its length to drill further (this is called a *connection* scenario in practice). If the reservoir is producing liquid during connection, this inequality does not always hold. Therefore, the bit model presented in Table 4.1 needs to be used with proper care to simulate the hydraulics in this situation. At this stage, it is important to mention that the dissipativity of the drill bit, standalone, is of less practical importance. The passivity of the composed system is the quantity of significant interest. We also recall that the dissipation effects are usually present in the drill string and the annulus, and that the passivity of the composed system could still hold even if (4.98) is violated. A more rigorous numerical check is advised/required before using the bit model to simulate any hydraulic operations, particularly the ones where (4.98) is violated. Notably, in cases when velocity of the flow from the reservoir is much higher than the velocity of the flow passing through the bit, usually the non-return valve is closed and the two subsystems become isolated. This situation, however, requires more investigation. The bit models are usually derived by curve fitting to experimental data obtained under certain conditions. To adapt the bit model, if required, experiments should be designed in such a way that $\mathscr{R}_{tot}$ (or the inequality (4.98)) is violated and a new model should be fitted to the new data.

4.7   Conclusions

One of the main results of this chapter is the dissipative Hamiltonian realization and definition of extended state-dependent Stokes-Dirac structure. Consequently, we obtained port-Hamiltonian representations for both single-phase and two-phase models (here, the Drift Flux Model) governing fluid flow across spatially and temporally varying cross-section. It should be mentioned that even though we focused on the Drift Flux Model, the Two-Fluid Model across a variable cross-section can also be formulated in a port-Hamiltonian representation. A key observation, based on our findings, is that a special treatment is required to formulate port-Hamiltonian representations (in terms of non-conservative state variables) of single- and two-phase flow models across a variable geometrical cross-section. For instance, we need to invoke the concept of port-Hamiltonian descriptor systems. Such a framework is also perceived to be useful in order to develop (descriptor) port-Hamiltonian representations without having to eliminate the auxilliary variables. We defer this aspect to future works. Furthermore, similar to Chapter 3, we have adopted a non-unique way to formulate boundary port variables. The parametrization of boundary port variables in the scope of the obtained port-Hamiltonian representations deserves a further study. Moreover, we need to stress that we considered smooth variations in geometrical cross-sections, although non-smooth variations in geometrical cross-sections is admissible in real drilling scenarios. However, such variations in geometrical cross-sections complicate the study of Stokes-Dirac structures in a functional analytic setting that largely is built upon the fundamental lemma of calculus of variations. Hence, the consideration of non-smooth variations in the cross-section needs further investigation.

Another main result is hinged around the compositional port-Hamiltonian modelling for Managed Pressure Drilling (MPD) systems. We interconnect the (single- and two-phase) flow dynamic models and the drill bit model in a power-preserving manner, while also accounting for the interaction between the reservoir and the wellbore. It is observed that the mathematical model of the bit, used to interconnect the drill string and the annulus, obeys dissipativity under a certain condition. However, this conditional dissipativity does not restrict the normal drilling operation region. The drill bit model may restrict the drilling operation where liquid influx from reservoir flows into the wellbore. A rigorous numerical analysis of $\mathscr{R}_{tot}$ is required in such situations. While we observe that the passivity of the composed system holds conditionally, it is also interesting to find whether the composed system is incrementally passive, if not passive. However, the aspect of incremental passivity is not pursued in this dissertation. It

should be emphasized that we considered constant cross-section in the analysis of the behaviour at the bit junction and the passivity of the composed system. However, a similar analysis can straightforwardly be carried out even for the setting with varying cross-sections. The essential component are the boundary port variables, which have already been obtained for varying geometrical shapes as well (see Section 4.2 and 4.3). We restricted our attention to the case where two boundary conditions are specified at the left-hand side of the annulus and other two physical boundary conditions are specified at the right-hand side of the annulus. Other type of physical boundary conditions, such as the 3-1 case as opposed to 2-2 case, deserve a further dedicated study. This aspect, however, does not seem to pose serious concerns.

In this chapter, we have not delved deeper into the aspects of composed Dirac structure for the MPD system. It is well-known that the composition of Dirac structures is also a Dirac structure. We have already proved the existence of Dirac structures for the single- and two-phase flow models. The only additional investigation required, in the scope of composed MPD system, is to assess whether the finite-dimensional interconnection at the bit junction is indeed a Dirac structure and under what conditions. Connecting a finite-dimensional component with an infinite-dimensional Dirac structure will again be a Dirac structure. In the light of this discussion, we would like to provide insights regarding the conditional existence of a finite-dimensional Dirac structure at the bit junction. This conditional existence is closely tied to the conditional passivity of the composed system. It can be argued that, overall, more conditions than those required for the existence of Dirac structure (individually) for fluid dynamics models, would be needed. A rather technical in-depth discussion is deferred to future works.

The framework proposed in this chapter enables an energy-based controller design for MPD while taking the infinite-dimensional nature of the dynamics into account. In the pursuit towards developing a controller, one of the options could be to develop an infinite-dimensional controller and then discretize the resulting closed-loop plant via a structure-preserving scheme. Another option could be to discretize the model in a structure-preserving manner, and then develop a finite-dimensional controller. In any case, discretizing a port-Hamiltonian model in a structure-preserving manner deserves important attention, and this forms the focus of the next chapter.

Chapter 5

Structure-preserving discretization of two-phase flow models

In this chapter, we present a structure-preserving spatial discretization method for infinite-dimensional non-linear port-Hamiltonian representations of commonly used one-dimensional two-phase flow models: the Two-Fluid Model (TFM) and the Drift Flux Model (DFM). We introduce the port-Hamiltonian representations of these two-phase flow models and then invoke a mixed finite element method to perform a structure-preserving spatial discretization, both in the presence and absence of dissipative/resistive effects. Consequently, we obtain a finite-dimensional realization of a proposed novel (extended) Stokes-Dirac structure for the models of interest. The properties of the resulting finite-dimensional realizations are assessed and the conditions under which these are known to respect the properties of finite-dimensional Dirac structures are discussed. Furthermore, we derive the complete finite-dimensional interconnected port-Hamiltonian model by invoking the notion of power-preserving interconnection. Moreover, we employ the concept of Gauss-Legendre collocation methods for temporal discretization of state-dependent infinite-dimensional port-Hamiltonian representations of two-phase flow models. We extend the collocation concept, a definition based on symplectic integration, by accounting for state-dependent system matrices and feedthrough terms in the state-space port-Hamiltonian representation. Consequently, we develop a framework for

structure-preserving temporal discretization of the TFM and the DFM. Finally, we test the proposed structure-preserving discretization method on a simple, yet representative test case.

5.1   Introduction

In Chapters 3 and 4, we developed infinite-dimensional port-Hamiltonian representations for the single- and two-phase flow models. In this chapter, we will develop numerical algorithms that preserve a port-Hamiltonian structure, in particular for a class of two-phase flow models, at the continuous-time and discrete-time level.

Structure-preserving algorithms (or mimetic methods) are defined as *numerical algorithms that conserve as much as possible the qualitative behaviour of the original (finite- or infinite-dimensional) system*. Such algorithms have been used for solving ordinary differential equations, partial differential equations and differential algebraic equations in [43,54,123,125]. Structure-preservation carries a different meaning for different research communities. It could mean preservation of several (if not all) important features of a (continuous) problem, such as positivity of a physical quantity, local or global mass and momentum conservation or preservation of structural invariants/Casimirs, preservation of steady-state behaviour and limit-cycle behaviour, control of entropy production/dissipation rate, mimicking the decay behaviour of the physical energy, and preserving the index at the discrete level [17, 39, 40, 80]. Structure-preserving numerical algorithms could also encompass preservation of geometrical and algebraic structure, for instance Poisson structure, underlying the phase space of Hamiltonian dynamical systems [82] and preservation of the saddle point nature of the problem. Recently, numerical algorithms preserving the (multi-)symplectic structure, Hamiltonian structure, port-Hamiltonian structure and port-Hamiltonian descriptor structure have also been investigated [32,63,123].

In the past, several scientific applications have been formulated in the (finite- and infinite-dimensional) port-Hamiltonian framework [13, 19, 34, 115, 207] and, subsequently, spatially discretized, temporally discretized or both spatially and temporally discretized in a structure-preserving manner [20, 48, 75, 96, 163]. Several methodologies exist for discretization of finite-dimensional port-Hamiltonian systems, i.e., explicit discrete-time port-Hamiltonian systems [96] and discrete-time port-Hamiltonian descriptor realizations [123]. In the scope of structure-preserving discretization of infinite-dimensional port-Hamiltonian systems, the spatial discretization and temporal discretization of the underlying

geometry of the model forms a key step.

We are interested in the (complete) structure-preserving discretization of port-Hamiltonian representations of Managed Pressure Drilling (MPD) systems [130], which are composed of compressible single- and two-phase flow models. To this end, a structure-preserving finite-dimensional realization is required both for the single- and two-phase flow models. The structure-preserving discretization has already been performed for the single-phase flow models in [48, 95]. Moreover, [95] considered (compressible) single-phase flow in a pipe segment and also accounted for several practically relevant scenarios, such as a series interconnection of two and N-pipe segments while also taking junction conditions into consideration. However, the structure-preserving discretization of the two-phase flow models, popularly modelled by the Two-Fluid Model (TFM) and the Drift Flux Model (DFM), is still an open challenge, and we seek to address it in this chapter. The development of a structure-preserving discretization framework for compressible two-phase flow models and the existing structure-preserving discretization framework for compressible single-phase flow models will serve as a stepping stone to obtain a structure-preserving discretization and, ultimately, also to control design of MPD systems.

In Chapters 3 and 4 (or [19, 21]), a port-Hamiltonian formulation has been defined (implicitly in terms of *state-dependent* Stokes-Dirac structures) for single-phase and two-phase models for flows across constant and variable geometrical cross-sections. The structure-preserving discretization of such *state-dependent* non-linear distributed-parameter port-Hamiltonian systems is a key step for the simulation purposes and for the observer and controller design based on energy-shaping and damping-injection principles. Several works, as in [75, 95, 126], have performed discretization of *state-independent* Stokes-Dirac structures. However, there is limited work in the direction of discretization of *state-dependent* Stokes-Dirac structures in a structure-preserving manner. We develop structure-preserving numerical methods for the class of (conditional) hyperbolic systems of conservation laws that give rise to *state-dependent* Stokes-Dirac structures. In particular, the aim of this work is to develop finite-dimensional port-Hamiltonian formulations for the two variants of two-phase flow models that possess a *state-dependent* underlying geometrical structure.

Port-Hamiltonian formulations are generally represented implicitly in terms of Stokes-Dirac structures [21, 53]. These underlying geometric structures are infinite-dimensional Dirac structures based on Stokes' theorem [19, 53, 101]. These are known to play a crucial role in characterizing the structural invariants of a port-Hamiltonian system [53]. Moreover, we show, for the first time to the best of our knowledge, that such structures also essentially help us to

obtain appropriate/compatible approximation spaces (of the flow, state and the effort variables), and thus eventually to obtain a finite-dimensional Dirac/port-Hamiltonian structure at the discrete level. It is well known that the resistive effects can be accounted for by constructing a so-called extended Stokes-Dirac structure, see [21, 53] for representative examples of such extended structures in the scope of *state-dependent* and *state-independent* operators. To this end, it will be required to approximate extended Stokes-Dirac structures. Lossy port-Hamiltonian systems have been discretized in the past [162]. However, the extended Stokes-Dirac structures were *state-independent*. Moreover, only the corresponding structure-preserving *spatial* discretization scheme was outlined. We are interested in the complete structure-preserving discretization of an extended Stokes-Dirac structure while accounting for dissipative effects.

In the past, spatial discretization methods preserving the port-Hamiltonian structure have been proposed in the scope of one- and multi-dimensional problems [35, 75, 97, 126, 161–163, 179]. Several well-known spatial discretization methods include pseudo-spectral methods [126], mixed finite element methods [75, 97], partitioned finite-element methods [35, 162], staggered finite-volume and finite-difference methods [179], and explicit simplicial discretization methods built on the framework of exterior calculus [163].

However, most of these existing works on structure-preserving spatial discretization of infinite-dimensional port-Hamiltonian representations have focused on the approximation of a constant Stokes-Dirac structure arising from a *state-independent* (and extended) skew-adjoint Hamiltonian operator. For instance, in [160], non-canonical but physically relevant Hamiltonian functionals lead to non-linear port-Hamiltonian systems with an underlying constant and an extended *state-independent* Stokes-Dirac structure. Such a geometric structure has subsequently been discretized in a structure-preserving manner using the partitioned finite-element method in [161]. Structure-preserving spatial discretization has also been performed for first- and higher-order state-independent Hamiltonian operators with distributed inputs, see [33]. However, for those methods, the Hamiltonian operators were again *state-independent*. As for applications, the (lossless) transmission line, wave equations, Maxwell's equations, rendering *state-independent* Hamiltonian operators, have generally been used as test-beds for the numerical validation of the structure-preserving discretization framework [75]. Additionally, most of the existing spatial discretization methods have been developed and tested for models with quadratic Hamiltonian functionals.

Non-quadratic Hamiltonian functionals and *state-dependent* skew-adjoint operators (under a certain choice of the state-variables) are specific features of the

TFM and the DFM. One of the existing works on structure-preserving spatial discretization of mathematical models with non-quadratic Hamiltonian functionals includes [95]. This work deals with the structure-preserving discretization of a port-Hamiltonian representation of a compressible single-phase flow model. It focuses on discretizing *state-independent* Hamiltonian operators or Stokes-Dirac structures (under a certain choice of the state variables) for the model of interest. In contrast, the first focus of this chapter is to spatially discretize *state-dependent* Stokes-Dirac structures (or port-Hamiltonian representations) in a structure-preserving manner. Moreover, it may be the case that any choice of state-variables will yield *state-dependent* skew-adjoint Hamiltonian operators; see [19] for port-Hamiltonian representation(s) of the Drift Flux Model.

Relevant works in the direction of structure-preserving discretization of a *state-dependent* Stokes-Dirac structure include [23,140] and [148]. These works are built upon the mixed finite element method (mFEM), which was initially proposed for canonical port-Hamiltonian systems in [75]. The methodology is based on the concept of using mixed or different finite elements for the approximation of flow and effort variables. This methodology has been used to deal with the spatial discretization of linear shallow water equations (accounting for Coriolis forces and topography) with a non-constant Stokes-Dirac structure in [148]. Spatial discretization essentially translates a Stokes-Dirac structure into a finite-dimensional Dirac structure [162]. In view of our interest to develop structure-preserving discretization techniques for a port-Hamiltonian representation of the TFM and the DFM, we recall that the corresponding Stokes-Dirac structure, recently introduced in [19], is novel and has not been spatially-discretized or both spatially- and temporally-discretized in a structure-preserving manner.

We employ the mFEM for the structure-preserving spatial discretization of the TFM and the DFM. We demonstrate that a finite-dimensional representation preserves the finite-dimensional Dirac structure, i.e., the geometric interconnection structure and the power balance equation. We also discuss the role of the boundary port-variables in guaranteeing a finite-dimensional Dirac structure. It is of interest to investigate whether the mFEM, which has so far been tested on linear(ized) problems or *state-independent* Stokes-Dirac structures, requires fundamental modifications to deal with the non-linearity of the mathematical model or the *state-dependent* nature of the Stokes-Dirac structure. Hence, we provide insights on the influence of the model non-linearity or *state-dependent* Stokes-Dirac structure in obtaining a structure-preserving spatial discretization method. Moreover, the mFEM is known to yield feedthrough terms in the finite-dimensional state-space model representation. We subsequently reason

that it is not necessary to eliminate the appearance of feedthrough terms from the finite-dimensional model approximation and that a special structure in the feedthrough matrix makes it amenable to use the mFEM for spatially discretizing hyperbolic systems, such as, the TFM and the DFM.

The structure-preserving spatial discretization yields continuous-time, finite-dimensional port-Hamiltonian representations of the models of interest. A next natural step is to develop a discrete-time, finite-dimensional port-Hamiltonian representation by employing suitable structure-preserving temporal discretization methods on the continuous-time port-Hamiltonian representation. Discrete-time port-Hamiltonian systems have been recently defined in [96] based on the concept of symplectic Gauss-Legendre collocation. Following the principles laid down in [96], we employ Gauss-Legendre collocation methods and seek to develop discrete-time port-Hamiltonian representations of the TFM and the DFM. However, the authors in [96] do not consider the presence of feedthrough terms in the state-space port-Hamiltonian model, which is the starting point for the application of the concept of the Gauss-Legendre collocation methods. It is worth to emphasize that the mFEM, employed for the structure-preserving spatial discretization, yields feedthrough terms in the continuous-time, finite-dimensional port-Hamiltonian representations of the TFM and the DFM. To this end, we extend the existing concepts in [96]. We account for the state-dependent system matrices and the feedthrough terms in the continuous-time port-Hamiltonian representations of the two-phase flow models of interest and, subsequently, develop corresponding discrete-time port-Hamiltonian representations.

The main contributions of the chapter are to: *(i)* perform a spatial discretization both for the conservative and the resistive setting, and construct a point-wise in time, finite-dimensional Dirac structure on a minimal discrete bond space, *(ii)* interconnect finite-dimensional models (obtained for each discretized lump) in a power-preserving manner, *(iii)* perform a temporal discretization on the state-space port-Hamiltonian model with feedthrough term, and obtain a discrete-time Dirac structure, and *(iv)* numerically assess the proposed structure-preserving discretization framework.

The chapter is organized as follows. In Section 5.2, the two-phase flow models of interest, i.e., the TFM and the DFM, are introduced. Mixed finite element based structure-preserving spatial discretization is then discussed in Section 5.3. The spatial discretization methodology, along with the framework of power-preserving interconnection of discretized lumps, yielding the state-space port-Hamiltonian model is extensively discussed for the TFM and the DFM in Section 5.4. Section 5.5 is dedicated to structure-preserving temporal discretization that is built upon the concept of Gauss-Legendre collocation

methods. Then, numerical experiments are discussed in Section 5.6 to validate the theoretical foundation. Finally, Section 5.7 ends with conclusions and potential future works.

5.2 Revisiting Two-Phase Flow Models

In this section, we state port-Hamiltonian formulations of the two-phase flow models and their equivalent extended Stokes-Dirac structure representations to set the foundation for developing a corresponding structure-preserving finite-dimensional realization.

5.2.1 Two-Fluid Model

The non-quadratic Hamiltonian functional $\mathcal{H}_T$ is given by

$$\mathcal{H}_T(q_1, q_2, q_3, q_4) := \int_\Omega \frac{q_3^2}{2q_1} + \frac{q_4^2}{2q_2} + q_1 c_g^2 \ln\left(\frac{p}{c_g^2}\right)$$

$$+ q_2 c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right) + \left(1 - \alpha_g\right)\beta \, \mathrm{d}z, \quad (5.1)$$

where p and α_g can be replaced by the relations (3.5) and (3.6), respectively, and where $q = [q_1, q_2, q_3, q_4]^T := [m_g, m_\ell, m_g v_g, m_\ell v_\ell]^T$ denotes the state of the system.

Theorem 5.2.1 *The Two-Fluid Model, governed by* (3.1) *together with the closure equations* (3.2)*, can be written in a port-Hamiltonian representation as follows:*

$$\partial_t q = \left(\mathcal{J}_T(q) - \mathcal{R}_T\right)\delta_q \mathcal{H}_T(q) \text{ or } f = \left(\mathcal{J}_T(q) - \mathcal{R}_T\right)e, \quad (5.2)$$

with state-variable q, the Hamiltonian functional $\mathcal{H}_T$ as in (5.1)*,*

$$\mathcal{J}_T(q) = -\begin{bmatrix} 0 & 0 & \partial_z(q_1 \cdot) & 0 \\ 0 & 0 & 0 & \partial_z(q_2 \cdot) \\ q_1\partial_z(\cdot) & 0 & \partial_z(q_3 \cdot) + q_3\partial_z(\cdot) & 0 \\ 0 & q_2\partial_z(\cdot) & 0 & \partial_z(q_4 \cdot) + q_4\partial_z(\cdot) \end{bmatrix},$$

$$\mathcal{R}_T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & b_g^M & -b_g^M \\ 0 & 0 & -b_g^M & b_g^M \end{bmatrix}, \quad (5.3)$$

and the boundary port flow and effort variables

$$
\begin{cases}
e_{b1}^B = e_{q_3}|_{z=L}, & f_{b1}^B = \left(q_1 e_{q_1} + q_3 e_{q_3}\right)|_{z=L}, \\
e_{b2}^B = e_{q_4}|_{z=L}, & f_{b2}^B = \left(q_2 e_{q_2} + q_4 e_{q_4}\right)|_{z=L}, \\
f_{a1}^B = -e_{q_3}|_{z=0}, & e_{a1}^B = \left(q_1 e_{q_1} + q_3 e_{q_3}\right)|_{z=0}, \\
f_{a2}^B = -e_{q_4}|_{z=0}, & e_{a2}^B = \left(q_2 e_{q_2} + q_4 e_{q_4}\right)|_{z=0},
\end{cases}
\tag{5.4}
$$

where the flow and effort variables have been introduced in (5.2). The flow variables f are defined as the derivative of the state-variables with respect to time i.e., $f = [f_{q_1}, f_{q_2}, f_{q_3}, f_{q_4}]^T := \partial_t q$. Moreover, the effort variables e are defined as the variational derivative of the Hamiltonian functional $\mathcal{H}_T(q)$ with respect to the state-variable, i.e., $e = [e_{q_1}, e_{q_2}, e_{q_3}, e_{q_4}]^T := \delta_q \mathcal{H}_T(q)$.

Remark 5.2.2 *The proof to Theorem 5.2.1 follows directly from Theorem 3.3.6, which laid the foundation of dissipative Hamiltonian representation of the TFM, and Theorem 3.4.9, which characterized the corresponding boundary port variables such that the underlying geometric structure exists.*

Using the formally skew-adjointness of $\mathcal{J}_T(q)$, the behaviour of the Hamiltonian along the solutions of the mathematical model is governed by the following power balance equation:

$$
\frac{d\mathcal{H}_T}{dt} = \int_\Omega (\delta_q \mathcal{H}_T(q))^T (-\mathcal{R}_T(q)) \delta_q \mathcal{H}_T(q) \, dz + \left(e_{q_3}\left(q_1 e_{q_1} + q_3 e_{q_3}\right)\right)|_{z=0} +
$$
$$
\left(e_{q_4}\left(q_2 e_{q_2} + q_4 e_{q_4}\right)\right)|_{z=0} - \left(e_{q_3}\left(q_1 e_{q_1} + q_3 e_{q_3}\right)\right)|_{z=L} - \left(e_{q_4}\left(q_2 e_{q_2} + q_4 e_{q_4}\right)\right)|_{z=L}.
$$

We can clearly observe that $\frac{d\mathcal{H}_T}{dt}$, in the absence of dissipation, is governed by the product of boundary port variables.

Remark 5.2.3 *We have considered a non-unique way to define the boundary port flow and effort variables. Such boundary port variables can be parametrized following the principles laid down in [101].*

The boundary flow and effort variables can be interpreted physically. Ignoring the sign associated to the boundary flow (and effort) variables; f_{a1}^B and f_{a2}^B (e_{b1}^B and e_{b2}^B) can be interpreted as gas and liquid volumetric flow rate, respectively, at the left end (and at the right end) of the spatial domain. The flow (and effort) variables f_{b1}^B and f_{b2}^B (e_{a1}^B and e_{a2}^B) have physical dimensions of energy per unit mass.

Equivalent to the representation in Theorem 5.2.1, the Two-Fluid Model can be represented in terms of an extended Stokes-Dirac structure as shown next.

Theorem 5.2.4 *The Two-Fluid Model, governed by (3.1) together with the closure equations (3.2), can be represented in an extended Stokes-Dirac structure representation as follows:*

$$
\begin{bmatrix} f_{q_1} \\ f_{q_2} \\ f_{q_3} \\ f_{q_4} \\ f_{R_1} \\ f_{R_2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \partial_z(q_1\cdot) & 0 & 0 & 0 \\ 0 & 0 & 0 & \partial_z(q_2\cdot) & 0 & 0 \\ q_1\partial_z(\cdot) & 0 & \partial_z(q_3\cdot)+q_3\partial_z(\cdot) & 0 & -1 & 0 \\ 0 & q_2\partial_z(\cdot) & 0 & \partial_z(q_4\cdot)+q_4\partial_z(\cdot) & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} e_{q_1} \\ e_{q_2} \\ e_{q_3} \\ e_{q_4} \\ e_{R_1} \\ e_{R_2} \end{bmatrix}, \quad (5.5)
$$

with closure relations

$$
e_{R_1} = b_g^M f_{R_1} - b_g^M f_{R_2}, \quad e_{R_2} = -b_g^M f_{R_1} + b_g^M f_{R_2}, \quad (5.6)
$$

and the boundary port flow and effort variables

$$
\begin{cases}
e_{b1}^B = e_{q_3}|_{z=L}, & f_{b1}^B = \left(q_1 e_{q_1} + q_3 e_{q_3}\right)|_{z=L}, \\
e_{b2}^B = e_{q_4}|_{z=L}, & f_{b2}^B = \left(q_2 e_{q_2} + q_4 e_{q_4}\right)|_{z=L}, \\
f_{a1}^B = -e_{q_3}|_{z=0}, & e_{a1}^B = \left(q_1 e_{q_1} + q_3 e_{q_3}\right)|_{z=0}, \\
f_{a2}^B = -e_{q_4}|_{z=0}, & e_{a2}^B = \left(q_2 e_{q_2} + q_4 e_{q_4}\right)|_{z=0}.
\end{cases} \quad (5.7)
$$

Remark 5.2.5 *Similar to the proof of Theorem 5.2.1, the proof of Theorem 5.2.4 also uses Theorems 3.3.6 and 3.4.9. In addition, it relies on the concept of extended Stokes-Dirac structure. We omit the detailed proof, and instead refer the reader to follow the lines of reasoning in the proof of Proposition 4.2.9 for obtaining extended Stokes-Dirac structure from the knowledge of dissipative Hamiltonian representation of the TFM.*

This extended geometric structure will be utilized later in this chapter in order to develop a structure-preserving discretization of the TFM in the presence of resistive or dissipative effects. In particular, the extended Stokes-Dirac structure will play a crucial role in quantification of the energy attributed to the resistive ports. Consequently, it will aid in assessing whether the finite-dimensional extended Stokes-Dirac structure preserves the energy due to the resistive effects.

5.2.2 Drift Flux Model

The non-linear, non-quadratic and non-separable Hamiltonian functional for the Drift Flux Model is given by:

$$\mathcal{H}_D(m_g, m_\ell, v) = \int_\Omega m_g \frac{v^2}{2} + m_\ell \frac{v^2}{2} + m_\ell c_\ell^2 \ln\left(\frac{p+\beta}{c_\ell^2}\right) + m_g c_g^2 \ln\left(\frac{p}{c_g^2}\right) +$$

$$\alpha_\ell \beta + (m_g + m_\ell)\left(\int_a^x g \sin(\theta(\xi)) d\xi\right) dx. \quad (5.8)$$

Theorem 5.2.6 *The Drift Flux Model (without slip), governed by (3.15) together with $v := v_g = v_\ell$, the closure equations (3.2a), (3.2e), (3.2f) and (3.16), can be written in a port-Hamiltonian representation as follows:*

$$\partial_t x = \left(\mathcal{J}_D(x) - \mathcal{R}_D(x)\right)\delta_x \mathcal{H}_D(x), \text{ or } f = \left(\mathcal{J}_D(x) - \mathcal{R}_D(x)\right)e \quad (5.9)$$

with $x := [x_1, x_2, x_3]^T = [m_g, m_\ell, v]^T$, the Hamiltonian functional $\mathcal{H}_D$ as in (5.8),

$$\mathcal{J}_D(x) = \begin{bmatrix} 0 & 0 & -\partial_z\left(\frac{m_g}{m_g+m_\ell}\cdot\right) \\ 0 & 0 & -\partial_z\left(\frac{m_\ell}{m_g+m_\ell}\cdot\right) \\ -\frac{m_g}{m_g+m_\ell}\partial_z(\cdot) & -\frac{m_\ell}{m_g+m_\ell}\partial_z(\cdot) & 0 \end{bmatrix}, \quad (5.10)$$

$$\mathcal{R}_D(x) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{32\mu_m}{d^2(m_g+m_\ell)^2} \end{bmatrix}, \quad (5.11)$$

and the boundary port flow and effort variables

$$\begin{bmatrix} f_a^B \\ f_b^B \end{bmatrix} = \begin{bmatrix} -(e_{m_g} A_g + e_{m_\ell} A_\ell)|_{z=0} \\ (e_{m_g} A_g + e_{m_\ell} A_\ell)|_{z=L} \end{bmatrix}, \begin{bmatrix} e_a^B \\ e_b^B \end{bmatrix} = \begin{bmatrix} e_v|_{z=0} \\ e_v|_{z=L} \end{bmatrix}, \quad (5.12)$$

where the flow and effort variables have been introduced in (5.9), and the short-hand notations A_g and A_ℓ are given by $A_g = \frac{m_g}{m_g+m_\ell}$ and $A_\ell = \frac{m_\ell}{m_g+m_\ell}$, respectively. The flow variables f are defined as the derivative of the state-variables with respect to time i.e., $f = [f_{x_1}, f_{x_2}, f_{x_3}]^T := \partial_t q$. Moreover, the effort variables e are defined as the variational derivative of the Hamiltonian functional $\mathcal{H}_D(x)$ with respect to the state-variable, i.e., $e = [e_{x_1}, e_{x_2}, e_{x_3}]^T := \delta_q \mathcal{H}_D(x)$.

Remark 5.2.7 *The proof to Theorem 5.2.6 follows directly from Theorem 3.3.10, which laid the foundation of dissipative Hamiltonian representation of the DFM, and Theorem 3.4.11, which characterized the corresponding boundary port variables such that the underlying geometric structure exists.*

The boundary flow and effort variables can be interpreted physically. Ignoring the sign associated to the boundary flow and effort variables; e_a^B and e_b^B represent total volumetric flow rate at the left end and the right end of the spatial domain, respectively. The flow variables f_a^B and f_b^B have physical dimensions of energy per unit mass.

Using the formally skew-adjointness of $\mathscr{J}_D(x)$, the behaviour of the Hamiltonian along the solutions of the mathematical model is governed by the following power balance equation:

$$\frac{\mathrm{d}\mathscr{H}_D}{\mathrm{d}t} = \int_\Omega (\delta_x \mathscr{H}_D(x))^T (-\mathscr{R}_D(x)) \delta_x \mathscr{H}_D(x) \; \mathrm{d}z + \left(e_v \Big(e_{m_g} A_g + e_{m_\ell} A_\ell \Big) \right)|_{z=0} -$$
$$\left(e_v \Big(e_{m_g} A_g + e_{m_\ell} A_\ell \Big) \right)|_{z=L}.$$

We observe that $\frac{\mathrm{d}\mathscr{H}_D}{\mathrm{d}t}$, in the absence of dissipation, is governed by the product of boundary port variables.

Similar to the structure-preserving discretization of the TFM, we will exploit an extended Stokes-Dirac structure representation of the DFM to develop a spatially discretized model for the DFM in the presence of resistive effects. We do not state the extended geometric structure here and instead refer to Proposition 4.2.9 for its corresponding description.

5.3   Structure-preserving Spatial Discretization

A closed form analytical solution does not exist for complicated models such as the two-phase flow models presented in Section 5.2. Numerical solvers play a key role to provide insight in the behaviour of the physical model under consideration. In the past, several numerical methods such as Finite Difference Methods [70], Finite Volume Methods [14, 57–61, 67, 128, 167] and Finite Element Methods [192] have been employed to numerically approximate two-phase flow models. However, these numerical methods are not necessarily structure-preserving by construction. Structure-preserving numerical discretization methods can be formulated either in the strong form [75] or in the weak/variational form [97].

We employ the mFEM in the scope of structure-preserving discretization of two-phase flow models. This methodology is also valid for systems on multi-dimensional spatial domains. In this chapter, we will focus on one-dimensional problems only. For an application of such an approach to multi-dimensional problems, we refer the reader to [75, 97].

The major challenge lies in identifying a finite-dimensional (continuous-time) Dirac structure with non-degenerate power pairings. Such a structure is crucial as it serves as a basis for the state-space port-Hamiltonian model.

We first set the background by recalling key mathematical preliminaries. We begin with defining Dirac structures.

A (finite-dimensional) Dirac structure $\mathscr{D}$ is defined as follows.

Definition 5.3.1 *[20] Consider $e \in \mathscr{E}$ and $f \in \mathscr{F}$, where $\mathscr{E}$ and $\mathscr{F}$ are real finite-dimensional normed vector spaces and where $\mathscr{E} = \mathscr{F}^*$, the dual space of $\mathscr{F}$. We define*

$$e(f) := \langle e, f \rangle := f^T e = \frac{1}{2} f^T e + \frac{1}{2} e^T f. \tag{5.13}$$

Moreover, we define the non-degenerate bilinear pairing $d : (\mathscr{E} \times \mathscr{F}) \times (\mathscr{E} \times \mathscr{F}) \to \mathbb{R}$ in the following way:

$$d\big((e_1, f_1), (e_2, f_2)\big) := \langle e_2, f_1 \rangle + \langle e_1, f_2 \rangle. \tag{5.14}$$

The subspace $\mathscr{D} \subset \mathscr{E} \times \mathscr{F}$ is a Dirac structure if $\mathscr{D} = \mathscr{D}^\perp$, where $\mathscr{D}^\perp$ is defined as follows:

$$\mathscr{D}^\perp := \{(e, f) \in \mathscr{E} \times \mathscr{F} \mid d\big((e_0, f_0), (e, f)\big) = 0 \quad \forall (e_0, f_0) \in \mathscr{D}\}. \tag{5.15}$$

Several matrix representations are possible for a finite-dimensional Dirac structure. Here, we only recall the definition of the image representation of a Dirac structure.

Definition 5.3.2 *Let $\mathscr{D} \subset \mathscr{E} \times \mathscr{F}$ with $dim(\mathscr{F}) = n$. A subspace*

$$\mathscr{D} = \{(e, f) \in \mathscr{E} \times \mathscr{F} \mid \exists \lambda \in \mathbb{R}^m \ \ s.t. \ \ e = E^T \lambda, f = F^T \lambda\}, \tag{5.16}$$

endowed with a non-degenerate bilinear pairing (5.14), is an image representation of a Dirac structure if and only if there exist two $n \times n$ real matrices E and F that satisfy

(i) $EF^T + FE^T = 0.$
(ii) $rank([E \quad F]) = n.$

Condition *(i)* above is a discrete equivalent of a power balance equation (or power-preservation), while, condition *(ii)* is representative of richness in the interconnection structure in order to interconnect discretized lumps in a power-preserving manner.

Definition 5.3.3 *[187] Let $\mathscr{D} \subset \mathscr{E} \times \mathscr{F}$, with dim($\mathscr{F}$) = n, be a Dirac structure. Then $\mathscr{D}$ can be represented in the kernel representation in the following manner:*

$$\mathscr{D} = \{(e, f) \in \mathscr{E} \times \mathscr{F} \mid Ee + Ff = 0\} \tag{5.17}$$

for any $n \times n$ real matrices E and F satisfying

(i) $FE^T + EF^T = 0$, *and* (ii) $rank(E \mid F) = n$.

Next, we first discuss about the structure-preserving spatial discretization of the TFM and then about the structure-preserving spatial discretization of the DFM (without slip).

5.3.1   Spatial Discretization of Two-Fluid Model

The structure-preserving spatial discretization procedure essentially consists of the following steps. We approximate the states, and the flow and effort variables by suitable discrete functions. We then use such an approximation to obtain finite-dimensional equations and ensure power balance. In addition to preserving the power balance equation, we assess the existence of the finite-dimensional Dirac structure.

It is known that it is easier to simulate the resulting finite-dimensional realization in the presence of resistive effects than in the absence of such resistive effects. However, even in the presence of resistive effects, the numerical solvers may produce solutions in which the numerical dissipated power is not (exactly) matching with the dissipated power of the infinite-dimensional system. The issue is that the numerical solvers distort or mismatch the resistive nature of a port-Hamiltonian system and this is, of course, most visible in the lossless case. Hence, we develop the structure-preserving discretization methodology for both the lossless and the lossy setting.

Conservative Setting (Lossless case)
Here, we consider the setting which abides by the strict conservation of energy

in the absence of dissipative terms. So, $b_g^M = 0$ in (5.3) and, therefore, $\mathcal{R}_T = 0$ in the following discussion.

Numerical discretization entails discretizing the spatial domain $\Omega = [0, L]$ into several finite elements. We apply the procedure to a single discretized lump $Z_{ab} = [a, b]$, with $0 \le a < b \le L$, in the scope of the TFM.

Each state variable is discretized using one spatial basis function as follows:

$$q_i(t,z) = q_i^{ab}(t)\omega_{q_i}^{ab}(z), \text{ for all } i \in \{1,2,3,4\}, \tag{5.18}$$

with $z \in Z_{ab}$. The normalization assumption on $\omega_{q_i}^{ab}(z)$ yields

$$\int_a^b \omega_{q_i}^{ab}(z)dz = 1, \text{ for all } i \in \{1,2,3,4\}. \tag{5.19}$$

Recall from Theorem 5.2.1 that we have:

$$f_{q_i} = \partial_t q_i, \quad e_{q_i} = \delta_{q_i}\mathcal{H}(q), \text{ for all } i \in \{1,2,3,4\}. \tag{5.20}$$

Each flow variable f_{q_i} is discretized in a manner similar to the state approximation, i.e.,

$$f_{q_i}(t,z) = f_{q_i}^{ab}(t)\omega_{q_i}^{ab}(z), \quad \text{for all } i \in \{1,2,3,4\}. \tag{5.21}$$

Each effort variable e_{q_i} is spatially discretized using two different basis functions $\omega_{q_i}^a(z)$ and $\omega_{q_i}^b(z)$ in the following manner:

$$e_{q_i}(t,z) = e_{q_i}^a(t)\omega_{q_i}^a(z) + e_{q_i}^b(t)\omega_{q_i}^b(z), \text{ for all } i \in \{1,2,3,4\}, \tag{5.22}$$

where the following boundary conditions hold:

$$\omega_{q_i}^a(a) = 1, \omega_{q_i}^a(b) = 0, \omega_{q_i}^b(a) = 0, \omega_{q_i}^b(b) = 1, \text{ for all } i \in \{1,2,3,4\}, \tag{5.23}$$

and

$$e_{q_i}^a(t) = e_{q_i}(t,a), \quad e_{q_i}^b(t) = e_{q_i}(t,b), \text{ for all } i \in \{1,2,3,4\}. \tag{5.24}$$

Here, $e_{q_i}(t,a)$ (resp. $e_{q_i}(t,b)$) is the value of the effort variable at the left (resp. the right) boundary of Z_{ab}.

Inserting the discrete approximation of the flow variables (5.21), state variables (5.18) and effort variables (5.22) into (5.2), we obtain

$$f_{q_1}^{ab}(t)\omega_{q_1}^{ab}(z) = -q_1^{ab}(t)e_{q_3}^a(t)\partial_z\left(\omega_{q_1}^{ab}(z)\omega_{q_3}^a(z)\right) - q_1^{ab}(t)e_{q_3}^b(t)\partial_z\left(\omega_{q_1}^{ab}(z)\omega_{q_3}^b(z)\right), \tag{5.25}$$

$$f_{q_2}^{ab}(t)\omega_{q_2}^{ab}(z) = -q_2^{ab}(t)e_{q_4}^{a}(t)\partial_z\left(\omega_{q_2}^{ab}(z)\omega_{q_4}^{a}(z)\right) - q_2^{ab}(t)e_{q_4}^{b}(t)\partial_z\left(\omega_{q_2}^{ab}(z)\omega_{q_4}^{b}(z)\right),$$
(5.26)

$$f_{q_3}^{ab}(t)\omega_{q_3}^{ab}(z) = -q_1^{ab}(t)\omega_{q_1}^{ab}(z)e_{q_1}^{a}(t)\partial_z(\omega_{q_1}^{a}(z)) - q_1^{ab}(t)\omega_{q_1}^{ab}(z)e_{q_1}^{b}(t)\partial_z(\omega_{q_1}^{b}(z))$$
$$- q_3^{ab}(t)e_{q_3}^{a}(t)\partial_z\left(\omega_{q_3}^{ab}(z)\omega_{q_3}^{a}(z)\right) - q_3^{ab}(t)e_{q_3}^{b}(t)\partial_z\left(\omega_{q_3}^{ab}(z)\omega_{q_3}^{b}(z)\right)$$
$$- q_3^{ab}(t)e_{q_3}^{a}(t)\omega_{q_3}^{ab}(z)\partial_z(\omega_{q_3}^{a}(z)) - q_3^{ab}(t)e_{q_3}^{b}(t)\omega_{q_3}^{ab}(z)\partial_z(\omega_{q_3}^{b}(z)),$$
(5.27)

$$f_{q_4}^{ab}(t)\omega_{q_4}^{ab}(z) = -q_2^{ab}(t)\omega_{q_2}^{ab}(z)e_{q_2}^{a}(t)\partial_z(\omega_{q_2}^{a}(z)) - q_2^{ab}(t)\omega_{q_2}^{ab}(z)e_{q_2}^{b}(t)\partial_z(\omega_{q_2}^{b}(z))$$
$$- q_4^{ab}(t)e_{q_4}^{a}(t)\partial_z\left(\omega_{q_4}^{ab}(z)\omega_{q_4}^{a}(z)\right) - q_4^{ab}(t)e_{q_4}^{b}(t)\partial_z\left(\omega_{q_4}^{ab}(z)\omega_{q_4}^{b}(z)\right)$$
$$- q_4^{ab}(t)e_{q_4}^{a}(t)\omega_{q_4}^{ab}(z)\partial_z(\omega_{q_4}^{a}(z)) - q_4^{ab}(t)e_{q_4}^{b}(t)\omega_{q_4}^{ab}(z)\partial_z(\omega_{q_4}^{b}(z)).$$
(5.28)

The relation between different basis functions, in the sense of compatibility of forms as in [75], needs to concur with the following points:

1. The basis function $\omega_{q_1}^{ab}(z)$ and functions $\omega_{q_3}^{a}(z)$ and $\omega_{q_3}^{b}(z)$ should be chosen in such a way that for every $e_{q_3}^{a}(t)$ and $e_{q_3}^{b}(t)$ we can find $f_{q_1}^{ab}(t)$ such that (5.25) is satisfied.

2. The basis function $\omega_{q_2}^{ab}(z)$ and functions $\omega_{q_4}^{a}(z)$ and $\omega_{q_4}^{b}(z)$ should be chosen in such a way that for every $e_{q_4}^{a}(t)$ and $e_{q_4}^{b}(t)$ we can find $f_{q_2}^{ab}(t)$ such that (5.26) is satisfied.

3. The basis function $\omega_{q_3}^{ab}(z)$ and functions $\omega_{q_1}^{a}(z)$, $\omega_{q_1}^{b}(z)$, $\omega_{q_3}^{a}(z)$ and $\omega_{q_3}^{b}(z)$ should be chosen in such a way that for every $e_{q_1}^{a}(t)$, $e_{q_1}^{b}(t)$, $e_{q_3}^{a}(t)$ and $e_{q_3}^{b}(t)$ we can find $f_{q_3}^{ab}(t)$ such that (5.27) is satisfied.

4. The basis function $\omega_{q_4}^{ab}(z)$ and functions $\omega_{q_2}^{a}(z)$, $\omega_{q_2}^{b}(z)$, $\omega_{q_4}^{a}(z)$ and $\omega_{q_4}^{b}(z)$ should be chosen in such a way that for every $e_{q_2}^{a}(t)$, $e_{q_2}^{b}(t)$, $e_{q_4}^{a}(t)$ and $e_{q_4}^{b}(t)$ we can find $f_{q_4}^{ab}(t)$ such that (5.28) is satisfied.

For instance, take $e_{q_3}^{b}(t) = 0$. Under this assumption, (5.25) is true if and only if $\partial_z\left(\omega_{q_1}^{ab}(z)\omega_{q_3}^{a}(z)\right) = c\omega_{q_1}^{ab}(z)$ for a constant c. Integrating this equation over Z_{ab} yields:

$$\omega_{q_1}^{ab}(b)\omega_{q_3}^{a}(b) - \omega_{q_1}^{ab}(a)\omega_{q_3}^{a}(a) = c\int_{Z_{ab}}\omega_{q_1}^{ab}(z)\,\mathrm{d}z.$$
(5.29)

Exploiting the normalization assumption (5.19) and the boundary values of the basis functions ($\omega_{q_3}^a$ and $\omega_{q_3}^b$) (5.23), we can find that $c = -\omega_{q_1}^{ab}(a)$ and, hence,

$$\partial_z\left(\omega_{q_1}^{ab}(z)\omega_{q_3}^a(z)\right) = -\omega_{q_1}^{ab}(a)\omega_{q_1}^{ab}(z). \tag{5.30}$$

Analogously, we can obtain the following relations:

$$\partial_z\left(\omega_{q_1}^{ab}(z)\omega_{q_3}^b(z)\right) = \omega_{q_1}^{ab}(b)\omega_{q_1}^{ab}(z),$$
$$\partial_z\left(\omega_{q_2}^{ab}(z)\omega_{q_4}^a(z)\right) = -\omega_{q_2}^{ab}(a)\omega_{q_2}^{ab}(z), \tag{5.31}$$
$$\partial_z\left(\omega_{q_2}^{ab}(z)\omega_{q_4}^b(z)\right) = \omega_{q_2}^{ab}(b)\omega_{q_2}^{ab}(z).$$

Now consider $e_{q_1}^b(t) = 0$, $e_{q_3}^a(t) = 0$ and $e_{q_3}^b(t) = 0$. Under this assumption, (5.27) is true if and only if $\omega_{q_1}^{ab}(z)\partial_z\omega_{q_1}^a(z) = c\omega_{q_3}^{ab}(z)$ for a different constant c. Integrating this equation over Z_{ab} gives the following result:

$$c\int_{Z_{ab}} \frac{\omega_{q_3}^{ab}(z)}{\omega_{q_1}^{ab}(z)}\mathrm{d}z = \omega_{q_1}^a(b) - \omega_{q_1}^a(a) = -1, \tag{5.32}$$

where we have exploited (5.19) and (5.23). Using (5.32) we have that:

$$\omega_{q_1}^{ab}(z)\partial_z\omega_{q_1}^a(z) = -\frac{1}{\int_{Z_{ab}} \frac{\omega_{q_3}^{ab}(z)}{\omega_{q_1}^{ab}(z)}\mathrm{d}z}\omega_{q_3}^{ab}(z). \tag{5.33}$$

In a similar manner, we can arrive at the following identities:

$$\omega_{q_1}^{ab}(z)\partial_z\omega_{q_1}^b(z) = \frac{1}{\int_{Z_{ab}} \frac{\omega_{q_3}^{ab}(z)}{\omega_{q_1}^{ab}(z)}\mathrm{d}z}\omega_{q_3}^{ab}(z),$$
$$\omega_{q_2}^{ab}(z)\partial_z\omega_{q_2}^a(z) = -\frac{1}{\int_{Z_{ab}} \frac{\omega_{q_4}^{ab}(z)}{\omega_{q_2}^{ab}(z)}\mathrm{d}z}\omega_{q_4}^{ab}(z), \tag{5.34}$$
$$\omega_{q_2}^{ab}(z)\partial_z\omega_{q_2}^b(z) = \frac{1}{\int_{Z_{ab}} \frac{\omega_{q_4}^{ab}(z)}{\omega_{q_2}^{ab}(z)}\mathrm{d}z}\omega_{q_4}^{ab}(z).$$

We now consider $e_{q_1}^a(t) = 0$, $e_{q_1}^b(t) = 0$ and $e_{q_3}^b(t) = 0$. Under this assumption, (5.27) is true if and only if

$$\partial_z\left(\omega_{q_3}^{ab}(z)\omega_{q_3}^a(z)\right) + \omega_{q_3}^{ab}(z)\partial_z\omega_{q_3}^a(z) = c\omega_{q_3}^{ab}(z), \tag{5.35}$$

for another constant c. Integrating the above equation over Z_{ab} yields:

$$c \int_{Z_{ab}} \omega_{q_3}^{ab}(z) = \omega_{q_3}^{ab}(b)\omega_{q_3}^{a}(b) - \omega_{q_3}^{ab}(a)\omega_{q_3}^{a}(a) + \int_{Z_{ab}} \omega_{q_3}^{ab}(z)\partial_z\omega_{q_3}^{a}(z)\ \mathrm{d}z. \qquad (5.36)$$

Exploiting the normalization assumption (5.19) and the boundary values of the basis functions (5.23), we can find that

$$c = -\omega_{q_3}^{ab}(a) + \int_{Z_{ab}} \omega_{q_3}^{ab}(z)\partial_z\omega_{q_3}^{a}(z)\ \mathrm{d}z. \qquad (5.37)$$

Using the above steps, we will analogously have the following results:

$$\begin{aligned}
\partial_z\left(\omega_{q_3}^{ab}(z)\omega_{q_3}^{a}(z)\right) + \omega_{q_3}^{ab}(z)\partial_z\omega_{q_3}^{a}(z) &= A\omega_{q_3}^{ab}(z), \\
\partial_z\left(\omega_{q_3}^{ab}(z)\omega_{q_3}^{b}(z)\right) + \omega_{q_3}^{ab}(z)\partial_z\omega_{q_3}^{b}(z) &= B\omega_{q_3}^{ab}(z), \\
\partial_z\left(\omega_{q_4}^{ab}(z)\omega_{q_4}^{a}(z)\right) + \omega_{q_4}^{ab}(z)\partial_z\omega_{q_4}^{a}(z) &= C\omega_{q_4}^{ab}(z), \\
\partial_z\left(\omega_{q_4}^{ab}(z)\omega_{q_4}^{b}(z)\right) + \omega_{q_4}^{ab}(z)\partial_z\omega_{q_4}^{b}(z) &= D\omega_{q_4}^{ab}(z),
\end{aligned} \qquad (5.38)$$

where

$$A = -\omega_{q_3}^{ab}(a) + \int_{Z_{ab}} \omega_{q_3}^{ab}(z)\partial_z\omega_{q_3}^{a}(z)\ \mathrm{d}z, \quad B = \omega_{q_3}^{ab}(b) + \int_{Z_{ab}} \omega_{q_3}^{ab}(z)\partial_z\omega_{q_3}^{b}(z)\ \mathrm{d}z, \quad (5.39)$$

$$C = -\omega_{q_4}^{ab}(a) + \int_{Z_{ab}} \omega_{q_4}^{ab}(z)\partial_z\omega_{q_4}^{a}(z)\ \mathrm{d}z, \quad D = \omega_{q_4}^{ab}(b) + \int_{Z_{ab}} \omega_{q_4}^{ab}(z)\partial_z\omega_{q_4}^{b}(z)\ \mathrm{d}z. \quad (5.40)$$

The time-derivative of the Hamiltonian in Z_{ab} is given by:

$$\frac{d}{dt}\mathcal{H}_T^{ab} = \int_{Z_{ab}} \left(f_{q_1}e_{q_1} + f_{q_2}e_{q_2} + f_{q_3}e_{q_3} + f_{q_4}e_{q_4} \right)\ \mathrm{d}z. \qquad (5.41)$$

Using the discrete approximation of the flow variables (5.21) and effort variables (5.22) in (5.41), we have:

$$\frac{d}{dt}\mathcal{H}_T^{ab} =$$

$$\int_{Z_{ab}} \Big(f_{q_1}^{ab}(t)\omega_{q_1}^{ab}(z)(e_{q_1}^{a}(t)\omega_{q_1}^{a}(z) + e_{q_1}^{b}(t)\omega_{q_1}^{b}(z)) + f_{q_2}^{ab}(t)\omega_{q_2}^{ab}(z)(e_{q_2}^{a}(t)\omega_{q_2}^{a}(z) + e_{q_2}^{b}(t)\omega_{q_2}^{b}(z))$$

$$+ f_{q_3}^{ab}(t)\omega_{q_3}^{ab}(z)(e_{q_3}^{a}(t)\omega_{q_3}^{a}(z) + e_{q_3}^{b}(t)\omega_{q_3}^{b}(z)) + f_{q_4}^{ab}(t)\omega_{q_4}^{ab}(z)(e_{q_4}^{a}(t)\omega_{q_4}^{a}(z) + e_{q_4}^{b}(t)\omega_{q_4}^{b}(z)) \Big)\ \mathrm{d}z.$$

$$(5.42)$$

Similar to [33], we define another type of effort-type variables in the following way:

$$e_{q_i}^{ab}(t) := e_{q_i}^a(t) \underbrace{\int_{Z_{ab}} \left(\omega_{q_i}^{ab}(z)\omega_{q_i}^a(z) \right) \mathrm{d}z}_{\alpha_i^a} + e_{q_i}^b(t) \underbrace{\int_{Z_{ab}} \left(\omega_{q_i}^{ab}(z)\omega_{q_i}^b(z) \right) \mathrm{d}z}_{\alpha_i^b}, \quad i = \{1,2,3,4\}.$$

$$(5.43)$$

The alternative effort-type variables $e_{q_i}^{ab}$ can be viewed as a projection of the original physically meaningful effort variables $e_{q_i}^a$ and $e_{q_i}^b$. Such projection maps are commonly required in order to preserve a finite-dimensional Dirac structure on a minimal discrete bond space [75, 97, 126]. As a result, the power balance can be expressed as a product of four flow variables and the newly defined four effort-type variables in the following manner:

$$\frac{d}{dt} \mathcal{H}_T^{ab} = \sum_{i=1}^{4} f_{q_i}^{ab} e_{q_i}^{ab}. \tag{5.44}$$

Let us define

$$\mathbf{f}_g = [f_{q_1}^{ab}, f_{q_2}^{ab}, f_{q_3}^{ab}, f_{q_4}^{ab}, f_{a1}^{B}, f_{a2}^{B}, f_{b1}^{B}, f_{b2}^{B}]^T,$$

$$\mathbf{e}_g = [e_{q_1}^{ab}, e_{q_2}^{ab}, e_{q_3}^{ab}, e_{q_4}^{ab}, e_{a1}^{B}, e_{a2}^{B}, e_{b1}^{B}, e_{b2}^{B}]^T,$$

and $\Delta z := b - a$. Using (5.25)-(5.28), (5.43) and discrete approximation of (5.7), we then have

$$\mathbf{f}_g = F_g^T \lambda_g, \quad \mathbf{e}_g = E_g^T \lambda_g, \tag{5.45}$$

where

$$F_g^T = \begin{bmatrix} 0 & 0 & 0 & 0 & q_1^{ab}\omega_{q_1}^{ab}(a) & -q_1^{ab}\omega_{q_1}^{ab}(b) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & q_2^{ab}\omega_{q_2}^{ab}(a) & -q_2^{ab}\omega_{q_2}^{ab}(b) \\ F_{31} & F_{32} & 0 & 0 & -Aq_3^{ab} & -Bq_3^{ab} & 0 & 0 \\ 0 & 0 & F_{43} & F_{44} & 0 & 0 & -Cq_4^{ab} & -Dq_4^{ab} \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & q_1^{ab}\omega_{q_1}^{ab}(b) & 0 & 0 & 0 & q_3^{ab}\omega_{q_3}^{ab}(b) & 0 & 0 \\ 0 & 0 & 0 & q_2^{ab}\omega_{q_2}^{ab}(b) & 0 & 0 & 0 & q_4^{ab}\omega_{q_4}^{ab}(b) \end{bmatrix},$$

$$(5.46)$$

$$
E_g^T = \begin{bmatrix}
\alpha_1^a & \alpha_1^b & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha_2^a & \alpha_2^b & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha_3^a & \alpha_3^b & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha_4^a & \alpha_4^b \\
q_1^{ab}\omega_{q_1}^{ab}(a) & 0 & 0 & 0 & q_3^{ab}\omega_{q_3}^{ab}(a) & 0 & 0 & 0 \\
0 & 0 & q_2^{ab}\omega_{q_2}^{ab}(a) & 0 & 0 & 0 & q_4^{ab}\omega_{q_4}^{ab}(a) & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \quad (5.47)
$$

and

$$
\lambda_g = [e_{q_1}^a, \, e_{q_1}^b, \, e_{q_2}^a, \, e_{q_2}^b, \, e_{q_3}^a, \, e_{q_3}^b, \, e_{q_4}^a, \, e_{q_4}^b]^T, \quad (5.48)
$$

with

$$
F_{31} = \frac{q_1^{ab}}{\int_{Z_{ab}} \frac{\omega_{q_3}^{ab}(z)}{\omega_{q_1}^{ab}(z)}\mathrm{d}z}, \quad F_{32} = \frac{-q_1^{ab}}{\int_{Z_{ab}} \frac{\omega_{q_3}^{ab}(z)}{\omega_{q_1}^{ab}(z)}\mathrm{d}z}, \quad F_{43} = \frac{q_2^{ab}}{\int_{Z_{ab}} \frac{\omega_{q_4}^{ab}(z)}{\omega_{q_4}^{ab}(z)}\mathrm{d}z}, \quad F_{44} = \frac{-q_2^{ab}}{\int_{Z_{ab}} \frac{\omega_{q_4}^{ab}(z)}{\omega_{q_2}^{ab}(z)}\mathrm{d}z}.
$$
$$(5.49)$$

It is not trivial to check whether the matrices F_g^T in (5.46) and E_g^T in (5.47), for any choice of basis functions consistent with the compatibility conditions, satisfy the requirements for the existence of a kernel or an image representation of a (finite-dimensional) Dirac structure; see Definition 5.3.3 and Definition 5.3.2. Also, in the scope of the system under consideration, a popular proposition as in [75] (see Proposition 1) is not straightforward to generalize. It could be said that the profile of the basis functions should be chosen such that the resulting F_g^T and E_g^T satisfy *(i)* and *(ii)* as per the Definition 5.3.2. However, still finding an explicit characterization of the basis functions remains a challenge.

As a simplification, we choose to investigate whether the geometric properties are preserved in the discrete case while associating the lowest order finite-element spaces. We consider

$$
\omega_{q_i}^{ab}(z) = \frac{1}{b-a}, \text{ for all } i = \{1,2,3,4\}, \quad (5.50)
$$

and

$$
\omega_{q_i}^a(z) = \frac{b-z}{b-a}, \quad \text{and} \quad \omega_{q_i}^b(z) = \frac{z-a}{b-a}, \text{ for all } i = \{1,2,3,4\}. \quad (5.51)
$$

Remark 5.3.4 *The time-dependence is omitted in the notation used in the sequel. We focus on obtaining a (point-wise in time) finite-dimensional Dirac structure.*

The basis functions for the flow and the effort variables should satisfy the system model in (5.2). Under the choice of the basis functions as in (5.50) and (5.51), we obtain the following finite-dimensional equations:

$$
\begin{aligned}
f_{q_1}^{ab} &= \frac{1}{b-a}\Big(q_1^{ab}e_{q_3}^a - q_1^{ab}e_{q_3}^b\Big), \\
f_{q_2}^{ab} &= \frac{1}{b-a}\Big(q_2^{ab}e_{q_4}^a - q_2^{ab}e_{q_4}^b\Big), \\
f_{q_3}^{ab} &= \frac{1}{b-a}\Big(q_1^{ab}e_{q_1}^a - q_1^{ab}e_{q_1}^b + 2q_3^{ab}e_{q_3}^a - 2q_3^{ab}e_{q_3}^b\Big), \\
f_{q_4}^{ab} &= \frac{1}{b-a}\Big(q_2^{ab}e_{q_2}^a - q_2^{ab}e_{q_2}^b + 2q_4^{ab}e_{q_4}^a - 2q_4^{ab}e_{q_4}^b\Big).
\end{aligned}
\tag{5.52}
$$

The above system (5.52) cannot be used for defining the Dirac structure as we have four flow-type variables and eight effort-type variables. Following [33], we motivate an alternative definition of (discrete) effort variables to ultimately obtain a Dirac structure at the finite-dimensional level. Alternative effort-type variables are defined as the average values of the efforts on the boundary of the discretization interval. Such a definition of effort-type variables is in the same spirit as in [33] (see Section 8.2.3), and can be viewed as a special case of the definition in (5.43) with $\alpha_i^a = \alpha_i^b = 0.5$. Such a definition essentially guarantees that the energy is preserved after (spatial) discretization. Mathematically, these alternative effort-type variables take the following form:

$$
e_{q_i}^{ab} = \frac{1}{2}(e_{q_i}^a + e_{q_i}^b), \quad \text{for all } i \in \{1,2,3,4\}.
\tag{5.53}
$$

Let us define

$$
\mathbf{f} = [f_{q_1}^{ab}, f_{q_2}^{ab}, f_{q_3}^{ab}, f_{q_4}^{ab}, f_{a1}^B, f_{a2}^B, f_{b1}^B, f_{b2}^B]^T,
$$

$$
\mathbf{e} = [e_{q_1}^{ab}, e_{q_2}^{ab}, e_{q_3}^{ab}, e_{q_4}^{ab}, e_{a1}^B, e_{a2}^B, e_{b1}^B, e_{b2}^B]^T,
$$

and $\Delta z := b - a$. Using (5.52), (5.53) and discrete approximation of (5.7), we

then have

$$
\mathbf{f}=\underbrace{\begin{bmatrix}
0 & 0 & 0 & 0 & \frac{q_1^{ab}}{\Delta z} & -\frac{q_1^{ab}}{\Delta z} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{q_2^{ab}}{\Delta z} & -\frac{q_2^{ab}}{\Delta z} \\
\frac{q_1^{ab}}{\Delta z} & -\frac{q_1^{ab}}{\Delta z} & 0 & 0 & \frac{2q_3^{ab}}{\Delta z} & -\frac{2q_3^{ab}}{\Delta z} & 0 & 0 \\
0 & 0 & \frac{q_2^{ab}}{\Delta z} & -\frac{q_2^{ab}}{\Delta z} & 0 & 0 & \frac{2q_4^{ab}}{\Delta z} & -\frac{2q_4^{ab}}{\Delta z} \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & \frac{q_1^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_3^{ab}}{\Delta z} & 0 & 0 \\
0 & 0 & 0 & \frac{q_2^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_4^{ab}}{\Delta z}
\end{bmatrix}}_{F^T}
\begin{bmatrix}
e_{q_1}^a \\ e_{q_1}^b \\ e_{q_2}^a \\ e_{q_2}^b \\ e_{q_3}^a \\ e_{q_3}^b \\ e_{q_4}^a \\ e_{q_4}^b
\end{bmatrix},
\tag{5.54}
$$

$$
\mathbf{e}=\underbrace{\begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
\frac{q_1^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_3^{ab}}{\Delta z} & 0 & 0 & 0 \\
0 & 0 & \frac{q_2^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_4^{ab}}{\Delta z} & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}}_{E^T}
\begin{bmatrix}
e_{q_1}^a \\ e_{q_1}^b \\ e_{q_2}^a \\ e_{q_2}^b \\ e_{q_3}^a \\ e_{q_3}^b \\ e_{q_4}^a \\ e_{q_4}^b
\end{bmatrix}.
\tag{5.55}
$$

Theorem 5.3.5 *Consider $\mathscr{E} \times \mathscr{F} = \mathbb{R}^8 \times \mathbb{R}^8$. The subspace*

$$
\mathscr{D}_{TFM} = \{(\mathbf{e},\mathbf{f}) \in \mathscr{E} \times \mathscr{F} \mid \exists \lambda \in \mathbb{R}^8 \ \ s.t. \ \ \mathbf{e} = E^T \lambda, \mathbf{f} = F^T \lambda\}, \tag{5.56}
$$

with E and F defined in (5.55) and (5.54), respectively, is a Dirac structure.

Proof: It is straightforward to check that F and E introduced respectively in (5.54) and (5.55) fulfill the two conditions, outlined in Definition 5.3.2, for all discrete states $q_1^{ab}, q_2^{ab}, q_3^{ab}$ and q_4^{ab}. It is also clearly observable that $\mathscr{D}_{TFM}$ is a Dirac structure irrespective of the value of the spatial step Δz. ∎

Remark 5.3.6 *Recently, the notion of power-preserving maps has been introduced in [97] to define a finite-dimensional Dirac structure abiding by the non-degenerate bilinear form.*

Remark 5.3.7 *The approximation space generated by the chosen polynomial basis may not be suitable for resolving moving discontinuities. It is well known that*

*not all choices of finite-element approximation spaces may lead to a stable mixed
Galerkin discretization or a well-behaved scheme. Hence, an alternative (discrete)
function approximation might be needed to resolve these sharp gradients.*

Using (5.54) and (5.55), the flow-effort relations are given by:

$$\mathbf{f}^{ab} = J\mathbf{e}^{ab} + B\mathbf{u}_\partial,$$
$$\mathbf{y}_\partial = C\mathbf{e}^{ab} + D\mathbf{u}_\partial,$$

(5.57)

with $\mathbf{f}^{ab} = [f_{q_1}^{ab}, f_{q_2}^{ab}, f_{q_3}^{ab}, f_{q_4}^{ab}]^T$, $\mathbf{e}^{ab} = [e_{q_1}^{ab}, e_{q_2}^{ab}, e_{q_3}^{ab}, e_{q_4}^{ab}]^T$, $\mathbf{y}_\partial = [f_{a1}^B, f_{a2}^B, f_{b1}^B, f_{b2}^B]^T$
and $\mathbf{u}_\partial = [e_{a1}^B, e_{a2}^B, e_{b1}^B, e_{b2}^B]^T$. The matrices J, B, C and D are as follows:

$$J = \begin{bmatrix} 0 & 0 & \frac{2q_1^{ab}}{\Delta z} & 0 \\ 0 & 0 & 0 & \frac{2q_2^{ab}}{\Delta z} \\ -\frac{2q_1^{ab}}{\Delta z} & 0 & 0 & 0 \\ 0 & -\frac{2q_2^{ab}}{\Delta z} & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & -\frac{2q_1^{ab}}{\Delta z} & 0 \\ 0 & 0 & 0 & -\frac{2q_2^{ab}}{\Delta z} \\ 2 & 0 & -\frac{2q_3^{ab}}{\Delta z} & 0 \\ 0 & 2 & 0 & -\frac{2q_4^{ab}}{\Delta z} \end{bmatrix},$$

$$C = -B^T, \quad D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}. \quad (5.58)$$

Remark 5.3.8 *Invoking the mFEM for the spatial discretization of the TFM is seen
to contain feedthrough terms, i.e., $D \neq 0$ in (5.58). However, the matrix D is
observed to have a special structure, i.e., it is skew-symmetric. We claim that
such a structure is not unnatural for hyperbolic systems and support it by the
following reasoning. The instantaneous information that is transferred across each
face of the finite-element is non-zero. However, the instantaneous power due to the
feedthrough matrix D is zero as shown below:*

$$\langle \mathbf{u}_\partial, \mathbf{y}_\partial \rangle = \mathbf{u}_\partial^T D \mathbf{u}_\partial = 0. \quad (5.59)$$

We now invoke several notations in order to define a state-space port-Hamiltonian
model that represents the TFM locally at the geometry of Z_{ab}. Using (5.18),
(5.50), (5.20) and (5.21), we can straightforwardly deduce that

$$\mathbf{f}^{ab} = [\frac{d}{dt}q_1^{ab}, \frac{d}{dt}q_2^{ab}, \frac{d}{dt}q_3^{ab}, \frac{d}{dt}q_4^{ab}]^T. \quad (5.60)$$

The discrete Hamiltonian H_d^{ab} is expressed as:

$$H_d^{ab} = \int_{Z_{ab}} \left(\frac{(q_3^{ab})^2}{2q_1^{ab}(b-a)} + \frac{(q_4^{ab})^2}{2q_2^{ab}(b-a)} + \frac{q_1^{ab}c_g^2}{b-a} \ln\left(\frac{q_1^{ab}}{b-a} + \frac{c_\ell^2}{c_g^2}\frac{q_2^{ab}}{b-a} - \right. \right.$$
$$\left. \left. \frac{\beta}{c_g^2}(1-\alpha_g^d) \right) + \frac{q_2^{ab}c_\ell^2}{b-a} \ln\left(\frac{q_2^{ab}}{b-a} + \frac{c_g^2}{c_\ell^2}\frac{q_1^{ab}}{b-a} + \frac{\beta}{c_\ell^2}\alpha_g^d \right) + (1-\alpha_g^d)\beta \right) dz, \quad (5.61)$$

where $\alpha_g^d = -\frac{q_1^{ab}c_g^2}{2\beta(b-a)} - \frac{q_2^{ab}c_\ell^2}{2\beta(b-a)} + \frac{1}{2} + \Delta$, with

$$\Delta = \sqrt{\left(\frac{q_1^{ab}c_g^2}{2\beta(b-a)} + \frac{q_2^{ab}c_\ell^2}{2\beta(b-a)} - \frac{1}{2} \right)^2 + \frac{q_1^{ab}c_g^2}{\beta(b-a)}}.$$

Using the discrete Hamiltonian (5.61) along with (5.20), (5.24) and (5.53), $\mathbf{e}^{ab}$ can also be computed for numerical purposes. For the sake of simplicity, we will use the notation $\dot{Q}^{ab}$ instead of $\mathbf{f}^{ab}$, and ∇H_d^{ab} instead of $\mathbf{e}^{ab}$ in the sequel, i.e.,

$$\dot{Q}^{ab} := \mathbf{f}^{ab}, \quad \nabla H_d^{ab} := \mathbf{e}^{ab}.$$

Using these notations, the finite-dimensional pH model at the geometric domain Z_{ab} is given by

$$\begin{aligned} \dot{Q} &= J\nabla H + BU, \\ Y &= C\nabla H + DU, \end{aligned} \quad (5.62)$$

where we have omitted the superscript "ab" in the above representation for the sake of generality of the structure. The subscript "d" is also dropped from the Hamiltonian for the sake of clarity in the sequel. Moreover, $U := [U^l, U^r]^T = \mathbf{u}_\partial$ and $Y := [Y^l, Y^r]^T = \mathbf{y}_\partial$, where we define U^l, U^r, Y^l and Y^r as follows:

$$U^l = [e_{a1}^B, e_{a2}^B]^T, \quad U^r = [e_{b1}^B, e_{b2}^B]^T, \quad (5.63)$$

$$Y^l = [f_{a1}^B, f_{a2}^B]^T, \quad Y^r = [f_{b1}^B, f_{b2}^B]^T. \quad (5.64)$$

Each discretized lump has been shown to possess the properties of a finite-dimensional Dirac structure. Given this fact, the lumps can be interconnected in a power-preserving manner. Before delving into the power-preserving interconnection, we consider resistive setting in the scope of the TFM and develop the corresponding structure-preserving finite-dimensional realization.

Resistive setting (Lossy case)

The resistive effects can be incorporated while developing a state-space port-Hamiltonian model by considering the discrete approximation of different variables as introduced earlier in this section. However, the image representation of a Dirac structure will be different, and conditions for the existence of a Dirac structure will need to be modified in accordance. This is necessary since we need to approximate an extended state-dependent Stokes-Dirac structure, as in Theorem 5.2.4, by using admissible function spaces for different variables. Such an approximation helps to develop a finite-dimensional extended Dirac structure, which is crucial to quantify the preservation of the dissipated power upon spatial discretization.

Additional ingredients are briefly discussed next. We approximate the flow and effort variables (corresponding to the resistive port) on Z_{ab} as follows:

$$f_{R_j} = f_{R_j}^{ab}(t)\omega_{f,R_j}^{ab}(z), \quad e_{R_j} = e_{R_j}^{ab}(t)\omega_{e,R_j}^{ab}(z), \quad j = \{1,2\}. \quad (5.65)$$

Furthermore, consider that

$$\omega_{e,R_1}^{ab}(z) = \omega_{q_3}^{ab}(z), \quad \omega_{e,R_2}^{ab}(z) = \omega_{q_4}^{ab}(z). \quad (5.66)$$

We state the non-degenerated bilinear product on $(\mathscr{L}^2(Z_{ab})^6 \times \mathbb{R}^4)^2$:

$$\int_{Z^{ab}} \left(f_{q_1}e_{q_1} + f_{q_2}e_{q_2} + f_{q_3}e_{q_3} + f_{q_4}e_{q_4} + f_{R_1}e_{R_1} + f_{R_2}e_{R_2} \right) \, dz + b.t. \quad (5.67)$$

Substituting the discrete approximations in the above equation, we have:

$$\int_{Z_{ab}} \left(f_{q_1}^{ab}(t)\omega_{q_1}^{ab}(z)(e_{q_1}^a(t)\omega_{q_1}^a(z) + e_{q_1}^b(t)\omega_{q_1}^b(z)) + f_{q_2}^{ab}(t)\omega_{q_2}^{ab}(z)(e_{q_2}^a(t)\omega_{q_2}^a(z) + e_{q_2}^b(t)\omega_{q_2}^b(z)) + \right.$$

$$\left. f_{q_3}^{ab}(t)\omega_{q_3}^{ab}(z)(e_{q_3}^a(t)\omega_{q_3}^a(z) + e_{q_3}^b(t)\omega_{q_3}^b(z)) + f_{q_4}^{ab}(t)\omega_{q_4}^{ab}(z)(e_{q_4}^a(t)\omega_{q_4}^a(z) + e_{q_4}^b(t)\omega_{q_4}^b(z)) \right) dz +$$

$$\int_{Z_{ab}} \left(f_{R_1}^{ab}\omega_{f,R_1}^{ab}e_{R_1}^{ab}\omega_{e,R_1}^{ab} + f_{R_2}^{ab}\omega_{f,R_2}^{ab}e_{R_2}^{ab}\omega_{e,R_2}^{ab} \right) dz + b.t.. \quad (5.68)$$

Upon using (5.43) with $\alpha_i^a = \alpha_i^b = 0.5$, (5.68) can be simplified to obtain:

$$f_{q_1}^{ab}\frac{1}{2}(e_{q_1}^a + e_{q_1}^b) + f_{q_2}^{ab}\frac{1}{2}(e_{q_2}^a + e_{q_2}^b) + f_{q_3}^{ab}\frac{1}{2}(e_{q_3}^a + e_{q_3}^b) + f_{q_4}^{ab}\frac{1}{2}(e_{q_4}^a + e_{q_4}^b) +$$

$$f_{R_1}^{ab}e_{R_1}^{ab}\int_{Z_{ab}} \omega_{f,R_1}^{ab}\omega_{e,R_1}^{ab} \, dz + f_{R_2}^{ab}e_{R_2}^{ab}\int_{Z_{ab}} \omega_{f,R_2}^{ab}\omega_{e,R_2}^{ab} \, dz + b.t.. \quad (5.69)$$

Recalling that the boundary port variables will remain unchanged in the presence of dissipation (also see Remark 3.4.7, Chapter 3), and using (5.66), we can write (5.69) as follows:

$$-e_{R_1}^{ab}\frac{1}{2}(e_{q_3}^a+e_{q_3}^b)-e_{R_2}^{ab}\frac{1}{2}(e_{q_4}^a+e_{q_4}^b)+$$
$$f_{R_1}^{ab}e_{R_1}^{ab}\int_{Z_{ab}}\omega_{f,R_1}^{ab}\omega_{e,R_1}^{ab}\,\mathrm{d}z+f_{R_2}^{ab}e_{R_2}^{ab}\int_{Z_{ab}}\omega_{f,R_2}^{ab}\omega_{e,R_2}^{ab}\,\mathrm{d}z. \quad (5.70)$$

As per the notion of non-degenerate bilinear product, (5.70) should be zero. This holds if:

$$f_{R_1}^{ab}\int_{Z_{ab}}\omega_{f,R_1}^{ab}\omega_{e,R_1}^{ab}\,\mathrm{d}z=\frac{1}{2}(e_{q_3}^a+e_{q_3}^b),\ \text{ and }\ f_{R_2}^{ab}\int_{Z_{ab}}\omega_{f,R_2}^{ab}\omega_{e,R_2}^{ab}\,\mathrm{d}z=\frac{1}{2}(e_{q_4}^a+e_{q_4}^b).$$

Imposing

$$\int_{Z_{ab}}\omega_{f,R_1}^{ab}\omega_{e,R_1}^{ab}\,\mathrm{d}z=1,\quad\text{and}\quad\int_{Z_{ab}}\omega_{f,R_2}^{ab}\omega_{e,R_2}^{ab}\,\mathrm{d}z=1, \quad (5.71)$$

we have that

$$f_{R_1}^{ab}=\frac{1}{2}(e_{q_3}^a+e_{q_3}^b),\quad f_{R_2}^{ab}=\frac{1}{2}(e_{q_4}^a+e_{q_4}^b). \quad (5.72)$$

Let us now define

$$\mathbf{f}_{ext}=[f_{q_1}^{ab},f_{q_2}^{ab},f_{q_3}^{ab},f_{q_4}^{ab},f_{R_1}^{ab},f_{R_2}^{ab},f_{a1}^B,f_{a2}^B,f_{b1}^B,f_{b2}^B]^T, \quad (5.73a)$$

$$\mathbf{e}_{ext}=[e_{q_1}^{ab},e_{q_2}^{ab},e_{q_3}^{ab},e_{q_4}^{ab},e_{R_1}^{ab},e_{R_2}^{ab},e_{a1}^B,e_{a2}^B,e_{b1}^B,e_{b2}^B]^T, \quad (5.73b)$$

and $\Delta z:=b-a$. Using (5.52), (5.53), (5.72) and discrete approximation of (5.7), we then have

$$\mathbf{f}_{ext}=\underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \frac{q_1^{ab}}{\Delta z} & -\frac{q_1^{ab}}{\Delta z} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{q_2^{ab}}{\Delta z} & -\frac{q_2^{ab}}{\Delta z} & 0 & 0 \\ \frac{q_1^{ab}}{\Delta z} & -\frac{q_1^{ab}}{\Delta z} & 0 & 0 & \frac{2q_3^{ab}}{\Delta z} & -\frac{2q_3^{ab}}{\Delta z} & 0 & 0 & -1 & 0 \\ 0 & 0 & \frac{q_2^{ab}}{\Delta z} & -\frac{q_2^{ab}}{\Delta z} & 0 & 0 & \frac{2q_4^{ab}}{\Delta z} & -\frac{2q_4^{ab}}{\Delta z} & 0 & -1 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & \frac{q_1^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_3^{ab}}{\Delta z} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{q_2^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_4^{ab}}{\Delta z} & 0 & 0 \end{bmatrix}}_{F_{ext}^T}\begin{bmatrix} e_{q_1}^a \\ e_{q_1}^b \\ e_{q_2}^a \\ e_{q_2}^b \\ e_{q_3}^a \\ e_{q_3}^b \\ e_{q_4}^a \\ e_{q_4}^b \\ e_{R_1}^{ab} \\ e_{R_2}^{ab} \end{bmatrix}, \quad (5.74)$$

$$\mathbf{e}_{ext} = \underbrace{\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{q_1^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_3^{ab}}{\Delta z} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{q_2^{ab}}{\Delta z} & 0 & 0 & 0 & \frac{q_4^{ab}}{\Delta z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{E_{ext}^T} \begin{bmatrix} e_{q_1}^a \\ e_{q_1}^b \\ e_{q_2}^a \\ e_{q_2}^b \\ e_{q_3}^a \\ e_{q_3}^b \\ e_{q_4}^a \\ e_{q_4}^b \\ e_{R_1}^{ab} \\ e_{R_2}^{ab} \end{bmatrix}. \tag{5.75}$$

Theorem 5.3.9 *Consider* $\mathscr{E}_{ext} \times \mathscr{F}_{ext} = \mathbb{R}^{10} \times \mathbb{R}^{10}$. *The subspace*

$$\mathscr{D}_{TFM}^{ext} = \{(\mathbf{e}_{ext}, \mathbf{f}_{ext}) \in \mathscr{E}_{ext} \times \mathscr{F}_{ext} \mid \exists \lambda_{ext} \in \mathbb{R}^{10} \ s.t. \ \mathbf{e}_{ext} = E_{ext}^T \lambda_{ext}, \mathbf{f}_{ext} = F_{ext}^T \lambda_{ext}\}, \tag{5.76}$$

with E_{ext} *and* F_{ext} *defined in* (5.75) *and* (5.74), *respectively, is a Dirac structure.*

Proof: It is straightforward to check that F_{ext} and E_{ext} introduced respectively in (5.74) and (5.75) fulfill the two conditions, outlined in Definition 5.3.2, for all discrete states $q_1^{ab}, q_2^{ab}, q_3^{ab}$ and q_4^{ab}. It is also clearly observable that $\mathscr{D}_{TFM}^{ext}$ is a Dirac structure irrespective of the value of the spatial step Δz. ∎

The above obtained continuous-time finite-dimensional structure $\mathscr{D}_{TFM}^{ext}$ (5.76), along with the approximate form of the closure relation (5.6) in consensus with (5.66), (5.71) and (5.72), represents the system that results after structure-preserving spatial discretization of the TFM with resistive effects. Unlike a state-space port-Hamiltonian ODE model obtained in the lossless case, we obtain a port-Hamiltonian DAE model in the lossy case. It is also worth mentioning that each discretized lump can be interconnected in a power-preserving manner due to the nature of $\mathscr{D}_{TFM}^{ext}$.

5.3.2 Spatial Discretization of Drift Flux Model

Similar to the discretization of the TFM in Section 5.3.1, we use the mFEM for structure-preserving (spatial) discretization of the DFM (without slip). We again apply the procedure to a single discretized lump $Z_{ab} = [a, b]$, with $0 \le a < b \le L$.

We will only consider the lossless equivalent of the DFM and develop a corresponding finite-dimensional realization. We refer the reader to the lossy case of the TFM in order to develop a structure-preserving discretization methodology of the DFM in the presence of resistive effects. Similar to the TFM, the resistive effects do not pose significant additional challenges as the port-Hamiltonian representation of the DFM is also devoid of any operator-type terms in the resistive matrix; see Theorem 5.2.6.

The basis function profile, used for discrete approximation, is recalled below:

$$\omega_{m_g}^{ab}(z) = \omega_{m_\ell}^{ab}(z) = \omega_v^{ab}(z) = \frac{1}{b-a}, \tag{5.77a}$$

$$\omega_{m_g}^{a}(z) = \omega_{m_\ell}^{a}(z) = \omega_v^{a}(z) = \frac{b-z}{b-a}, \tag{5.77b}$$

$$\omega_{m_g}^{b}(z) = \omega_{m_\ell}^{b}(z) = \omega_v^{b}(z) = \frac{z-a}{b-a}. \tag{5.77c}$$

The basis function approximation (5.77) possesses the following properties:

$$\int_a^b \omega_{m_g}^{ab}(z)\omega_{m_g}^{a}(z)\,dz = 0.5. \tag{5.78}$$

A similar property also holds for other state variables, i.e., m_ℓ and v. The finite-dimensional equivalent of (5.9) takes the following form:

$$f_{m_g}^{ab} = -\frac{A_g^{ab}}{\Delta z}\left(e_v^b - e_v^a\right), \tag{5.79a}$$

$$f_{m_\ell}^{ab} = -\frac{A_\ell^{ab}}{\Delta z}\left(e_v^b - e_v^a\right), \tag{5.79b}$$

$$f_v^{ab} = -\frac{A_g^{ab}}{\Delta z}\left(e_{m_g}^b - e_{m_g}^a\right) - \frac{A_\ell^{ab}}{\Delta z}\left(e_{m_\ell}^b - e_{m_\ell}^a\right), \tag{5.79c}$$

where

$$A_g^{ab} := \frac{m_g^{ab}}{m_g^{ab} + m_\ell^{ab}}, \quad A_\ell^{ab} := \frac{m_\ell^{ab}}{m_g^{ab} + m_\ell^{ab}}. \tag{5.80}$$

Similar to (5.53), an alternative definition of effort-type variables in the scope

of the DFM is given by:

$$e_{m_g}^{ab} = \frac{1}{2}(e_{m_g}^a + e_{m_g}^b), \tag{5.81}$$

$$e_{m_\ell}^{ab} = \frac{1}{2}(e_{m_\ell}^a + e_{m_\ell}^b), \tag{5.82}$$

$$e_v^{ab} = \frac{1}{2}(e_v^a + e_v^b). \tag{5.83}$$

Let us define

$$\mathbf{f}_d = [f_{m_g}^{ab}, f_{m_\ell}^{ab}, f_v^{ab}, f_a^B, f_b^B]^T, \tag{5.84}$$

and

$$\mathbf{e}_d = [e_{m_g}^{ab}, e_{m_\ell}^{ab}, e_v^{ab}, e_a^B, e_b^B]^T. \tag{5.85}$$

Using (5.79), (5.81) and discrete approximation of (5.12), we have

$$\mathbf{f}_d = \underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 & \frac{A_g^{ab}}{\Delta z} & -\frac{A_g^{ab}}{\Delta z} \\ 0 & 0 & 0 & 0 & \frac{A_\ell^{ab}}{\Delta z} & -\frac{A_\ell^{ab}}{\Delta z} \\ \frac{A_g^{ab}}{\Delta z} & -\frac{A_g^{ab}}{\Delta z} & \frac{A_\ell^{ab}}{\Delta z} & -\frac{A_\ell^{ab}}{\Delta z} & 0 & 0 \\ -\frac{A_g^{ab}}{\Delta z} & 0 & -\frac{A_\ell^{ab}}{\Delta z} & 0 & 0 & 0 \\ 0 & \frac{A_g^{ab}}{\Delta z} & 0 & \frac{A_\ell^{ab}}{\Delta z} & 0 & 0 \end{pmatrix}}_{F_d^T} \begin{pmatrix} e_{m_g}^a \\ e_{m_g}^b \\ e_{m_\ell}^a \\ e_{m_\ell}^b \\ e_v^a \\ e_v^b \end{pmatrix}, \tag{5.86}$$

and

$$\mathbf{e}_d = \underbrace{\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}_{E_d^T} \begin{pmatrix} e_{m_g}^a \\ e_{m_g}^b \\ e_{m_\ell}^a \\ e_{m_\ell}^b \\ e_v^a \\ e_v^b \end{pmatrix}. \tag{5.87}$$

Theorem 5.3.10 *Consider $\mathscr{E}_d \times \mathscr{F}_d = \mathbb{R}^5 \times \mathbb{R}^5$. The subspace*

$$\mathscr{D}_{DFM} = \{(\mathbf{e}_d, \mathbf{f}_d) \in \mathscr{E}_d \times \mathscr{F}_d \mid \exists \lambda \in \mathbb{R}^6 \ \ s.t. \ \mathbf{e}_d = E_d^T \lambda, \mathbf{f} = F_d^T \lambda\}, \tag{5.88}$$

with E_d and F_d defined in (5.87) and (5.86), respectively, is a Dirac structure.

Proof: The two conditions that need to hold in order to ensure that $\mathscr{D}_{DFM}$ is a Dirac structure are as follows:
(i) $E_d F_d^T + F_d E_d^T = 0$, and *(ii)* rank($[E_d \quad F_d]$) = 5 (also see Definition 5.3.2). It is straightforward to check that F_d and E_d introduced respectively in (5.86) and (5.87) fulfill the above two conditions for all discrete states m_g^{ab}, m_ℓ^{ab}, and v^{ab}. It is also clearly observable that $\mathscr{D}_{DFM}$ is a Dirac structure irrespective of the value of the spatial step Δz. ■

Remark 5.3.11 *Reversing the role of the boundary flow and effort variables will not impact the existence of the above proved finite-dimensional Dirac structure. This is straightforward to identify by using the basic principles of linear algebra.*

Let us consider $\mathbf{f}_{dm} = \mathbf{f}_d$ and $\mathbf{e}_{dm} = \mathbf{e}_d$. Using the observation in Remark 5.3.11, we have:

$$
\mathbf{f}_{dm} = \underbrace{\begin{pmatrix}
0 & 0 & 0 & 0 & \frac{A_g^{ab}}{\Delta z} & -\frac{A_g^{ab}}{\Delta z} \\
0 & 0 & 0 & 0 & \frac{A_\ell^{ab}}{\Delta z} & -\frac{A_\ell^{ab}}{\Delta z} \\
\frac{A_g^{ab}}{\Delta z} & -\frac{A_g^{ab}}{\Delta z} & \frac{A_\ell^{ab}}{\Delta z} & -\frac{A_\ell^{ab}}{\Delta z} & 0 & 0 \\
-\frac{A_g^{ab}}{\Delta z} & 0 & -\frac{A_\ell^{ab}}{\Delta z} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}}_{F_{dm}^T}
\underbrace{\begin{pmatrix}
e_{m_g}^a \\
e_{m_g}^b \\
e_{m_\ell}^a \\
e_{m_\ell}^b \\
e_v^a \\
e_v^b
\end{pmatrix}}_{\lambda_{dm}},
\quad
\mathbf{e}_{dm} = \underbrace{\begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & \frac{A_g^{ab}}{\Delta z} & 0 & \frac{A_\ell^{ab}}{\Delta z} & 0 & 0
\end{pmatrix}}_{E_{dm}^T}
\begin{pmatrix}
e_{m_g}^a \\
e_{m_g}^b \\
e_{m_\ell}^a \\
e_{m_\ell}^b \\
e_v^a \\
e_v^b
\end{pmatrix}.
$$

$$\tag{5.89}$$

In order to develop an explicit representation of $\mathbf{f}_{dm}$ in terms of $\mathbf{e}_{dm}$, we need to eliminate λ_{dm} and eventually invert E_{dm}^T. We can clearly observe that the matrices F_{dm}^T and E_{dm}^T in (5.89) are not square and, hence, not invertible (in a usual sense). However, we can exploit the principle of Moore-Penrose pseudo inverse. It can be observed that E_{dm}^T is a full rank 5×6 matrix, i.e., rank(E_{dm}^T) = min(5,6) = 5. The rows of E_{dm}^T are linearly independent, and in such a setting, the right-inverse of E_{dm}^T can be easily obtained. From now on, we will denote the right-inverse of E_{dm}^T as $(E_{dm}^T)_r^{-1}$. An explicit representation can now be obtained as follows:

$$\mathbf{f}_{dm} = F_{dm}^T (E_{dm}^T)_r^{-1} \mathbf{e}_{dm}. \tag{5.90}$$

The flow-effort relations are given by:

$$
\begin{aligned}
\mathbf{f}_d^{ab} &= J_d \mathbf{e}_d^{ab} + B_d \mathbf{u}_{\partial,d}, \\
\mathbf{y}_{\partial,d} &= C_d \mathbf{e}_d^{ab} + D_d \mathbf{u}_{\partial,d},
\end{aligned}
\tag{5.91}
$$

where $\mathbf{f}_d^{ab} = [f_{m_g}^{ab}, f_{m_\ell}^{ab}, f_v^{ab}]^T$, $\mathbf{e}_d^{ab} = [e_{m_g}^{ab}, e_{m_\ell}^{ab}, e_v^{ab}]^T$, $\mathbf{y}_{\partial,d} = [f_a^B, f_b^B]^T$ and $\mathbf{u}_{\partial,d} = [e_a^B, e_b^B]^T$. The matrices J_d and D_d are as follows:

$$
J_d = \begin{bmatrix} 0 & 0 & -2\frac{A_g^{ab}}{\Delta z} \\ 0 & 0 & -2\frac{A_\ell^{ab}}{\Delta z} \\ 2\frac{A_g^{ab}}{\Delta z} & 2\frac{A_\ell^{ab}}{\Delta z} & 0 \end{bmatrix}, D_d = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.
$$

The matrices B_d and C_d can also be deduced. We can observe that both J_d and D_d are skew-symmetric. Our claim, as before, regarding the zero instantaneous power due to the feedthrough matrix D_d holds since the matrix D_d is skew-symmetric.

The main challenge of obtaining the required properties on system matrices corresponding to each individual lump have been fully addressed. It is now straightforward to define a state-space port-Hamiltonian model that represents the DFM (without slip) locally at the geometry of Z_{ab}. We omit the related discussion and instead refer the reader to follow the principles used in the scope of the TFM in Section 5.3.1.

Each discretized lump, both in the scope of the TFM and the DFM, has been shown to possess the properties of a finite-dimensional Dirac structure. Given this fact, the lumps can be interconnected to obtain an aggregated finite-dimensional port-Hamiltonian representation, which is shown next.

5.4   Power-preserving interconnection

In this section, we will base our study on the TFM. We show the procedure to interconnect two lumps (in series), obtained in the scope of the TFM, in a power-preserving manner. A similar analysis can be carried out to develop an aggregated finite-dimensional port-Hamiltonian model for the DFM.

We consider that the spatial domain $\Omega = [0, L]$ with $0 < a < b < c < L$ is discretized into N lumps. We focus our attention on two discretized lumps, i.e., $Z_{ab} = [a, b]$ and $Z_{bc} = [b, c]$. Using the structure in (5.62), the first discretized lump across Z_{ab} is governed by the following state ODEs:

$$
\begin{aligned}
\dot{Q}_1 &= J_1 \nabla H_1 + B_1 U_1 = J_1 \nabla H_1 + B_1^l U_1^l + B_1^r U_1^r, \\
Y_1 &= C_1 \nabla H_1 + D_1^l U_1^l + D_1^r U_1^r,
\end{aligned}
\tag{5.92}
$$

where B_1^l and D_1^l, respectively, refer to the matrix composed from the first two columns of B and D introduced earlier. Anologously, B_1^r and D_1^r are composed

of the last two columns of B and D, respectively. Matrices B_1^l, D_1^l, B_1^r, and D_1^r are each of size 4×2. The subscript "1" refers to the quantities (introduced in (5.62), (5.63) and (5.64)) for the first element Z_{ab}.

Similarly, for the second discretized lump Z_{bc}, we have

$$
\begin{aligned}
\dot{Q}_2 &= J_2 \nabla H_2 + B_2 U_2 = J_2 \nabla H_2 + B_2^l U_2^l + B_2^r U_2^r, \\
Y_2 &= C_2 \nabla H_2 + D_2^l U_2^l + D_2^r U_2^r,
\end{aligned}
\tag{5.93}
$$

where the subscript "2" refers to the quantities corresponding to the second finite-element Z_{bc}.

The following power-preserving interconnection relations hold:

$$
\begin{bmatrix} U_1^r \\ U_2^l \end{bmatrix} = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \begin{bmatrix} Y_1^r \\ Y_2^l \end{bmatrix} = \begin{bmatrix} -C_2^{12} \nabla H_2 - D_2^{r,12} U_2^r \\ C_1^{34} \nabla H_1 + D_1^{l,34} U_1^l \end{bmatrix},
$$

where the subscripts "1" and "2" carry the same meaning as before. Matrix C_2^{12} refers to the first two rows, i.e., row 1 and 2 of the matrix C for the second element Z_{bc}. Similarly, matrix C_1^{34} refers to the last two rows, i.e., row 3 and 4 of the matrix C for the first element Z_{ab}. Analogously, $D_2^{r,12}$ refers to the first two rows, i.e., row 1 and 2 of the matrix D_2^r for the second element Z_{bc}. Similar explanation holds for $D_1^{l,34}$.

Interconnecting the 2 discretized lumps in a power-preserving manner yields the following structure:

$$
\begin{bmatrix} \dot{Q}_1 \\ \dot{Q}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} J_1 & -B_1^r C_2^{12} \\ B_2^l C_1^{34} & J_2 \end{bmatrix}}_{J_{\text{assembled}}} \begin{bmatrix} \nabla H_1 \\ \nabla H_2 \end{bmatrix} + \begin{bmatrix} B_1^l & -B_1^r D_2^{r,12} \\ B_2^l D_1^{l,34} & B_2^r \end{bmatrix} \begin{bmatrix} U_1^l \\ U_2^r \end{bmatrix},
\tag{5.94}
$$

$$
\begin{bmatrix} Y_1^l \\ Y_2^r \end{bmatrix} = \begin{bmatrix} C_1^{12} & -D_1^{r,12} C_2^{12} \\ D_2^{l,34} C_1^{34} & C_2^{34} \end{bmatrix} \begin{bmatrix} \nabla H_1 \\ \nabla H_2 \end{bmatrix} + \begin{bmatrix} \mathbf{O}_{2\times2} & -D_1^{r,12} D_2^{r,12} \\ D_2^{l,34} D_1^{l,34} & \mathbf{O}_{2\times2} \end{bmatrix} \begin{bmatrix} U_1^l \\ U_2^r \end{bmatrix},
\tag{5.95}
$$

where C_1^{12}, $D_1^{r,12}$, C_2^{34} and $D_2^{l,34}$ can be computed by following the notational conventions introduced earlier.

Remark 5.4.1 *The assembled system is conservative in the sense that $\frac{dH}{dt} = \frac{dH_1}{dt} + \frac{dH_2}{dt} = -\left((Y_1^\ell(t))^T U_1^\ell(t) + (Y_2^r(t))^T U_2^r(t) \right)$ represents the net power across the left and right ports of the assembled system, and the system has no resistive effects in its*

dynamic behaviour. This behaviour is inherent and a representative feature of hyperbolic partial differential equations of the type that are studied here. The conservative (lossless) nature of the approximate mathematical model is therefore physically meaningful and replicates the conservative (lossless) nature of the infinite-dimensional port-Hamiltonian model.

We have described the procedure to interconnect two discretized lumps. Analogously, N discretized lumps can be interconnected in a power-preserving manner to obtain a complete port-Hamiltonian model as follows:

$$
\begin{bmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ \dot{Q}_3 \\ \vdots \\ \vdots \\ \dot{Q}_{n-1} \\ \dot{Q}_n \end{bmatrix} = \begin{bmatrix} J_1 & -B_1^r C_2^{12} & B_1^r D_2^{r,12} C_3^{12} & \dots & \dots & M_{1,n-1} & M_{1,n} \\ B_2^l C_1^{34} & J_2 & -B_2^r C_3^{12} & \dots & \dots & M_{2,n-1} & M_{2,n} \\ B_3^l D_2^{l,34} C_1^{34} & B_3^l C_2^{34} & J_3 & \dots & \dots & M_{3,n-1} & M_{3,n} \\ \dots & \dots & \dots & \ddots & \ddots & \dots & \dots \\ \dots & \dots & \dots & \ddots & \ddots & \dots & \dots \\ M_{n-1,1} & M_{n-1,2} & M_{n-1,3} & \dots & \dots & J_{n-1} & -B_{n-1}^r C_n^{12} \\ M_{n,1} & M_{n,2} & M_{n,3} & \dots & \dots & B_n^l C_{n-1}^{34} & J_n \end{bmatrix} \begin{bmatrix} \nabla H_1 \\ \nabla H_2 \\ \nabla H_3 \\ \vdots \\ \vdots \\ \nabla H_{n-1} \\ \nabla H_n \end{bmatrix}
$$

$$
+ \begin{bmatrix} B_1^l & \mp B_1^r D_2^{r,12}\dots D_n^{r,12} \\ B_2^l D_1^{l,34} & \pm B_2^r D_3^{r,12}\dots D_n^{r,12} \\ B_3^l D_2^{l,34} D_1^{l,34} & \mp B_3^r D_4^{r,12}\dots D_n^{r,12} \\ \dots & \dots \\ \dots & \dots \\ B_{n-1}^l D_{n-2}^{l,34}\dots D_1^{l,34} & \mp B_{n-1}^r D_n^{r,12} \\ B_n^l D_{n-1}^{l,34}\dots D_1^{l,34} & \pm B_n^r \end{bmatrix} \begin{bmatrix} U_1^l \\ U_n^r \end{bmatrix}, \quad (5.96)
$$

$$
\begin{bmatrix} Y_1^l \\ Y_n^r \end{bmatrix} = \begin{bmatrix} C_1^{12} & -D_1^{r,12} C_2^{12} & G_{1,3} & \dots & \dots & G_{1,n-1} & G_{1,n} \\ G_{n,1} & G_{n,2} & G_{n,3} & \dots & \dots & D_n^{l,34} C_{n-1}^{34} & C_n^{34} \end{bmatrix} \begin{bmatrix} \nabla H_1 \\ \nabla H_2 \\ \nabla H_3 \\ \vdots \\ \vdots \\ \nabla H_{n-1} \\ \nabla H_n \end{bmatrix}
$$

$$
+ \begin{bmatrix} 0 & -D_1^{r,12} D_2^{r,12}\dots D_{n-1}^{r,12} D_n^{r,12} \\ D_1^{r,12} D_2^{r,12}\dots D_{n-1}^{r,12} D_n^{r,12} & 0 \end{bmatrix} \begin{bmatrix} U_1^l \\ U_n^r \end{bmatrix}, \quad (5.97)
$$

where

$$M_{1,n-1} = (-1)^n B_1^r D_2^{r,12} D_3^{r,12} \ldots D_{n-2}^{r,12} C_{n-1}^{12}, \quad M_{1,n} = (-1)^{1+n} B_1^r D_2^{r,12} D_3^{r,12} \ldots D_{n-1}^{r,12} C_n^{12},$$
$$M_{2,n-1} = (-1)^{n+1} B_2^r D_3^{r,12} D_4^{r,12} \ldots D_{n-2}^{r,12} C_{n-1}^{12}, \quad M_{2,n} = (-1)^{2+n} B_2^r D_3^{r,12} D_4^{r,12} \ldots D_{n-1}^{r,12} C_n^{12},$$
$$M_{3,n-1} = (-1)^{n+2} B_3^r D_4^{r,12} D_5^{r,12} \ldots D_{n-2}^{r,12} C_{n-1}^{12}, \quad M_{3,n} = (-1)^{3+n} B_3^r D_4^{r,12} D_5^{r,12} \ldots D_{n-1}^{r,12} C_n^{12},$$
$$M_{n-1,1} = B_{n-1}^l D_{n-2}^{l,34} \ldots D_2^{l,34} C_1^{34}, \quad M_{n-1,2} = B_{n-1}^l D_{n-2}^{l,34} \ldots D_3^{l,34} C_2^{34},$$
$$M_{n-1,3} = B_{n-1}^l D_{n-2}^{l,34} \ldots D_4^{l,34} C_3^{34}, \quad M_{n,1} = B_n^l D_{n-1}^{l,34} \ldots D_2^{l,34} C_1^{34},$$
$$M_{n,2} = B_n^l D_{n-1}^{l,34} \ldots D_3^{l,34} C_2^{34}, \quad M_{n,3} = B_n^l D_{n-1}^{l,34} \ldots D_4^{l,34} C_3^{34}, \quad (5.98)$$

and

$$G_{1,3} = D_1^{r,12} D_2^{r,12} C_3^{12}, \quad G_{1,n-1} = \pm D_1^{r,12} D_2^{r,12} \ldots D_{n-3}^{r,12} D_{n-2}^{r,12} C_{n-1}^{12},$$
$$G_{1,n} = \mp D_1^{r,12} D_2^{r,12} \ldots D_{n-2}^{r,12} D_{n-1}^{r,12} C_n^{12}, \quad G_{n,1} = D_n^{l,34} D_{n-1}^{l,34} \ldots D_2^{l,34} C_1^{34},$$
$$G_{n,2} = D_n^{l,34} D_{n-1}^{l,34} \ldots D_3^{l,34} C_2^{34}, \quad G_{n,3} = D_n^{l,34} D_{n-1}^{l,34} \ldots D_4^{l,34} C_3^{34}. \quad (5.99)$$

We now have all the ingredients to proceed with the development of the structure-preserving temporal discretization method.

5.5 Structure-preserving Temporal Discretization

Explicit methods such as Runge Kutta methods or linear multi-step methods fail to preserve the discrete-time port-Hamiltonian structure. We build on the work of [96] and exploit Gauss-Legendre collocation methods [80] for the temporal discretization of the state-space port-Hamiltonian representations obtained in the scope of the two-phase flow models. These collocation methods are a well known family of A-stable numerical methods for ordinary differential equations. The symplectic version of these collocation methods is known to be structure-preserving in the sense of the Dirac structure [96] and is known to work well for both conservative and dissipative systems [80]. It is also well known that the symplectic time-integration methods are essentially implicit methods, particularly for the non-quadratic Hamiltonian setting. At this point, it is worth mentioning that preserving the symplectic nature is not equivalent to preserving a discrete energy balance in an exact manner. For instance, it is known that symplectic schemes can preserve an exact discrete energy balance only conditionally even in the case of linear port-Hamiltonian systems with a quadratic

Hamiltonian functional [96]. A standard implementation of existing symplectic schemes in the scope of our work does not essentially guarantee the exact preservation of discrete energy balance as we are dealing with non-linear port-Hamiltonian systems with non-quadratic Hamiltonian functional and, hence, deserves further careful treatments. Furthermore, we have *state-dependent* system matrices and feedthrough terms in the continuous-time finite-dimensional realization of the models of interest.

We recall the power balance equation obtained in the context of single (spatially) discretized lump. The power balance equation reads as

$$\mathbf{e}^{ab}(t)^T \mathbf{f}^{ab}(t) + \mathbf{y}_\partial(t)^T \mathbf{u}_\partial(t) = 0, \tag{5.100}$$

where $\mathbf{e}^{ab}, \mathbf{f}^{ab}, \mathbf{y}_\partial$ and $\mathbf{u}_\partial$ carry the same meaning as introduced earlier in Section 5.3. On integrating the power balance equation (5.100) over every time interval $[t_1, t_2]$, we obtain

$$\int_{t_1}^{t_2} \left(\mathbf{e}^{ab}(s)^T \mathbf{f}^{ab}(s) + \mathbf{y}_\partial(s)^T \mathbf{u}_\partial(s) \right) ds = 0. \tag{5.101}$$

Preserving the energy balance in the integral form (as in (5.101)) will form the key step in obtaining a fully discrete port-Hamiltonian model. We recall that a (continuous-time) state-space port-Hamiltonian model has been defined with respect to a continuous-time Dirac structure; see Theorem 5.3.5, Theorem 5.3.9 and Theorem 5.3.10. Similarly, a discrete-time port-Hamiltonian model can be defined with respect to a discrete-time Dirac structure. We seek to obtain a (finite-dimensional) discrete-time Dirac structure by applying the Gauss-Legendre collocation method to the class of state-space port-Hamiltonian systems and by defining the discrete flow, effort, input and output variables for every sampling interval in an appropriate manner.

5.5.1   Discrete approximation of variables

Similar to [96], we obtain the numerical approximation of the solution $\mathbf{x}^{ab}(t) \in \mathbb{R}^n$ in the form of a vector $\bar{\mathbf{x}}^{ab}(t) \in \mathbb{R}^n$ of collocation polynomials of degree s. Here, $\mathbf{x}^{ab}$ indicates the general vector and can symbolize the flow, effort, input or output vector in the sequel. We consider equidistant or uniform sampling intervals $I^k = [t_0^k, t_{s+1}^k] = [(k-1)h, kh], k \in \mathbb{N}$, where $t_{s+1}^k = t_0^k + h$. Generically speaking, we consider the initial value $\mathbf{x}_0^{ab,k} := \bar{\mathbf{x}}^{ab}(t_0^k) = \mathbf{x}^{ab}(t_0^k)$ to be known. The continuous numerical solution is obtained by matching the time derivative

of the vector of polynomials to the right-hand side of the ODE model description in the s collocation points.

The flow, effort, input and output coordinates can be defined based on identical or different collocation points. In this chapter, similar to [96], the flow, effort, input and output coordinates are defined based on identical collocation points.

Consider the general form of the (spatially discretized) port-Hamiltonian model as

$$\mathbf{f}^{ab} = J\mathbf{e}^{ab} + B\mathbf{u}_\partial, \tag{5.102a}$$

$$\mathbf{y}_\partial = C\mathbf{e}^{ab} + \mathbf{D}u_\partial \tag{5.102b}$$

with $\mathbf{f}^{ab} \in \mathbb{R}^n, \mathbf{e}^{ab} \in \mathbb{R}^n, \mathbf{y}_\partial \in \mathbb{R}^m$ and $\mathbf{u}_\partial \in \mathbb{R}^m$.
The interpolation formula for the state/flow variables reads as follows:

$$\dot{\mathbf{q}}^{ab}(t_0^k + \tau h) := \tilde{\mathbf{f}}^{ab}(t_0^k + \tau h) = \sum_{i=1}^{s} \mathbf{f}_i^{ab,k} l_i(\tau), \tag{5.103}$$

where l_i is the i−th Lagrange interpolation polynomial

$$l_i(\tau) = \prod_{\substack{j=1 \\ j \neq i}}^{s} \frac{\tau - c_j}{c_i - c_j}, \quad \tau \in [0,1]. \tag{5.104}$$

The numerical solution $\tilde{\mathbf{q}}^{ab}(t_0^k + \tau h)$ can be obtained by integrating (5.103) as shown below:

$$\tilde{\mathbf{q}}^{ab}(t_0^k + \tau h) = \tilde{\mathbf{q}}^{ab}(t_0^k) + h \sum_{j=1}^{s} \left(\mathbf{f}_j^{ab,k} \int_0^\tau l_j(\sigma) d\sigma \right). \tag{5.105}$$

The numerical solution $\mathbf{q}_i^{ab,k} := \tilde{\mathbf{q}}^{ab}(t_i^k)$ inside and at the end of the interval I^k can be computed in the following manner:

$$\mathbf{q}_i^{ab,k} = \mathbf{q}_0^{ab,k} + h \sum_{j=1}^{s} a_{ij} \mathbf{f}_j^{ab,k}, \quad i = 1, 2, \dots, s,$$

$$\mathbf{q}_{s+1}^{ab,k} = \mathbf{q}_0^{ab,k} + h \sum_{j=1}^{s} b_j \mathbf{f}_j^{ab,k}, \tag{5.106}$$

where a_{ij} and b_j are the coefficients of the Butcher Tableau for the Runge-Kutta interpretation of the collocation method [80].

The interconnection matrix J and the input matrix B need to be evaluated at the flow collocation points c_i. These evaluated matrices are given by:

$$J_i^k := J(\mathbf{q}_i^{ab,k}), \quad B_i^k := B(\mathbf{q}_i^{ab,k}). \tag{5.107}$$

Similarly, we also need to evaluate the effort vector $\mathbf{e}^{ab}$ and the input vector $\mathbf{u}_\partial$ at the flow collocation points. Such an evaluation yields:

$$u_{\partial i}^{\,k} := \mathbf{u}_\partial(t_0^k + c_i h), \quad e_i^{ab,k} := \mathbf{e}^{ab}\,|_{\mathbf{q}_i^{ab,k}}. \tag{5.108}$$

The polynomial approximation of the effort vector $\mathbf{e}^{ab}$ is given as follows:

$$\tilde{\mathbf{e}}^{ab}(t_0^k + \tau h) = \sum_{i=1}^{s} e_i^{ab,k} l_i(\tau). \tag{5.109}$$

Similar to the approximation of the effort variables, the polynomial approximation of the input vector $\mathbf{u}_\partial$ is given by

$$\tilde{\mathbf{u}}_\partial(t_0^k + \tau h) = \sum_{i=1}^{s} u_{\partial i}^{\,k} l_i(\tau). \tag{5.110}$$

The flow, effort, input and output coordinates are respectively merged in the discrete-time flow, effort, input and output vector as follows:

$$f^{ab,k} := [(f_1^{ab,k})^T \, \dots \, (f_s^{ab,k})^T]^T \in \mathbb{R}^{sn}, \tag{5.111a}$$

$$e^{ab,k} := [(e_1^{ab,k})^T \, \dots \, (e_s^{ab,k})^T]^T \in \mathbb{R}^{sn}, \tag{5.111b}$$

$$u_\partial^k := [(u_{\partial 1}^{\,k})^T \, \dots \, (u_{\partial s}^{\,k})^T]^T \in \mathbb{R}^{sm}, \tag{5.111c}$$

$$y_\partial^k := [(y_{\partial 1}^{\,k})^T \, \dots \, (y_{\partial s}^{\,k})^T]^T \in \mathbb{R}^{sm}. \tag{5.111d}$$

The discrete-time counterpart of (5.102a) is given by

$$f_i^{ab,k} = J_i^k e_i^{ab,k} + B_i^k u_{\partial i}^{\,k}, \quad i = 1, 2, \dots, s. \tag{5.112}$$

On the sampling interval I^k, (5.112) can be rewritten as

$$f^{ab,k} = J^k e^{ab,k} + B^k u_\partial^k, \tag{5.113}$$

where the block-diagonal matrix J^k is given by

$$J^k = -(J^k)^T = \text{blockdiag}(J_1^k, \ldots, J_s^k), \tag{5.114}$$

and the block-diagonal matrix B^k is given by

$$B^k = \text{blockdiag}(B_1^k, \ldots, B_s^k). \tag{5.115}$$

We consider the discrete-time equivalent of the output equation (as in (5.102b)) as follows:

$$y_\partial^k = -(B^k)^T e^{ab,k} + D^k u_\partial^k, \tag{5.116}$$

where we have utilized the relation between system matrices B and C, i.e., $C = -B^T$ (see Section 5.3.1). Here, the block-diagonal matrix D^k is given by

$$D^k = \text{blockdiag}(D_1^k, \ldots, D_s^k). \tag{5.117}$$

5.5.2   Discrete-time Dirac structure

We seek to obtain an approximation of the supplied energy, denoted as $\Delta \tilde{H}^k$, on the sampling interval I^k. To attain this, we integrate the polynomial approximation of the instantaneous power $\tilde{\mathbf{e}}^{ab}(t_0^k + \tau h)^T \tilde{\mathbf{f}}^{ab}(t_0^k + \tau h)$ over the normalized time interval $[0,1]$, and obtain:

$$\Delta \tilde{H}^k := h \int_0^1 \left(\tilde{\mathbf{e}}^{ab}(t_0^k + \tau h) \right)^T \tilde{\mathbf{f}}^{ab}(t_0^k + \tau h) \, d\tau. \tag{5.118}$$

Substituting the definitions of the polynomial approximation of the flow and the effort variables, we can re-write (5.118) in the following manner:

$$\begin{aligned}
\Delta \tilde{H}^k &= h \int_0^1 \left(\sum_{i=1}^s e_i^{ab,k} l_i(\tau) \right)^T \left(\sum_{i=1}^s f_i^{ab,k} l_i(\tau) \right) d\tau, \\
&= h(e^{ab,k})^T M f^{ab,k},
\end{aligned} \tag{5.119}$$

where $M \in \mathbb{R}^{sn \times sn}$ is a symmetric matrix given by:

$$M = \begin{bmatrix} m_{11} & \ldots & m_{1s} \\ \vdots & \ddots & \vdots \\ m_{s1} & \ldots & m_{ss} \end{bmatrix} \otimes I_{n \times n}, \quad \text{with } m_{ij} = \int_0^1 l_i(\tau) l_j(\tau) \, d\tau. \tag{5.120}$$

Substituting $f^{ab,k}$ in (5.119) with the right-hand side of (5.113), we have:

$$\Delta \tilde{H}^k = h(e^{ab,k})^T M\Big(J^k e^{ab,k} + B^k u_\partial^k\Big), \tag{5.121}$$

which is re-written as:

$$\Delta \tilde{H}^k = h(e^{ab,k})^T M J^k e^{ab,k} + h(e^{ab,k})^T M B^k u_\partial^k. \tag{5.122}$$

We recall that

$$h(e^{ab,k})^T M f^{ab,k} = h(e^{ab,k})^T M J^k e^{ab,k} + h(e^{ab,k})^T M B^k u_\partial^k. \tag{5.123}$$

We seek to recover a discrete-time equivalent of the energy balance given by (5.101). In the quest to achieve this, as discussed in [96], the first term in the right-hand side of (5.123) must vanish for all $e^{ab,k} \in \mathbb{R}^{sn}$, i.e.,

$$h(e^{ab,k})^T M J^k e^{ab,k} \overset{!}{=} 0. \tag{5.124}$$

The conditions under which (5.124) holds have been discussed in [96], and are recalled below:
(i) $m_{ij} = 0$ for all $i \neq j$,
(ii) $J_i^k = J_j^k = const.$ for all $i, j = 1, 2, \ldots, s$.
We recall that the interconnection matrices obtained in the scope of the two-phase flow models are *state-dependent*, i.e., non-constant (see Section 5.3) and, hence, condition *(ii)* is violated. As a consequence, condition *(i)* need to necessarily hold in order to make the first term in the right-hand side of (5.123) vanish for all $e^{ab,k} \in \mathbb{R}^{sn}$. Condition (i) is an orthogonality requirement on the choice of the approximation basis in the discretization method. In other words, the interpolation polynomials must form a system of orthogonal functions to have $m_{ij} = m_{ji} = 0$, for $i \neq j$, with m_{ij} defined as in (5.120). This holds if the collocation points $c_1, c_2, \ldots, c_s$ are taken as the zeros of the shifted Legendre polynomials.
Continuing with the pursuit of discrete-time energy balance, multiplying (5.116) by yet to be determined $P^T \in \mathbb{R}^{sm \times sm}$, we have

$$P^T y_\partial^k = -P^T (B^k)^T e^{ab,k} + P^T D^k u_\partial^k. \tag{5.125}$$

Applying the transpose operation on both the left and the right-hand side of (5.125), and imposing $P^T (B^k)^T = (B^k)^T M^T$, we have:

$$\Big(P^T y_\partial^k\Big)^T = -(e^{ab,k})^T \Big(P^T (B^k)^T\Big)^T + \Big(P^T D^k u_\partial^k\Big)^T \tag{5.126a}$$

$$= -(e^{ab,k})^T M B^k + \Big(P^T D^k u_\partial^k\Big)^T. \tag{5.126b}$$

The imposition of $P^T(B^k)^T = (B^k)^T M^T$, a constraint used in the above computation, helps to determine P since B^k and M are already known.

Now using (5.126b), we have

$$(e^{ab,k})^T M B^k = -\left(P^T y_\partial^k\right)^T + \left(P^T D^k u_\partial^k\right)^T. \tag{5.127}$$

Plugging (5.127) into (5.123), we obtain:

$$h(e^{ab,k})^T M f^{ab,k} - h\left[-\left(P^T y_\partial^k\right)^T + \left(P^T D^k u_\partial^k\right)^T\right] u_\partial^k = 0. \tag{5.128}$$

Simplifying the above expression, we arrive at

$$h(e^{ab,k})^T M f^{ab,k} - h\left[-(y_\partial^k)^T P + (u_\partial^k)^T (D^k)^T P\right] u_\partial^k = 0. \tag{5.129}$$

Enforcing $(D^k)^T P$ to be skew-symmetric, the expression in (5.129) reduces to

$$h(e^{ab,k})^T M f^{ab,k} + h(y_\partial^k)^T P u_\partial^k = 0, \tag{5.130}$$

which represents the balance equation at the discrete-time level in terms of the discrete-time conjugate port variables at the nodes of the sampling interval I^k. At this stage, we have all the ingredients to define a discrete-time Dirac structure. This structure is formalized below.

Theorem 5.5.1 *Consider the system of equations*

$$f^{ab,k} = J^k e^{ab,k} + B^k u_\partial^k, \tag{5.131a}$$

$$y_\partial^k = -(B^k)^T e^{ab,k} + D^k u_\partial^k, \tag{5.131b}$$

with discrete flow, effort, input and output vectors $f^{ab,k}, e^{ab,k}, u_\partial^k, y_\partial^k$ according to (5.111a), (5.111b), (5.111c) and (5.111d), respectively, and the block-diagonal matrices J^k, B^k and D^k according to (5.114), (5.115) and (5.117), respectively. Moreover, consider the time interval $I^k = [(k-1)h, kh)]$ along with the s collocation points $0 \leq c_i \leq 1$, $i = 1, 2, \ldots, s$ Then, the corresponding discrete-time Dirac structure, which approximates the continuous-time Dirac structure, is given by:

$$\begin{bmatrix} \hat{f}^{ab,k} \\ \hat{y}_\partial^k \end{bmatrix} + \begin{bmatrix} -MJ^k & -MB^k \\ (B^k)^T M^T & -P^T D^k \end{bmatrix} \begin{bmatrix} e^{ab,k} \\ u_\partial^k \end{bmatrix} = 0, \tag{5.132}$$

if the following conditions hold:

(C1) $m_{ij} = 0$ *for all* $i \neq j$,

(C2) $P^T (B^k)^T = (B^k)^T M^T$,

(C3) $(D^k)^T P \in \mathbb{R}^{sm \times sm}$ *is skew-symmetric,*

where $M = M^T$ according to (5.120), and $\hat{f}^{ab,k}$ and $\hat{y}^k_{\partial}$ are defined as $\hat{f}^{ab,k} :=$ $M f^{ab,k}$ and $\hat{y}^k_{\partial} := P^T y^k_{\partial}$, respectively.

Proof: Multiplying (5.131a) by M and (5.131b) by M^T, and using the definitions of $\hat{f}^{ab,k}$ and $\hat{y}^k_{\partial}$ we have

$$\hat{f}^{ab,k} = M J^k e^{ab,k} + M B^k u^k_{\partial}, \tag{5.133a}$$

$$\hat{y}^k_{\partial} = -P^T (B^k)^T e^{ab,k} + P^T D^k u^k_{\partial}. \tag{5.133b}$$

The above system (5.133) can be equivalently written as

$$\underbrace{\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}}_{F} \begin{bmatrix} \hat{f}^{ab,k} \\ \hat{y}^k_{\partial} \end{bmatrix} + \underbrace{\begin{bmatrix} -M J^k & -M B^k \\ (B^k)^T M^T & -P^T D^k \end{bmatrix}}_{E} \begin{bmatrix} e^{ab,k} \\ u^k_{\partial} \end{bmatrix} = 0, \tag{5.134}$$

which is equivalent to (5.132). The system (5.134) is the kernel representation of a finite-dimensional Dirac structure if the matrices F and E in (5.134) satisfy: *(i)* $E F^T + F E^T = 0$, and *(ii)* rank$[F \mid E] = sn + sm$.

Due to the nature of F in (5.134), it is only required to check whether E in (5.134) is skew-symmetric. It is easy to see that the matrix E is skew-symmetric when conditions (C1) - (C3) are satisfied simultaneously. The foundation of conditions (C1) - (C3) is clearly evident from the theoretical development provided in Section 5.5.2 before stating this theorem.

The balance equation on the discrete-time Dirac structure is given by:

$$(e^{ab,k})^T \hat{f}^{ab,k} + (\hat{y}^k_{\partial})^T u^k_{\partial} = 0. \tag{5.135}$$

Substituting the definitions of $\hat{f}^{ab,k}$ and $\hat{y}^k_{\partial}$, and multiplying by h, we obtain (5.130), which is a quadrature formula for the energy balance. ∎

The collocation method, under the conditions of Theorem 5.5.1, yields the approximation of energy balance through the port $(u^k_{\partial}, \hat{y}^k_{\partial})$:

$$\Delta \tilde{H}^k = h (e^{ab,k})^T M f^{ab,k} = h (y^k_{\partial})^T M u^k_{\partial} = h (\hat{y}^k_{\partial})^T u^k_{\partial}. \tag{5.136}$$

The change of stored energy based on the integration of the numerical solution is expressed as:

$$\Delta \bar{H}^k = H(q_{s+1}^{ab,k}) - H(q_0^{ab,k}). \tag{5.137}$$

Under an exact discrete energy balance, $\Delta \tilde{H}^k = \Delta \bar{H}^k$ holds. However, numerically this exactness does not always hold. Based upon the definition provided in [96], if under a given numerical integration scheme of order p, the increment of stored energy satisfies

$$\Delta \bar{H}^k = h(\hat{y}_\partial^k)^T u_\partial^k + O(h^p), \tag{5.138}$$

we will call (5.138) a discrete energy balance, which is consistent with the discretization scheme.

5.6   Numerical Experiments

This section deals with the preliminary results in the scope of the numerical implementation of the proposed structure-preserving discretization framework. In particular, we test the performance of the mFEM-based structure-preserving spatial discretization scheme and employ the (non structure-preserving) explicit first-order Forward Euler method for the temporal discretization. We pursue the aforementioned implementation for the following reasons. Firstly, such a test would help to check whether the spatial discretization methodology is implemented well and works as per the developed theory. Secondly, the explicit Forward Euler (and Runge Kutta) methods, although not structure-preserving, are not expected to produce diverging results.

Remark 5.6.1 *In principle, it would be desired to implement both the structure-preserving spatial and temporal discretization to validate the theory developed in Sections 5.3 - 5.5. However, we defer such an implementation to future works.*

We base our numerical considerations on the (challenging) conservative setting of the TFM, i.e., without any frictional effects. We assume that the spatial domain is of length $L = 100$. We consider a test case with following uniform initial conditions:

$$m_g = 0.2, \quad m_\ell = 800, \quad v_g = 0, \quad v_\ell = 10 \quad \forall z \in [0, 100]. \tag{5.139}$$

Under the aforementioned conditions, the system is expected to stay uniform $\forall t$ since there are no sources or any sharp gradient at any spatial location at

time $t = 0$ that may drive the emanation of wavefronts and, subsequently, the topological changes in the solution.

We solve the state-space port Hamiltonian representation (5.96) with the system matrices J_i, B_i, C_i, D_i for all $i = \{1, 2, \ldots, n\}$ of the form (5.58). Here, n stands for the number of elements that the spatial domain has been discretized into. We observe the numerical behavior of the states of the system, i.e., $m_g, m_\ell, m_g v_g$, after 1, 2 and finite number of time-steps with different choices of the number of elements and the temporal step size. We find out that the states remain uniform after 1 time step for any temporal step size and the choice of the number of discretized elements on the spatial domain. However, we observe numerical differences in the evolution of the states after 2 or more time steps under different choices of the temporal and the spatial step size. We also observe that the states do not remain uniform and may diverge after a finite number of time steps. Suppose we focus on the evolution of m_ℓ. In order to demonstrate the behaviour of the perturbation around the uniform state of m_ℓ,

Figure 5.1: Perturbation of the mass of the liquid phase per unit length, i.e., m_ℓ at $2 \times dt$ for different choices of the temporal step sizes. Here, the domain has been discretized into 200 elements.

we define a metric p_m in the following manner:

$$p_m(z, t) = \frac{m_\ell^{\text{numerical}}(z, t) - 800}{\text{dt}}, \tag{5.140}$$

where dt denotes the temporal step size, $m_\ell^{\text{numerical}}(z, t)$ stands for the numerically predicted value of m_ℓ at location z and time t. Figures 5.1 and 5.2, respectively, demonstrate the behaviour of p_m at $2 \times$ dt for different choices of the temporal step sizes under the consideration that the spatial domain has been discretized into 200 and 400 elements. We can observe a similar oscillatory pattern in both figures, and that these perturbations around the uniform state are of the frequency which is a function of the spatial step size. It is also visible that the amount of perturbations reduce for smaller temporal step sizes, and even reproduce the desired uniform state; for instance, see Figure 5.1 at dt $= 5e-6$. The maximum value of the perturbation metric p_m is seen to reduce by an order of magnitude for a reduction in the temporal step size by a factor of 10. The perturbations, that are seemingly bounded, are observed to diverge after the first few time steps. At this stage, the pattern of the oscillatory response is not
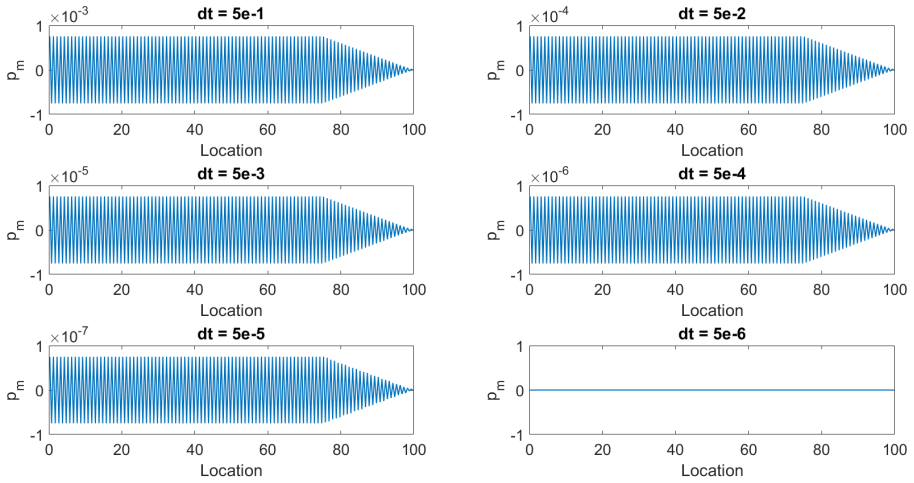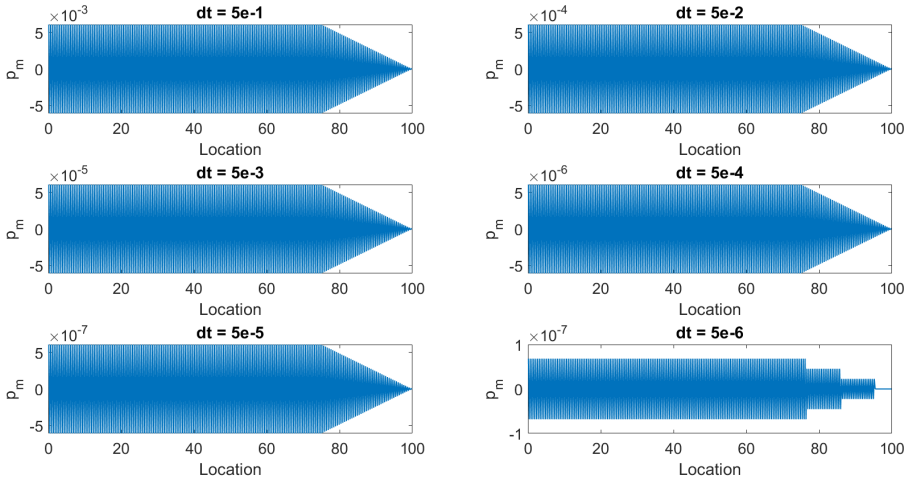

Figure 5.2: Perturbation of the mass of the liquid phase per unit length, i.e., m_ℓ at $2 \times$ dt for different choices of the temporal step sizes. Here, the domain has been discretized into 400 elements.

clear to us, and this aspect deserves further investigations. However, it is worth mentioning that such an oscillatory behaviour is certainly not due to physical reasons, but most likely due to numerical reasons, which include the choice of the non structure-preserving temporal discretization method.

Now, suppose that the spatial domain is discretized into 100 elements. Basically, we resort to a coarser mesh and perform a similar study as before. It is observed that the states remain uniform for any time horizon under any choice of the temporal step size. In other words, the perturbation metric p_m is found to be zero at each spatial location and time instant on a discretized space-time domain. We now seek to observe the behaviour of the Hamiltonian over time to validate the developed theory. Figure 5.3 demonstrates the behaviour of the Hamiltonian H_d over time under the choice of dt = 0.1. Here, the Hamiltonian H_d is obtained by summing up the discrete Hamiltonian of the form (5.61) over all the elements on the spatial domain. The observed behavior concurs well with the theory since we can see that H_d remains constant over time.

It is surprising that $p_m = 0$ for a coarser spatial mesh, and possesses oscil-
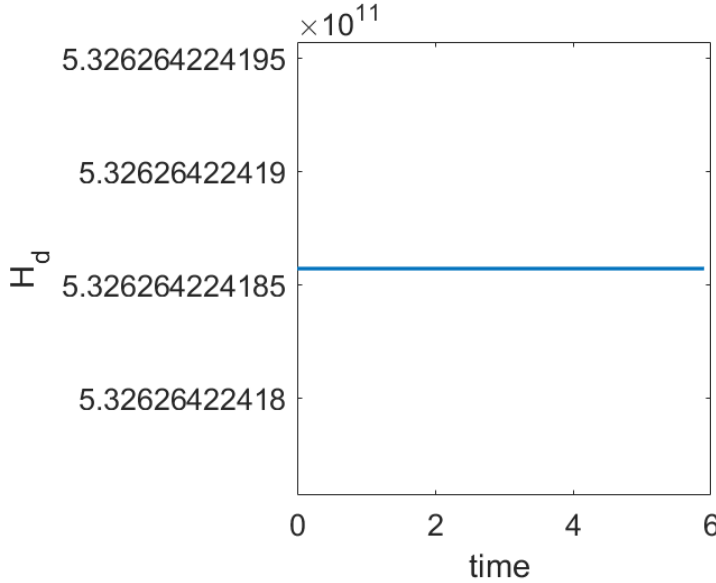
Figure 5.3: Behaviour of the Hamiltonian H_d over time. Here, 100 elements have been used and a time step size dt = 0.1 has been considered.

latory response for finer meshes. Given the observation that the evolution of m_ℓ is uniform under the choice of 100 elements even with the non structure-preserving temporal discretization technique, the reasons of deviation of m_ℓ (and other quantities) from the uniform state seem to be due to numerical error sources which possibly aggravate as we refine the mesh. This aspect, however, deserves further investigation.

Furthermore, given that the states remain uniform after the first time step under any (space and time) mesh resolution, we can (possibly) say that the structure-preserving spatial discretization methodology has been correctly implemented, and also works in consensus with the theory. However, further detailed investigation needs to be carried out in the future by assessing the proposed discretization framework on numerically challenging test cases.

5.7   Conclusions

We presented the spatial and temporal discretization of non-linear distributed-parameter port-Hamiltonian systems associated with state-dependent Stokes-Dirac structures. We performed spatial discretization on an infinite-dimensional port-Hamiltonian representation of the Two-Fluid Model and the Drift Flux Model using a mixed finite element method. We demonstrated that such a discretization preserves the (extended) continuous-time finite-dimensional Dirac structure even for the underlying (extended) *state-dependent* Stokes-Dirac structure. We obtained an aggregated finite-dimensional port-Hamiltonian model using the notion of power-preserving interconnection. Furthermore, we exploited the principles of symplectic integration methods to obtain a discrete-time Dirac structure. We derived the conditions under which the existence of a discrete-time Dirac structure can be guaranteed even in the presence of *state-dependent* system matrices and feedthrough terms.

Such a finite-dimensional approximation is amenable for control and observer design. Moreover, such finite-dimensional representations in conjunction with symmetry reduction could be a potential crucial ingredient towards obtaining a lowest-dimensional structure-preserving reduced-order model of the Drift Flux Model and the Two-Fluid Model.

The proposed spatial discretization methodology, i.e., mixed finite element method, can be adapted, though not straightforwardly, in order to consider higher-order finite-element spaces by using a similar spatial discretization procedure. Moreover, the proposed approach shares similarities with the recently developed mixed Galerkin discretization [97] applicable for multi-dimensional

models and, hence, it possesses potential for generalization to N–dimensional models. The presented approach can also be viewed from a domain decomposition point of view. The port-Hamiltonian representations are particularly attractive due to the property of compositionality. Hence, the basic challenge in the direction towards generalization to multi-dimensional models lies in obtaining an underlying finite-dimensional Dirac structure on individual finite-element of triangular or quadrilateral shape. This basic challenge can be dealt with modifications to the proposed setting. The applicability of the proposed approach for N–dimensional models will be considered extensively in future works.

We discussed the structure-preserving temporal discretization for an underlying ordinary differential equation based state-space port-Hamiltonian representation. Such an underlying representation is, for instance, obtained while considering lossless cases. It is however worth recalling that mixed finite element based spatial discretization yields a system of differential algebraic equations in the presence of resistive effects. In this chapter, we did not delve into developing a structure-preserving temporal discretization for a system of differential algebraic equations. However, structure-preserving temporal discretization can be obtained by following the principles laid down in [123].

We also numerically assessed the proposed discretization framework using a simple test case. Some numerical phenomena, which is not yet fully understandable, can be observed at finer mesh resolutions. Future research is required to understand such behaviour and also to numerically implement the proposed structure-preserving temporally-discretized scheme. Furthermore, the future study should encompass the validation of the proposed structure-preserving spatial and temporal discretization framework along with the verification of the numerical convergence.

In view of the still remaining open issues, we will not pursue the structure-preserving discretization approach in the rest of this dissertation. Instead, we resort to the Finite Volume Method for the high-fidelity numerical computations of the two-phase flow models of interest (see Chapter 6) and for building efficient reduced-order models (see Chapters 7-9).

Chapter 6

Numerical assessment of the two-phase flow model

Accurately capturing wave propagation phenomena is a central challenge in the hydraulic modelling of Managed Pressure Drilling (MPD) systems. In this chapter, the numerical analysis of two-phase flow models, particularly the one-dimensional Drift Flux Model (DFM), is pursued. We obtain a model for the speed of sound in the two-phase mixture by accounting for the compressibility effects of the phases under consideration and, subsequently, numerically study the importance of its role. Furthermore, to the best of our knowledge for the first time, we test the model and the numerical method with exact reference solutions. Moreover, we employ the non-linear error transport approach to numerically estimate the discretization error and, consequently, understand its generation and evolution mechanism. We also perform various challenging numerical case studies to test relevant numerical properties, such as Abgrall's principle, and also study the behaviour of the solutions under grid refinements. Finally, we use all the obtained knowledge to establish the relative order of merit of the introduced numerical schemes.

6.1 Introduction

Accurate multi-phase flow models are essential for wellbore simulations to support drilling operations [57, 130]. Namely, drilling scenarios may involve gas influx from the reservoir. In such situations, the flow inside the annulus section

is primarily of a two-phase nature and composed of gas-liquid mixture; also see Chapter 2. Often, the drilling fluids pumped inside the drill string are mixed with gas, which in effect makes the study of two-phase flow inside the drill string and annulus of relevance.

In this chapter, we are particularly interested in the numerical analysis of a one-dimensional two-phase flow model which is popularly known as the Drift Flux Model (DFM). Such a two-phase flow model is often used for wellbore simulations [57, 90, 130, 201]. We recall that multi-phase models, such as the DFM, pose several theoretical and computational challenges; see Chapter 1. Capturing the wave propagation phenomena, induced by slow or fast transients, is a central challenge in hydraulic modelling. Here, we are particularly interested to accurately resolve fast and slow evolving (discontinuous) wavefronts across the spatial domain (of the wellbore) while maintaining the positivity of a physical quantity.

Several approaches exist for solving hyperbolic or transport problems, such as the DFM. These include the Finite Volume Method [7, 14, 57–61, 67, 128, 167], Finite Element Method [98], Finite Difference Method [70], Discontinuous Galerkin Finite Element Method [192], etc. Among existing schemes, Finite Volume Methods are popularly used due to their versatility and good conservation properties. Hence, we use a Finite Volume Method to discretize the hyperbolic system model under consideration. In particular, conservative upwind methods are employed as they are known to accurately predict convection effects, which are responsible for the evolution of the discontinuities in the two-phase mixture.

It is already well-known that classical solutions to non-linear conservation laws are unique, but may not always exist due to discontinuous physical features. As a consequence, weak solutions are generally investigated and the numerical solution is expected to converge to the weak solution. Unlike the classical solutions, the weak solutions may be non-unique. Under grid refinement, an often suggested method for numerical validation, numerical solutions computed from various numerical schemes are theoretically expected to converge to the same solution. It is, however, observed that different numerical schemes yield different physical representations at a finer grid resolution [57]. Such an observation is surely related to the convergence behaviour of numerical schemes, which has to-date lacked proper attention in the existing literature in the scope of the DFM. Furthermore, numerical schemes developed for simulating the DFM have so far not been benchmarked against the exact solutions. Usually, a numerical scheme belonging to the family of Advection Upstream Splitting Methods, namely AUSMV, has been used as a reference to compare

the performance of any existing or newly developed numerical schemes. However, a sound physical reasoning behind considering AUSMV to be exhibiting the "true" behaviour is lacking. Moreover, AUSMV is not "Total Variation Diminishing" by nature; see [58] for the discussion in the scope of another variant of a two-phase flow model. The aforementioned reasons clearly point out the need to perform further numerical analysis to *(i)* test the model and the numerical method with exact reference solutions, if existing, and *(ii)* establish the relative order of merit of numerical schemes employed for the DFM.

In the pursuit of establishing the relative order of merit of the numerical schemes, the discretization error plays a fundamental role. The estimate of the discretization error can be obtained by employing residual and recovery methods [12], error transport methods inspired from defect correction techniques [18,38,143,164,169], etc. Residual and recovery methods have been primarily developed for linear elliptic and parabolic problems. And, error transport methods have been useful and are gaining popularity in the scope of estimating the discretization error of a hyperbolic model that even admits discontinuities [18, 143]. Due to the shortcomings of residual and recovery methods, we will resort to utilizing error transport methods. In the literature [18, 38, 143, 205], there exist linear and non-linear error transport approaches. The linear error transport approach has been shown to be inaccurate for problems with strong discontinuities [18]. Given the fact that the DFM admits such sharp gradient features, we do not pursue linear error transport approach. Instead, we resort to non-linear error transport approach [18, 143], that has met success in estimating the discretization error for problems with discontinuities. The foundation of non-linear error transport approach has been laid in [18]. In the past, the discretization error has been estimated for the Burgers' equation, wave equation, advection-diffusion equation, Euler equations in gas dynamics and Navier-Stokes equation [18, 38, 143, 205]. The effect of boundary condition implementation on the accuracy of the discretization error estimates has also been studied in the scope of Burgers' equation, Euler equations and Navier-Stokes equations [143]. However, the estimation of discretization error, with or without boundary effects, has not yet been pursued in the scope of the DFM. We will use the principle of non-linear error transport to numerically estimate the discretization error incurred upon employing different numerical schemes. This will serve as a stepping stone to obtain one of the metrics that will lead us to quantify the performance of the numerical scheme in the scope of the DFM.

In line with the objective of developing a robust, accurate and stable numerical scheme, it is important to mention that the speed of sound in the two-phase mixture plays an important role from a numerical perspective. Intermediate

steps in the construction of a finite-volume based numerical scheme include the computation of the numerical flux, which comprises computing the numerical convective flux and the numerical pressure flux [57]; also see Section 6.3. The numerical convective flux and the numerical pressure flux are both affected by the speed of sound in the two-phase mixture [57, 90]. Moreover, the numerical dissipation depends on the speed of sound in the two-phase mixture [57]. Furthermore, accurate acoustic velocity enables the correct determination of locations and speeds of the shock wave front, the rarefaction wave front and the contact discontinuity wave front. The speed of sound in the two-phase mixture also has a practical significance. It is linked to the propagation of pressure pulses, which is of great interest in the understanding of the transient operations in Managed Pressure Drilling (MPD). Namely, hydraulics models for MPD must be able to capture the pressure propagation irrespective of the flow rates of the drilling fluid and also incorporate water hammer effects. In general, the propagation of pressure pulses depends on the operating pressure and injection rates of the gaseous and the liquid phase. The propagation of the pressure pulses is proportional to the speed of sound in the two-phase mixture and is also affected by the fluid densities. Hence, the speed of sound in multi-phase medium dictates the pressure transient time (say, the time taken by the pressure pulse to travel from one end to the other end of the pipe). To summarize, the speed of sound in the two-phase mixture plays a very significant role, both numerically and practically.

Few surrogate models [4, 57] exist for calculating the speed of sound in gas-liquid mixtures. However, they are built upon the assumption of the incompressibility of the liquid phase. The major difference between compressible flows and incompressible flows is the propagation aspect of the pressure waves. Given the fact that the drilling fluids are highly compressible, the assumption of incompressiblity, to derive an approximate speed of sound in the two-phase mixture, is highly unrealistic for drilling applications. In addition, these surrogate models are singular[1] for very low and very high gas void fractions. Moreover, the model, generally used to compute an approximate speed of sound in the two phase mixtures (constituting higher gas void fractions), becomes singular before rendering the DFM to be non-hyperbolic. The existing models are also unreliable in terms of capturing the realistic effects at high operating pressures (of the order of 100 bars) or high operating temperatures, which occur in deep drilling scenarios. In view of all of the above factors, and the numerical and physical importance of the sound speed in the two-phase mixture, there is a

---

[1]See Section 6.2 for the discussion regarding the sense of singularity.

clear need for accounting compressiblity effects in the sound-speed model. In addition, it is important to numerically assess the behaviour of the solution in the presence of these compressibility effects.

The novel contributions of this chapter are as follows. Firstly, a model for the computation of the speed of sound in the two-phase mixture is proposed, and the importance of the role of sound speed of a two-phase mixture is studied numerically. The obtained numerical behaviour of the DFM is also compared against the one computed on the basis of surrogate models for the speed of sound in the two-phase mixture. Secondly, we test the model and numerical method with exact reference solutions, and thus lay a stronger basis to assess or benchmark the performance of a numerical scheme. Thirdly, we invoke the non-linear error transport approach to estimate the discretization error for several challenging numerical case studies. Finally, we exploit the knowledge of the discretization error estimate along with the other performance metrics, such as (formal and observed) order of accuracy, etc. to establish the relative order of merit of a numerical scheme.

The rest of the chapter is organized as follows. Conditional hyperbolicity and the model for the speed of sound in the two-phase mixture are then discussed in Section 6.2. The spatial discretization, numerical boundary conditions and the temporal discretization is detailed in Section 6.3. In Section 6.4, the error transport approach, is elaborated in the context of the DFM to estimate discretization error and eventually aid certification of the numerical schemes. A large assortment of numerical case studies are performed in Section 6.5 paving the road for establishing the relative order of merit of numerical schemes. Finally, Section 6.6 ends the chapter with conclusions and outlook.

6.2 Hyperbolicity and speed of sound

The model of interest, i.e., the DFM has already been extensively discussed in Section 2.3.1. The DFM is governed by (2.2) and (2.3) - (2.6). The model (2.2), under the assumption of constant cross-section, can be rewritten in the following general form:

$$\partial_t \mathbf{q} + \partial_x \mathbf{f}(\mathbf{q}, \mathbf{q_x}) = \mathbf{S}(\mathbf{q}, \mathbf{q_x}), \tag{6.1}$$

where

$$\mathbf{q} = \begin{pmatrix} \alpha_\ell \rho_\ell \\ \alpha_g \rho_g \\ \alpha_\ell \rho_\ell v_\ell + \alpha_g \rho_g v_g \end{pmatrix}, \quad \mathbf{f}(\mathbf{q}, \mathbf{q_x}) = \begin{pmatrix} \alpha_\ell \rho_\ell v_\ell \\ \alpha_g \rho_g v_g \\ \alpha_\ell \rho_\ell v_\ell^2 + \alpha_g \rho_g v_g^2 + P \end{pmatrix}, \quad \mathbf{S}(\mathbf{q}, \mathbf{q}_x) = \begin{pmatrix} 0 \\ 0 \\ Q_g + Q_v \end{pmatrix},$$

with $Q_g + Q_v = \left(g(m_g + m_\ell) \sin\theta + \frac{32\mu_m v_m}{d^2} \right)$; refer to Section 2.3.1 for the meaning of the physical variables.

Remark 6.2.1 *We will omit the argument $\mathbf{q}_x$ of the flux function $\mathbf{f}$ and the argument $\mathbf{q}_x$ of the source term $\mathbf{S}$ in the sequel. Such terms can be omitted as we are only considering flows across a constant geometrical cross-section. The omitted arguments will play a pivotal role for flows across variable cross-section [7] and will also influence the study of the hyperbolicity of the corresponding model.*

We now introduce the definition of the hyperbolic system to pave the way for related analysis in the context of the DFM.

Definition 6.1 *[105, 177] A system of the form* (6.1)*, with $\mathbf{f} \in \mathbb{R}^m, \mathbf{q} \in \mathbb{R}^m$, is said to be hyperbolic at a point (x,t) if $\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \in \mathbb{R}^{m \times m}$ has m real eigenvalues $\lambda_1, ..., \lambda_m$ and a corresponding set of m linearly independent right eigenvectors $\mathbf{K}^{(1)}, ..., \mathbf{K}^{(m)}$. The system is said to be strictly hyperbolic if the eigenvalues λ_i, $i = 1, 2, ..., m$, are all distinct.*

The eigensystem of the DFM is complicated and lacks a closed form representation [57, 201]. However, there exist surrogates of eigenvalues [2, 57]. Under the assumption of the incompressiblity of the liquid phase and $\alpha_g \rho_g \ll \alpha_\ell \rho_\ell$, the three surrogate eigenvalues for the DFM are given by $v_\ell - \omega_m$, v_g and $v_\ell + \omega_m$, where the sound speed of the mixture ω_m is given by [57]:

$$\omega_m = \begin{cases} c_\ell, & \alpha_g < \epsilon \\ c(P, \alpha_g, \rho_\ell, K), & \epsilon \leq \alpha_g \leq 1 - \epsilon, \\ c_g, & \alpha_g > 1 - \epsilon \end{cases} \tag{6.2}$$

with

$$c(P, \alpha_g, \rho_\ell, K) = \sqrt{\frac{P}{\alpha_g \rho_\ell (1 - K\alpha_g)}}. \tag{6.3}$$

Here, $\epsilon > 0$ is a small parameter whose purpose is to ensure the transition from a two-phase to a single-phase sound velocity. Two of the three eigenvalues $v_\ell - \omega_m$ and $v_\ell + \omega_m$ are linked to compressibility effects. The third eigenvalue v_g is coincident to the gaseous velocity. It can clearly be observed that the operating pressure P has a strong effect on the eigenvalues of the system. Also, the increase in pressure increases the absolute eigenvalues for a given gas-void fraction. In view of the eigenvalues $v_\ell - \omega_m$, v_g and $v_\ell + \omega_m$ and the expression of ω_m, the hyperbolicity of the model is ensured if $1 - K\alpha_g > 0$.

Another existing model for the sound speed of the mixture is obtained by using variable transformation in accordance with Gavryiluk and Fabre [2, 71], and under the assumption of the incompressibility of the liquid phase. The variable transformation is given by

$$\mathbf{u}^T = (\chi_\ell, \rho, v_g) := \left(\frac{(\alpha_\ell - \alpha_\ell{}^*)\rho_\ell}{\rho_{mix} - \alpha_\ell{}^* \rho_\ell}, \rho_{mix} - \alpha_\ell{}^* \rho_\ell, v_g \right), \tag{6.4}$$

where

$$\alpha_\ell{}^* = \frac{K-1}{K}, \quad \alpha_\ell{}^* \in [0,1). \tag{6.5}$$

The sound speed is given by

$$\omega_m(\mathbf{u}) = \sqrt{ \chi_\ell(\chi_\ell - 1)(v_g - v_\ell(\mathbf{u}))^2 + \frac{(1 - \alpha_\ell{}^*)c_g^2 \rho_g(\mathbf{u})}{\alpha_g(\mathbf{u})\rho} }. \tag{6.6}$$

From now on, we will refer the sound speed of the two-phase mixture in (6.2) and (6.6) as ω_m^1 and ω_m^2, respectively.

Having introduced the surrogate models for the sound speed of the two-phase mixture, we will now account for compressiblity effects of the liquid medium and shed insights on the resulting sound-speed model.

It is known that the formulations based on variables other than the conservative variables are bound to fail at shock-waves [177]. Despite this fact, we adopt the quasilinear formulation of the governing partial differential equations. We will not develop numerical schemes using the quasilinear formulation. We will only use the quasilinear formulation for analyzing the hyperbolicity of the model and eventually in obtaining sound speed model of the two-phase mixture. Such an analysis is supported by the point, as put forward in [165], that the computation of eigenvalues, Riemann invariants, and the genuine non linearity or linear degeneracy of the characteristic field are independent of the choice of the equations, i.e., they are invariant under coordinate transformations.

The quasilinear form of (6.1) in conservative variables is written as:

$$\partial_t \mathbf{q} + A(\mathbf{q})\partial_x \mathbf{q} = \mathbf{S}(\mathbf{q}), \tag{6.7}$$

where $\mathbf{q} = [\alpha_\ell \rho_\ell; \alpha_g \rho_g; \alpha_\ell \rho_\ell v_\ell + \alpha_g \rho_g v_g]^T$. The Jacobian matrix $A(\mathbf{q}) = \frac{\partial \mathbf{f}}{\partial \mathbf{q}}$ of the corresponding flux function $\mathbf{f}(\mathbf{q})$ cannot be computed in a straightforward manner only in terms of the conservative variables [57]. The Jacobian matrix

can however be easily obtained in terms of primitive variables. The quasilinear form of (6.1) in terms of the primitive (or non-conservative) variables can be written as:

$$\partial_t \mathbf{w} + B(\mathbf{w})\partial_x \mathbf{w} = s(\mathbf{w}), \tag{6.8}$$

where the actual set of primitive variables $\mathbf{w}$, $B(\mathbf{w})$, the relation between $B(\mathbf{w})$ and $A(\mathbf{q})$, and the relation between $s(\mathbf{w})$ and $\mathbf{S}(\mathbf{q})$ will be discussed in the sequel.

The primitive variables $\mathbf{w}$ can be obtained from conservative variables $\mathbf{q}$ and vice versa. Using the fact that $\mathbf{w}$ depends on $\mathbf{q}$, we rewrite (6.8) as

$$\frac{\partial \mathbf{w}}{\partial \mathbf{q}}\frac{\partial \mathbf{q}}{\partial t} + B(\mathbf{w})\frac{\partial \mathbf{w}}{\partial \mathbf{q}}\frac{\partial \mathbf{q}}{\partial x} = s(\mathbf{w}). \tag{6.9}$$

We now introduce $J = \partial_\mathbf{w}\mathbf{q}$ and $J^{-1} = \partial_\mathbf{q}\mathbf{w}$. Using this and (6.9), we have

$$J^{-1}\frac{\partial \mathbf{q}}{\partial t} + BJ^{-1}\frac{\partial \mathbf{q}}{\partial x} = s. \tag{6.10}$$

We have omitted the arguments in the above equation for the sake of simplicity. Relating (6.10) to (6.7), we have that

$$A(\mathbf{q}) = JBJ^{-1}, \quad S(\mathbf{q}) = Js. \tag{6.11}$$

If there are no sources, (6.9) can be written as follows:

$$C\partial_t \mathbf{w} + D\partial_x \mathbf{w} = 0. \tag{6.12}$$

Now, using (6.8) and (6.12), $B(\mathbf{w})$ can be defined as $B(\mathbf{w}) = C^{-1}D$.

The DFM, governed by (6.1) and (2.3) - (2.6), can be written in the form (6.12), as shown next. We consider that $\mathbf{w} := [\alpha_\mathrm{g},\, P,\, v_\ell]^T$. Using $\mathbf{q}$ and $\mathbf{w}$, we can obtain $J = \partial_\mathbf{w}\mathbf{q}$ as follows:

$$J = \begin{pmatrix} \frac{\partial q_1}{\partial w_1} & \frac{\partial q_1}{\partial w_2} & \frac{\partial q_1}{\partial w_3} \\ \frac{\partial q_2}{\partial w_1} & \frac{\partial q_2}{\partial w_2} & \frac{\partial q_2}{\partial w_3} \\ \frac{\partial q_3}{\partial w_1} & \frac{\partial q_3}{\partial w_2} & \frac{\partial q_3}{\partial w_3} \end{pmatrix} = \begin{pmatrix} -\rho_\ell & \frac{(1-\alpha_\mathrm{g})}{c_\ell^2} & 0 \\ \rho_\mathrm{g} & \frac{\alpha_\mathrm{g}}{c_\mathrm{g}^2} & 0 \\ (-\rho_\ell v_\ell + \rho_\mathrm{g} v_\mathrm{g}) & \frac{\alpha_\ell v_\ell}{c_\ell^2} + \frac{\alpha_\mathrm{g} v_\mathrm{g}}{c_\mathrm{g}^2} & \alpha_\ell \rho_\ell + \alpha_\mathrm{g} \rho_\mathrm{g}\frac{\partial v_\mathrm{g}}{\partial v_\ell} \end{pmatrix}.$$

The model (6.1) can be written in the form (6.12) with C and D given by:

$$C = \begin{pmatrix} -\rho_\ell & \frac{1-\alpha_\mathrm{g}}{c_\ell^2} & 0 \\ \rho_\mathrm{g} & \frac{\alpha_\mathrm{g}}{c_\mathrm{g}^2} & 0 \\ C_{31} & \frac{(1-\alpha_\mathrm{g})v_\ell}{c_\ell^2} + \frac{\alpha_\mathrm{g}\left(\frac{Kv_\ell(1-\alpha_\mathrm{g})+S}{1-K\alpha_\mathrm{g}}\right)}{c_\mathrm{g}^2} & (1-\alpha_\mathrm{g})\rho_\ell + \alpha_\mathrm{g}\rho_\mathrm{g}\left(\frac{K(1-\alpha_\mathrm{g})}{1-K\alpha_\mathrm{g}}\right) \end{pmatrix}, \tag{6.13}$$

$$D = \begin{pmatrix} -\rho_\ell v_\ell & \frac{(1-\alpha_g)v_\ell}{c_\ell^2} & \rho_\ell(1-\alpha_g) \\ D_{21} & \frac{\alpha_g\left(\frac{Kv_\ell(1-\alpha_g)+S}{1-K\alpha_g}\right)}{c_g^2} & \alpha_g\rho_g\left(\frac{K(1-\alpha_g)}{1-K\alpha_g}\right) \\ D_{31} & D_{32} & 2\rho_\ell(1-\alpha_g)v_\ell + 2\rho_g\alpha_g v_g\left(\frac{K(1-\alpha_g)}{1-K\alpha_g}\right) \end{pmatrix}, \qquad (6.14)$$

where

$$C_{31} = \rho_g\left(\frac{Kv_\ell(1-\alpha_g)+S}{1-K\alpha_g}\right) - \rho_\ell v_\ell + \alpha_g\rho_g\left(K\frac{\left(Kv_\ell(1-\alpha_g)+S\right)}{(1-K\alpha_g)^2} - \frac{Kv_\ell}{1-K\alpha_g}\right),$$

$$D_{21} = \rho_g\left(\frac{Kv_\ell(1-\alpha_g)+S}{1-K\alpha_g}\right) + \alpha_g\rho_g K\frac{\left(Kv_\ell(1-\alpha_g)+S\right)}{(1-K\alpha_g)^2} - \frac{\alpha_g\rho_g Kv_\ell}{1-K\alpha_g},$$

$$D_{31} = -v_\ell^2\rho_\ell + \left(\frac{Kv_\ell(1-\alpha_g)+S}{1-K\alpha_g}\right)^2\rho_g + 2\alpha_g\rho_g v_g\left(K\frac{\left(Kv_\ell(1-\alpha_g)+S\right)}{(1-K\alpha_g)^2} - \frac{Kv_\ell}{1-K\alpha_g}\right),$$

$$D_{32} = \left(1 + \frac{\left(\frac{Kv_\ell(1-\alpha_g)+S}{1-K\alpha_g}\right)^2\alpha_g}{c_g^2} + \frac{v_\ell^2(1-\alpha_g)}{c_\ell^2}\right).$$

Under the chosen set of primitive variables **w**, and using (6.13) and (6.14), we can obtain $B(\mathbf{w}) = C^{-1}D$. We can now analyze the eigenvalues of $B(\mathbf{w})$ to comment on the hyperbolic nature of the DFM and compare the obtained eigenvalues against the surrogate eigenvalues.

Remark 6.2.2 *It is correct here not to consider the effect of the source terms on the hyperbolicity of the model as the source terms are devoid of any derivative-type terms. The source terms will, however, play a crucial role in the hyperbolicity of the model for flows across a variable geometrical cross-section as the source terms are of the form $p\frac{dA}{dx}$. Still, the role of such terms was not considered in the hyperbolicity analysis for two-phase flow across a variable cross-section in [7], and only the role of the convective subset in dictating the hyperbolicity was taken into account.*

In addition to real eigenvalues, the matrix $B(\mathbf{w})$ is found to also admit complex eigenvalues. The complex eigenvalues imply that the solutions not only depend on the past state, but also on the future state. The existence of complex eigenvalues indicates the loss of hyperbolicity. It is observed that the DFM does not preserve hyperbolicity at high and low gas void fractions and, hence, global unconditional hyperbolicity cannot be ensured. Such a conditional hyperbolicity of the DFM is representative of the behaviour of most of the two-phase flow models, which are known not to maintain hyperbolicity across the whole parameter spectrum (here, gas void fraction). For instance, the Two-fluid Model [56] also

does not always yield real eigenvalues, and, hence, cannot be assumed hyperbolic for all flow conditions either. The existence of complex eigenvalues, in the scope of the DFM, hints towards the change of the type of the partial differential equation under consideration. The nature of the partial differential equation dictates the nature of the numerical method to be employed, and, hence, the study of hyperbolicity becomes even more important in view of accuracy sought in the numerical results. For instance, hyperbolic partial differential equations are generally solved using upwind finite volume methods.

Remark 6.2.3 *It is not clear if the hyperbolicity should be regarded as a necessary condition for well-posedness of the model. In other words, the physical validity of the model outside the hyperbolic regime remains an open question. Some works modify few terms in the model of interest to ensure the hyperbolicity of the model across all or vast range of operating conditions. For instance, it is known that interfacial forces influence the hyperbolicity of the two-fluid model [170]. In [51, 149, 157, 166, 170], the authors have proposed to add hyperbolic correction terms in order to retain the hyperbolicity of the two-fluid model and other variants of two-phase flow models. Such correction terms have not been assessed/studied yet in the scope of the DFM. And, we neither pursue it in the scope of this work. Furthermore, as in other research works pertaining to the DFM, we restrict the numerical experiments to the setting which abides by the hyperbolic nature of the model. However, it is worth mentioning that, in principle, the numerical implementation for (conditionally) hyperbolic DFM could encompass a switching behaviour to enforce suitable numerical scheme as per the type of the partial differential equation for a certain state of the system.*

Remark 6.2.4 *The hyperbolicity of the DFM will become even more restrictive for rather practical equations of state and complicated slip laws.*

We consider that the general form of the three eigenvalues of $B(\mathbf{w})$ is same as the one in [57], i.e., $\lambda_1 = v_\ell - \omega_m$, $\lambda_2 = v_g$, and $\lambda_3 = v_\ell + \omega_m$. The behaviour of ω_m is supposed to be different from (6.2) and (6.6) due to the inclusion of compressibility effects and no assumptions on relative velocity of the phases. For the no-slip case, we observe that one of the numerically computed eigenvalues is a mere perturbation on top of the flow velocity of the gaseous phase. This concurs with the surrogate eigenvalue v_g. Computing the other two eigenvalues and using the general form of λ_1 and λ_3, the modified sound speed ω_m^{new} for the two-phase mixture at $K = 1$ and $S = 0$, i.e., at no slip is given by

$$\omega_m^{new} = c_g c_\ell \sqrt{\frac{\rho_g \rho_\ell}{(\rho_\ell + \alpha_g \rho_g - \alpha_g \rho_\ell)(c_g^2 \rho_g - \alpha_g c_g^2 \rho_g + \alpha_g c_\ell^2 \rho_\ell)}}. \tag{6.15}$$
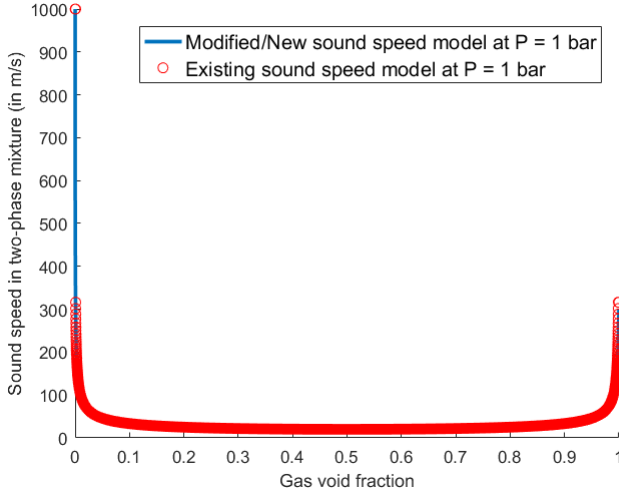
Figure 6.1: Comparison of sound speed models at P = 1 bar.

Remark 6.2.5 *Interestingly, the mixture sound speed obtained above for the DFM without slip is exactly equal to the mixture sound speed of the homogeneous two-phase mixture, which, contrary to the DFM, is governed by the mixture mass, momentum and energy conservation laws.*

The analytical expression of one of the earlier introduced surrogate model (see (6.2)) for the no-slip case yields:

$$\omega_m^1 = \begin{cases} c_\ell, & \alpha_g < \epsilon \\ \sqrt{\frac{P}{\alpha_g \rho_\ell (1-\alpha_g)}}, & \epsilon \le \alpha_g \le 1-\epsilon \\ c_g, & \alpha_g > 1-\epsilon \end{cases} \tag{6.16}$$

Figures 6.1 and 6.2 demonstrate the behaviour of the proposed and the existing model of the sound speed in the two-phase mixture versus gas void fractions for the no-slip case. While Figure 6.1 shows the comparison at $P = 1$ bar, Figure 6.2 shows the comparison across different values of the pressure. It can be clearly observed that the existing surrogate sound speed models admit non-smooth (discontinuous) transition in the regions of low and high gas

Figure 6.2: Zoomed in view for low gas void fractions and various pressures.

void fractions. Moreover, the surrogate models admit non-physical values of the speed of sound in the two-phase mixture. However, the obtained sound speed model under no-slip (6.15) is continuous and smooth throughout the whole spectrum of gas void fraction. Moreover, unlike (6.16), the 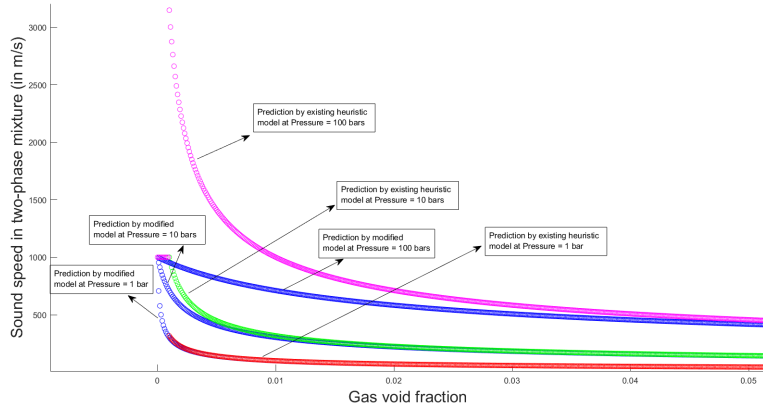model (6.15) only possesses physically consistent values of the speed of the sound in the two-phase mixture for varied phase compositions. A mixture sound speed model, as in (6.16) or (6.15), plays an important role in the velocity and the pressure splitting formulas for the computation of the numerical flux (see Section 6.3 for further details). Non-smooth transitions in the sound speed of the mixture, as in surrogate model (6.16), are a kind of hex that works well in practice for numerical simulations. However, non-smoothness (or a type of switching) in the sound speed model complicates the monotonicity analysis of a numerical scheme. Analysing the property of monotonicity of a numerical scheme is essential as numerical experiments are observed to show bounded oscillations [57]; also see Section 6.5. The presence of bounded oscillations is not attributed to the loss of stability. Instead, it is attributed to the loss of the property of monotonicity of a numerical scheme. The attempt towards having a smooth transition of the sound speed of the mixture, as observed for the modified model (6.15), will aid to study the property of monotonicity of a numerical scheme.

A closed form expression of the modified sound speed of the two-phase mixture could be obtained for the no-slip case. However, it is observed that the

quasilinear formulation, in general, does not aid in obtaining a closed expression of the eigenvalues and eigenvectors. A closed expression is not feasible for the DFM even for simplified equations of state of the phases under consideration. It should also be mentioned that the "true" hyperbolicity condition via the quasilinear approach is more restrictive than the hyperbolicity condition of the surrogate model (6.2). However, in view of the advantages of the modified model (6.15) discussed in the previous paragraph and that an accurate sound speed model is linked to a correct resolution of the moving wavefronts, we propose to use numerical eigenvalues (obtained using the quasilinear formulation) for the computation of the numerical flux instead of using surrogate eigenvalues. We propose to do so even though the closed expression of eigenvalues, eigenvectors and sound speed model is not feasible, and extra computational time will be required for the eigenvalue-computation at every time step. The differences incurred in the numerical results upon using the proposed and the existing surrogate model of the sound speed of the mixture are discussed in Section 6.5.

The identification of the type of the partial differential equation and the region of hyperbolicity is crucial for the reasons stated next. The numerical schemes should be developed such that they mimic the underlying hyperbolic nature. For instance, a finite speed of propagation in a certain direction should be well resolved by a numerical scheme. Furthermore, the numerical schemes tailored to solve the hyperbolic partial differential equations, should only be used within the hyperbolic regime[2]. Having discussed about the hyperbolicity of the model, we will next delve into spatial and temporal discretization methods along with the implementation of boundary conditions to solve the DFM numerically.

6.3   Numerical Methods

Given the fact that the one-dimensional DFM does not have any analytical solution for generic initial conditions, boundary conditions, equations of state and slip laws, it becomes important to solve the model numerically for concrete understanding of the underlying physical behaviour. While exact Riemann solvers have been developed for Euler equations [177], it is extremely challenging to construct exact Riemann solvers for the DFM. Hence, we need to heavily rely

---

[2]As mentioned in Remark 6.2.3, we will restrict the numerical experiments to the setting which abides by the hyperbolic nature of the model of interest.

on numerics. In the scope of our interest, the numerical schemes must be able to handle dynamic processes such as: *(i)* propagation of pressure pulses, and *(ii)* propagation of mass transport phenomena. The numerical schemes should be such that they allow the treatment of various linear or non-linear equations of state with slight modifications to the algorithmic framework. Furthermore, certified numerical schemes are sought in order to guarantee that the transient behaviour of the solution is not created due to numerical artifacts, and that the numerical solution is physically viable. Physical viability of the numerical scheme refers to obeying the non-linear conservation laws in discrete form, preserving the steady state solution at the discrete level and preserving the non-negativity of the physical variables.

Each numerical method should have the following characteristics [64]:

- **Consistency**: The truncation error should vanish under grid refinement, i.e., with the mesh spacing $\Delta x \to 0$ and/or $\Delta t \to 0$.

- **Stability**: The numerical errors should not be magnified when the iteration process advances. That is, for finite values of the time-step Δt and spatial-step Δx, the error (defined as the difference between the numerical solution and the exact solution of the numerical scheme) has to remain bounded, when the number of time steps, n, tends to infinity.

- **Convergence**: The numerical solution should approach the exact solution of the partial differential equation as the spatial and temporal step size tend to zero. Note that a consistent scheme is not necessarily convergent.

- **Conservation**: The conservation laws should be respected at the discrete level. The discrete conservation implies that any wave front computed by the conservative numerical scheme must be in the correct location, and the physically conserved quantities must be preserved/conserved at the discrete level.

In particular, we employ conservative numerical schemes in this work. It is well known that hyperbolic partial differential equations may develop discontinuities in finite time even for smooth initial data. Considering this basic feature, we do not consider non-conservative numerical schemes as such schemes do not converge to the correct solution if a shock wave is present in the flow, and the identified shock position may be wrong [177].

Numerical methods, popularly known by the name of method of lines, exist that first discretize in space and then in time [57–60, 67, 128, 167]. Also, there

exist numerical methods, popularly called Rothe's method, where the discretization in space is performed after the discretization in time [79, 100, 158]. Space-time discretization methods are also popular for numerically solving partial differential equations [72, 104, 190]. These aid to simultaneously discretize both space-time dimensions. In principle, any of the aforementioned methodologies can be used to discretize the DFM, which is a first-order (conditionally) hyperbolic partial differential equation. However, the method of lines approach is very popular to discretize the DFM [57, 90]. The lack of popularity or utilization of other numerical approaches to discretize the DFM is largely unknown. However, the lack of interest could be supported via following arguments. Space-time discretization methods are known to be computationally expensive and, hence, potentially not much popular in the scope of the discretization of the DFM. While Rothe's approach has been employed for hyperbolic partial differential equations, [79, 158], it has largely been used for discretizing second-order problems in this class. While delving into Rothe's method or space-time methods for numerically solving DFM could be the research topic in its own right, here we opt to pursue the popular discretization approach, i.e., the method of lines, in the scope of the DFM. Spatial discretization and temporal discretization, in the sense of method of lines, could be solved via explicit or implicit approaches. To be precise, the numerical flux approach and the time discretization approach could be explicit or implicit. On the one hand, explicit schemes are known to be efficient, but suffer from severe time-step restrictions in accordance with the limiting time-scale of the problem. On the other hand, implicit schemes are a bit more involved and complex from an implementation point of view. Without delving further into merits and demerits of each of these approaches, we use/-analyze the numerical methods that discretize in space (with explicit numerical flux approximation) followed by an explicit discretization in time.

The methodology for spatial and temporal discretization is critical in the construction of robust, accurate and efficient numerical methods. The discontinuities should be well captured and the numerical approach should not suffer from any spurious oscillations that may hamper the (accurate) approximation of the evolution response of the solution. Moreover, the numerical smearing should be minimal. It should be mentioned that most numerical methods are known to produce stable solutions under sufficient numerical damping, which is kind of an artificial damping that has no physical meaning. The stability of the numerical solutions alone will not be the criterion for rating the numerical scheme, which is one of the focal objectives of this work. It would be preferable to have a numerical scheme that is stable under zero/minimal numerical damping and without any other stabilization mechanisms.
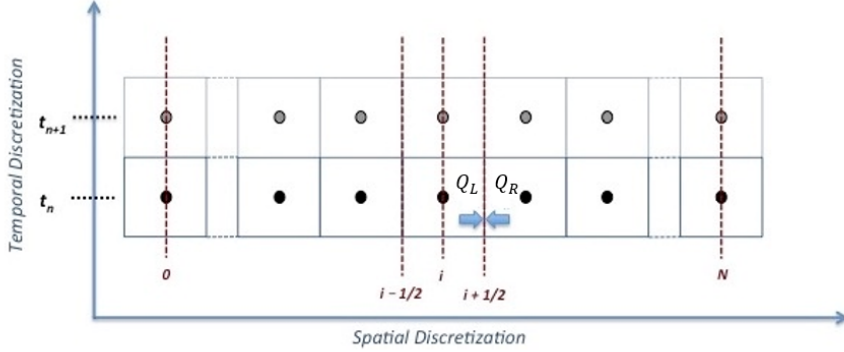
Figure 6.3: Stencil for numerical discretization.

The rest of this section is organized as follows. We first discuss the spatial discretization methodology extensively. First-order and higher-order spatial discretization is discussed in Section 6.3.1. Next, we discuss numerical treatment of boundary conditions in Section 6.3.2. The temporal discretization methodology is then discussed in Section 6.3.3.

6.3.1 Spatial Discretization

Spatial discretization constitutes the first step of establishing a computational method. The general form of a spatially discretized partial differential equation over the stencil represented in Figure 6.3 is given by

$$\partial_t \mathbf{q} = -\frac{1}{\Delta x}\left\{ F_{i+\frac{1}{2}}(Q_L, Q_R) - F_{i-\frac{1}{2}}(Q_L, Q_R) \right\} + S_i, \qquad (6.17)$$

where $\mathbf{q}$ is a state vector composed of conservative variables. Here, Q_L and Q_R are numerical value of the q-variables at left and right cell interface, respectively. Numerical flux $F^n_{i\pm\frac{1}{2}}(Q_L, Q_R)$ (at the right and the left interface of the finite volume cell i) can be estimated by various numerical schemes, and S_i represents source terms that can be approximated in various (consistent) ways.

In the finite-volume formulation, as mentioned earlier, the differences among

all numerical schemes essentially lie in the definition of the numerical flux evaluated at the cell interface. The different classes of numerical schemes are categorized as follows:

- **Upwind Schemes** [57, 90, 105]: Upwind schemes attempt to discretize hyperbolic PDEs by using differencing biased in the direction determined by the sign of the characteristic speeds. These schemes make use of the grid point where the information is coming from. Hence, upwinding is kind of directional bias while performing spatial discretization to respect the flow of the information determined by the characteristics. Thereby, such implementation incorporates aspects of the characteristic wave structure of the hyperbolic system under consideration, and is a natural choice for the spatial discretization of the system of interest.

- **Centered Schemes** [105]: These schemes do not rely on the specific eigenstructure of the problem and can be implemented in a straightforward manner. Such schemes, however, suffer from large numerical dissipation, and, hence, are incapable in sharp resolution of stationary or moving discontinuities.

- **Centered-Upwind Schemes** [59]: These schemes aim at exploiting the best properties of the upwind schemes and the centered schemes. The centered-upwind schemes utilize directional bias for discretization, and, in addition, retain the simplicity in construction and implementation.

The above mentioned class of spatial discretization schemes can be constructed to be (formally) first-order or higher-order accurate. The use of centered schemes is outrightly rejected due to its shortcomings. In this work, we will also not consider centered-upwind schemes as the main focus is not to enable simplicity in construction and implementation, but to address numerical issues (see Section 6.1) which are also observed in the class of upwind schemes. While centered-upwind schemes could also be considered in establishing the relative order of merit of a numerical scheme, we defer to do it in this chapter due to the aforementioned reasons, and to limit the number of schemes under investigation.

The first- and higher-order (upwind) approximations of numerical flux are discussed next.

First-order Spatial Discretization

First-order spatial discretization could be performed by employing various numerical schemes as listed below:

- Flux Difference Splitting (FDS) schemes [67, 177, 181]: Such schemes require the repeated computation of the eigenstructure, i.e., the Jacobian matrix of the system with respect to the conservative variables. Roe's Scheme and Osher's Scheme [177] belong to the class of FDS schemes. In the past, FDS schemes have been implemented for single-phase gas dynamics under ideal equation of state [177] and a variant of a two-phase flow model [181]. Several FDS based approximate Riemann solvers are used to solve hyperbolic-type partial differential equations, including the DFM [67]. However, such schemes are not much popular for discretizing the DFM due to the reliance on the eigenstructure of the system.

- Flux Vector Splitting (FVS) schemes [57, 59, 168]: Such schemes are simpler and in a way more efficient than the FDS schemes. FVS schemes rely on flux splitting based on the identification of the upwind direction(s). These schemes are, however, known to lead to numerical diffusion of contact discontinuity at rest. Van Leer type schemes are quite similar to the FVS schemes and, hence, are also classified within this category[3]. FVS and Van Leer type schemes are also commonly employed in the scope of the DFM, see [57].

- Hybrid schemes [57, 108, 109, 111, 198, 199]: The upwinding idea prevails throughout the construction of the AUSM scheme, which was initially developed by Liou and Steffen [111]. The AUSM scheme is not based on the Jacobian matrix computation and, hence, is independent of the eigenstructure. As a consequence, it is attractive for the DFM whose eigenstructure is not known explicitly. Instead of Jacobian computations, the AUSM scheme requires (numerical) flux to be split into separate components so that each one may be properly upwind stenciled. The convective and pressure fluxes are separated in the AUSM scheme. An upwind principle is used for the discretization of the convective terms. Unlike Roe's or Osher's scheme, it does not require redefining the intermediate states. It does not involve differentiation of (numerical) fluxes. The pressure component of the flux is split similar to the VanLeer concept. Unlike FVS schemes, the AUSM scheme allows capturing a stationary contact discontinuity with reduced numerical diffusion even with first-order discretization. In the past, several (improved) variants of the AUSM scheme, such as AUSM[+], AUSMV, AUSMD, AUSMDV, etc. have been developed [108,109,198,199].

---

[3]Minor technical differences between FVS and Van Leer type schemes will be pointed out later in this section while discussing about the computation of the numerical fluxes.

These variants have also been employed in the scope of the DFM, see [57].

We will not consider FDS methods in the sequel due to its shortcomings. Moreover, we will restrict our attention to only a few variants within the AUSM-family, namely, AUSM, AUSMV, AUSMD, and AUSMDV, which are popularly used for numerically solving the DFM. We will employ FVS, Van Leer scheme, and the aforementioned AUSM-family schemes in the scope of the DFM and, ultimately, study the pros and cons of each of these under different challenging test cases in Section 6.5.

Next, we discuss mathematical details of schemes belonging to the class of FVS, Van Leer and AUSM-family, and outline the corresponding numerical flux computations.

Flux Vector Splitting (FVS and Van Leer type)

It has been found that it is convenient to split the flux contributions into a part due to convective terms and a part due to pressure terms, and treat them separately. The (numerical convective and pressure) flux is split into respective contributions from each wave speed, with the positive flux component coming from positive wave speeds and the negative flux component coming from the negative wave speeds. As per (6.17), we need to compute the numerical flux at the interface $i + 1/2$ and at the interface $i - 1/2$. The numerical flux, in the sense of FVS, at the interface $i + 1/2$, i.e., $F_{i+1/2}^{FVS}$ is given by:

$$F_{i+1/2}^{FVS}(Q_L, Q_R) = \underbrace{\overbrace{(\alpha_\ell \rho_\ell)_L \Psi_{\ell,L}^+ + (\alpha_\ell \rho_\ell)_R \Psi_{\ell,R}^-}^{\text{Liquid}} + \overbrace{(\alpha_g \rho_g)_L \Psi_{g,L}^+ + (\alpha_g \rho_g)_R \Psi_{g,R}^-}^{\text{Gas}}}_{\text{Numerical Convective Flux}} +$$

$$\underbrace{(F_p)_{i+1/2}}_{\text{Numerical Pressure Flux}} , \quad (6.18)$$

where

$$\Psi_{\ell,L}^+ = \Psi_\ell^+(v_{\ell,L}, \omega_{i+1/2}), \quad \Psi_{\ell,R}^- = \Psi_\ell^-(v_{\ell,R}, \omega_{i+1/2}),$$
$$\Psi_{g,L}^+ = \Psi_g^+(v_{g,L}, \omega_{i+1/2}), \quad \Psi_{g,R}^- = \Psi_g^-(v_{g,R}, \omega_{i+1/2}), \quad (6.19)$$

and

$$(F_p)_{i+1/2} = \begin{pmatrix} 0 & 0 & p_{i+1/2} \end{pmatrix}^T, \quad (6.20)$$

with $p_{i+1/2} = P^+(v_\ell, \omega_{i+1/2})p_L + P^-(v_R, \omega_{i+1/2})p_R$, and $\Psi_\ell^+, \Psi_\ell^-, \Psi_g^+, \Psi_g^-$ as shown

below:

$$\Psi_\ell^+(v,\omega) = V^+(v,\omega)\begin{pmatrix}1\\0\\v\end{pmatrix}, \quad \Psi_\ell^-(v,\omega) = V^-(v,\omega)\begin{pmatrix}1\\0\\v\end{pmatrix},$$

$$\Psi_g^+(v,\omega) = V^+(v,\omega)\begin{pmatrix}0\\1\\v\end{pmatrix}, \quad \Psi_g^-(v,\omega) = V^-(v,\omega)\begin{pmatrix}0\\1\\v\end{pmatrix}. \quad (6.21)$$

The notation $v_{\ell,L}$ indicates the value of v_ℓ at the left of the interface under consideration. Other notations $v_{\ell,R}, v_{g,L}, v_{g,R}$ can be similarly inferred. And, the usage of $i+1/2$ in the subscript refers to the computation at the interface $i+1/2$. The variable ω stands for the speed of sound in the two-phase mixture. The velocity splitting function $V^\pm$ and pressure splitting function $P^\pm$ are given by:

$$V^\pm(v,\omega) = \begin{cases} \pm\frac{1}{4\omega}(v\pm\omega)^2 & \text{if } |v| \le \omega \\ \frac{1}{2}(v\pm|v|) & \text{otherwise} \end{cases}, \quad P^\pm(v,\omega) = V^\pm(v,\omega)\begin{cases} \frac{1}{\omega}(\pm2-\frac{v}{\omega}) & \text{if } |v| \le \omega \\ \frac{1}{v} & \text{otherwise} \end{cases}.$$
$$(6.22)$$

Here, $V^\pm$ and $P^\pm$ are the functions that satisfy the consistency, upwinding, monotonicity, differentiability and positivity property, see [58, 108]. The numerical flux, in the sense of FVS, at the interface $i-1/2$, i.e., $F_{i-1/2}^{FVS}$ can be analogously computed. We refer to [57] and the references therein for further technical details.

The FVS scheme coincides with the Van Leer scheme for the mass conservation equations. However, the discretization of mixed momentum equation is performed differently in these two approaches, as is clarified next while discussing about the numerical flux computation in the Van Leer approach. The numerical flux, again composed of the convective contribution due to the liquid and gaseous phase, and the pressure contribution, is given by the sum of the quantities $F_{i+1/2}^{\text{liquid, convective}}, F_{i+1/2}^{gas,convective}$ and $(F_p)_{i+1/2}$, where components $F_{i+1/2}^{\text{liquid, convective}}$ and $F_{i+1/2}^{gas,convective}$ are given by:

$$F_{i+1/2}^{\text{liquid, convective}} = \frac{1}{2}[(\alpha_\ell\rho_\ell v_\ell)_{i+1/2}(\Phi_{\ell,L}+\Phi_{\ell,R}) - |(\alpha_\ell\rho_\ell v_\ell)_{i+1/2}|(\Phi_{\ell,R}-\Phi_{\ell,L})],$$
$$(6.23)$$

$$F_{i+1/2}^{gas,convective} = \frac{1}{2}[(\alpha_g\rho_g v_g)_{i+1/2}(\Phi_{g,L}+\Phi_{g,R}) - |(\alpha_g\rho_g v_g)_{i+1/2}|(\Phi_{g,R}-\Phi_{g,L})],$$
$$(6.24)$$

and the numerical pressure flux $(F_p)_{i+1/2}$ is computed exactly in the manner as depicted for the FVS scheme. The quantities $\Phi_{\ell,L}, \Phi_{\ell,R}, \Phi_{g,L}, \Phi_{g,R}$ in (6.23) and (6.24) are given by:

$$\Phi_{\ell,L} = \begin{pmatrix} 1 & 0 & v_{\ell,L} \end{pmatrix}^T, \quad \Phi_{\ell,R} = \begin{pmatrix} 1 & 0 & v_{\ell,R} \end{pmatrix}^T,$$
$$\Phi_{g,L} = \begin{pmatrix} 0 & 1 & v_{g,L} \end{pmatrix}^T, \quad \Phi_{g,R} = \begin{pmatrix} 0 & 1 & v_{g,R} \end{pmatrix}^T. \quad (6.25)$$

And, the mass fluxes in (6.23) and (6.24), generically written as $\alpha \rho v$, are expressed as:

$$(\alpha \rho v)_{i+1/2}^{\text{Van Leer}} = V^+(v_L, \omega_{i+1/2}) \alpha_L \rho_L + V^-(v_R, \omega_{i+1/2}) \alpha_R \rho_R. \quad (6.26)$$

The velocity and pressure splitting functions are governed by the same definition as in (6.22). To summarize, the numerical flux, in the sense of Van Leer, at the interface $i+1/2$, i.e., $F_{i+1/2}^{\text{Van Leer}}$ is given by:

$$F_{i+1/2}^{\text{Van Leer}} = F_{i+1/2}^{\text{liquid, convective}} + F_{i+1/2}^{gas,convective} + (F_p)_{i+1/2}. \quad (6.27)$$

Analogously, $F_{i-1/2}^{\text{Van Leer}}$ can also be computed.

AUSM-Family

The schemes falling in this class also rest upon the principle to split the contributions due to the convective and pressure terms. While the numerical pressure flux is computed exactly as done in the FVS scheme, the numerical convective fluxes are treated differently. This is mainly owing to the treatment of mass fluxes, see (6.30). The numerical convective flux due to the liquid and gaseous contribution is shown below:

$$F_{i+1/2,AUSM}^{liquid,convective} = 0.5[(\alpha_\ell \rho_\ell v_\ell)_{i+1/2}(\Phi_{\ell,L} + \Phi_{\ell,R}) - |(\alpha_\ell \rho_\ell v_\ell)_{i+1/2}|(\Phi_{\ell,R} - \Phi_{\ell,L})], \quad (6.28)$$

$$F_{i+1/2,AUSM}^{gas,convective} = 0.5[(\alpha_g \rho_g v_g)_{i+1/2}(\Phi_{g,L} + \Phi_{g,R}) - |(\alpha_g \rho_g v_g)_{i+1/2}|(\Phi_{g,R} - \Phi_{g,L})], \quad (6.29)$$

where the quantities $\Phi_{\ell,L}, \Phi_{\ell,R}, \Phi_{g,L}, \Phi_{g,R}$ in (6.28) and (6.29) carry the same meaning as in (6.25). And, the mass fluxes (generically written as $\alpha \rho v$), appearing in (6.28) and (6.29) are given by:

$$(\alpha \rho v)_{i+1/2}^{AUSM} = 0.5[v_{i+1/2}(\alpha_L \rho_L + \alpha_R \rho_R) - |v_{i+1/2}|(\alpha_R \rho_R - \alpha_L \rho_L)] \quad (6.30)$$

where, $v_{i+1/2} = v_L^+ + v_R^-$ and $v_L^+ = V^+(v_L, \omega_{i+1/2})$ and $v_R^- = V^-(v_R, \omega_{i+1/2})$.

Next, we discuss about the variants of the AUSM scheme. AUSMV stands for Advection Upstream Splitting Method, and V at the end refers to the modified velocity splitting function, which is expressed as:

$$\tilde{V}^{\pm}(v, \omega, \chi) = \begin{cases} \chi V^{\pm}(v, \omega) + (1 - \chi)\frac{v \pm |v|}{2} & \text{if } |v| \leq \omega \\ \frac{1}{2}(v \pm |v|) & \text{otherwise} \end{cases}. \tag{6.31}$$

The only difference in the computation of the numerical flux via AUSMV lies in the treatment of mass flux type terms, as in (6.30). The mass fluxes, in the sense of AUSMV, are computed as:

$$(\alpha \rho v)_{i+1/2} = (\alpha \rho)_L \tilde{V}_L^+ + (\alpha \rho)_R \tilde{V}_R^-, \tag{6.32}$$

where, $\tilde{V}_L^+ = \tilde{V}^+(v_L, \omega_{i+1/2}, \chi_L)$ and $\tilde{V}_R^- = \tilde{V}^-(v_R, \omega_{i+1/2}, \chi_R)$. Many different choices can be made for the weighting functions χ_L and χ_R. However, $\chi_L = \alpha_R$ and $\chi_R = \alpha_L$ are popularly used [57].

AUSMD also replies on the modified definition of velocity splitting functions. AUSMD scheme can be viewed as a Van Leer scheme where modified velocity splitting functions replace the original velocity splitting functions. Similar to AUSMV, the mass fluxes, in the sense of AUSMD, are computed as:

$$(\alpha \rho v)_{i+1/2} = (\alpha \rho)_L \tilde{V}_L^+ + (\alpha \rho)_R \tilde{V}_R^- \tag{6.33}$$

Again, $\chi_L = \alpha_R$ and $\chi_R = \alpha_L$ are generally used [57].

In the above discussion we only delved into the aspects related to the computation of the numerical flux at the interface $i + 1/2$. The computations at the interface $i - 1/2$ can be carried out analogously. So far, we discussed about the first-order numerical scheme for spatial discretization. Next, we will discuss about the higher-order variants of the above introduced numerical schemes.

Higher-order Spatial Discretization

The correct description of fluid transport and pressure waves requires high-resolution schemes possessing little to zero numerical diffusion. It is well known that smearing dissipation plagues first-order methods [177]. In the past, a lot of work has been put in towards the development of the theory of high resolution

schemes for the numerical solution of the governing partial differential equations; see Chapters 13 and 14 of [177] and Chapter 9 of [175] for an overview. As a consequence of Godunov's Theorem (see Theorem 13.51 in [177]), we need to construct non-linear methods to obtain numerical solutions with the high-order of accuracy and with no spurious oscillations. In the pursuit of looking for such non-linear numerical schemes, the concept of limiters plays a key role. Limiters add the numerical non-linearity into the fully or semi-discrete system. A suitable limiter helps to design high resolution and non-oscillatory schemes by controlling the process of generation of over- and undershoots by preventing gradients to exceed certain limits. There exist several limiters, for example, Vanleer limiter, Min-Mod limiter, Osher limiter, Superbee limiter, van Albada limiter, Sweby limiter, generalized Min-Mod, monotonized central limiter, etc [171, 175, 177]. Each of these limiters have different switching characteristics. These limiters could be directly applied to the numerical fluxes or to the conservative/primitive variables. The former methodology is referred to as the flux-limiting, while the latter is referred to as the slope-limiting in the literature [171, 175, 177]. The slope-limiting can be implemented in two ways. One of the methodologies is to apply the limiting on the conservative variables. The other methodology includes the application of the limiter on the primitive variables. In certain situations, the flux-limiter can be shown to be exactly equivalent to the slope-limiter. While there are no significant differences between the two limiter classes, we will utilize the slope-limiting concept to perform higher-order spatial discretization of the DFM in this chapter. The slope-limiting methodology in the scope of the DFM is briefly illustrated next.

Consider the following one-dimensional semi-discrete system:

$$\partial_t \mathbf{q} = -\frac{1}{\Delta x}\left\{F_{i+\frac{1}{2}}(Q_L^\star, Q_R^\star) - F_{i-\frac{1}{2}}(Q_L^\star, Q_R^\star)\right\} + S_i. \qquad (6.34)$$

The above equation differs from (6.17) in the manner of computation of the $\mathbf{q}$–variables at the left and the right-hand side of the interface. Unlike first-order approximation, a piecewise constant value of the state variables is no longer used. A higher-order variant of various schemes for the DFM is obtained by using an approach similar to the classical Monotonic Upstream Scalar Conservation Laws (MUSCL) technique [177, 191]. In other words, the primitive or the conserved variables are extrapolated to get the values of the variables at the cell faces. The values of $Q_L^\star$ and $Q_R^\star$ at the interface $i + 1/2$ can be computed as shown below:

$$Q_L^\star = \mathbf{q}_i + \frac{1}{2}\phi(r_i)(\mathbf{q}_{i+1} - \mathbf{q}_i), \quad Q_R^\star = \mathbf{q}_{i+1} - \frac{1}{2}\phi(r_{i+1})(\mathbf{q}_{i+2} - \mathbf{q}_{i+1}), \qquad (6.35)$$

where $\phi(\cdot)$, a limiter function, is a measure of the smoothness of the solution profile, the argument r_i to the limiter function is given by:

$$r_i = \frac{u_i - u_{i-1}}{u_{i+1} - u_i},\tag{6.36}$$

and the subscripts to the **q**– variables represent the finite-volume cell at which the computation must be performed. The values of $Q_L^\star$ and $Q_R^\star$ at the interface $i-1/2$ can be analogously computed. The limiting, as in (6.35), has been applied on the conservative variables. A similar procedure could be used to limit the primitive variables, which can easily be obtained from the knowledge of the conservative variables.

The limiters should satisfy the following criteria in order to be "total variation diminishing" (TVD) by nature:

- $r \le \phi(r) \le 2r; (0 \le r \le 1)$;

- $1 \le \phi(r) \le r; (1 \le r \le 2)$;

- $1 \le \phi(r) \le 2; (r > 2)$;

- $\phi(1) = 1$.

The class of limiters which satisfy the above TVD constraints include Van Albada 1, VanLeer, Min-Mod, Osher, Monotonized Central, Superbee, Sweby, and Generalized Min-Mod limiter, and these will be exploited later in Section 6.5.

Numerical schemes, under consideration of different limiters, will show obvious differences. In Section 6.5, we will study the effects of different limiters on the performance of the numerical scheme. The observations will consequently aid in establishing the relative order of merit of the numerical schemes.

6.3.2   Numerical Implementation of Boundary Conditions

The numerical treatment of the boundary conditions is an essential aspect for finding the numerical solution of the model of interest. The numerical boundary conditions must be formulated and discretized appropriately so as to be compatible with the order of accuracy and stability conditions of the numerical scheme at internal domain cells. The number of positive eigenvalues of the Jacobian matrix, i.e., $\frac{\partial \mathbf{f}}{\partial \mathbf{q}}$, is equal to the number of physical boundary conditions on the inlet side of the spatial domain of interest. And, the number of negative eigenvalues of the Jacobian matrix corresponds to the number of physical boundary

conditions on the outlet side of the spatial domain. The numerical boundary conditions serve as the compatibility equations for the outgoing characteristics, which need to be added to the imposed physical boundary conditions in order to obtain the missing equations for the cells at the leftmost and the rightmost boundary of the spatial domain. See the Appendix in [90] for the characteristic treatment of the boundary conditions. It is an appealing method for specifying boundary conditions for hyperbolic partial differential equations as it uses relations based on the characteristic lines, i.e., on the analysis of the different waves crossing the boundary. This method does not use any extrapolation procedure, thereby it suppresses the arbitrariness in the imposition of the numerical boundary conditions.

In the scope of the drilling operations, the flow is usually in the sub-sonic regime. As a consequence of the eigenstructure of the DFM in such sub-sonic regime, two physical boundary conditions have to be specified at the inlet and one physical boundary condition at the outlet. The gas and liquid flowrates at the pipe inlet are generally known. Similarly, the pressure is known or prescribed at the outlet of the spatial domain; also see Section 2.3.1.

Remark 6.3.1 *An imposition of physical boundary conditions, as discussed above, will only hold if the eigenvalues of the Jacobian matrix do not flip during the evolution of the states over time. It is, hence, important to check whether the eigenvalues of the Jacobian matrix flip during numerical simulations. We will carry out this sanity check for all the numerical experiments discussed in Section 6.5.*

Extrapolation of say, pressure, using pressure values at interior points might be adopted for the implementation of the boundary conditions. However, the compatibility of such extrapolation with the original set of physical boundary conditions becomes unclear. This is due to the fact that the extrapolation acts as an additional physical condition imposing zero or pre-conceived pressure gradient and, hence, (could) overspecify the boundary conditions. Various ways to enforce extrapolation-based boundary treatments are briefly mentioned below. Considering the setting where m indicates the cell at the boundary, and n indicates the time sample, any physical quantity X can be extrapolated as follows:

Space Extrapolation, Zero-Order Extrapolation:

$$X_m^{n+1} = X_{m-1}^{n+1}. \tag{6.37}$$

Space Extrapolation, First-Order Extrapolation:

$$X_m^{n+1} = 2X_{m-1}^{n+1} - X_{m-2}^{n+1}. \tag{6.38}$$

Space-Time Extrapolation, First-Order in space/Zero-order in time:

$$X_m^{n+1} = 2X_{m-1}^n - X_{m-2}^n. \tag{6.39}$$

Time Extrapolation, First-Order:

$$X_m^{n+1} = 2X_m^n - X_m^{n-1}. \tag{6.40}$$

The choice among the above mentioned extrapolation approaches is a bit arbitrary. We can analyze the sensitivity of numerical schemes under different kinds of extrapolation-based boundary treatment. Moreover, we could compare the behaviour of the numerical schemes for characteristic-based boundary treatment and for the extrapolation-based boundary treatment. Both characteristic-based implementation of boundary conditions and extrapolation-based implementation of boundary conditions have been investigated while performing numerical simulations. We defer further discussion to Section 6.5.

6.3.3   Temporal Discretization

The spatial discretization of the partial differential equation results in a vector-valued semi-discrete system as in (6.17). It is easy to observe that the semi-discrete system obtained with or without implementing flux/slope limiters is highly non-linear. Now consider $L(\mathbf{q}) = -\frac{1}{\Delta x}(F_{i+1/2} - F_{i-1/2}) + S_i$. The semi-discrete system, equivalent to (6.17), can now be written as:

$$\frac{d\mathbf{q}}{dt} = L(\mathbf{q}). \tag{6.41}$$

The temporal accuracy of a numerical scheme is as important as the spatial accuracy. Hence, the choice of the method for the time discretization is also important. Moreover, the choice also affects the stability properties, which is one of the central issues while considering any numerical discretization technique. Furthermore, the solutions should remain non-negative after temporal discretization. As mentioned earlier, we restrict ourselves to the consideration of explicit time discretization schemes.

Remark 6.3.2 *The spatial discretization yields a semi-discrete system which is constituted by simultaneous ordinary differential equations (ODEs). If these ODEs have widely separated eigenvalues, then very small time-steps will be required while using explicit methods to obtain stable numerical solutions.*

In the scope of the DFM [57, 90], the numerical solution of (6.41) is usually advanced by employing a first-order Forward Euler scheme as shown below:

$$\mathbf{q}^{n+1} = \mathbf{q}^n + \Delta t L(\mathbf{q}^n), \tag{6.42}$$

where the superscripts n and $n+1$ are representative of the time-instants.

Remark 6.3.3 *For hyperbolic conservation laws, such as the DFM, the spectrum of the upwind spatial differential operator constitutes eigenvalues that lie in the left-half complex plane near the imaginary axis. The absolute stability region of the first-order Forward Euler method intersects the imaginary axis only at the origin. As a result, the first-order Forward Euler method is typically not a stable choice for temporal discretization, particularly when the (discrete) system contains pure imaginary eigenvalues which is representative of hyperbolic partial differential equations. Furthermore, it is only first-order accurate. If the system eigenvalues are on the imaginary axis, the temporal discretization should be such that the part of the imaginary axis is contained in the stability region. If the discrete system has eigenvalues near the imaginary axis, we would be much better off using one of the higher-order Runge Kutta methods.*

In view of Remark 6.3.3, in addition to the first-order Forward Euler, we explore family of (TVD) Runge Kutta methods for temporal discretization. The second, third-order and fourth-order (TVD) Runge Kutta scheme are recalled below.

Runge Kutta 2b Time Integration- The evolution of the states is as follows:

$$\mathbf{q}^{(0)} = \mathbf{q}^n, \tag{6.43a}$$

$$\mathbf{q}^{(1)} = \mathbf{q}^{(0)} + 1\Delta t L(\mathbf{q}^{(0)}), \tag{6.43b}$$

$$\mathbf{q}^{(2)} = \mathbf{q}^{(0)} + \frac{1}{2}\Delta t L(\mathbf{q}^{(1)}) + \frac{1}{2}\Delta t L(\mathbf{q}^{(0)}), \tag{6.43c}$$

$$\mathbf{q}^{n+1} = \mathbf{q}^{(2)}, \tag{6.43d}$$

where the bracketed supercripts denote the state at the intermediate stages, and $\mathbf{q}^n, \mathbf{q}^{n+1}$ indicate the value of $\mathbf{q}$ at time-instant t^n and t^{n+1}, respectively.

Runge Kutta 3b Time Integration- The evolution of the states is as follows:

$$\mathbf{q}^{(0)} = \mathbf{q}^n, \tag{6.44a}$$

$$\mathbf{q}^{(1)} = \mathbf{q}^{(0)} + 1\Delta t L(\mathbf{q}^{(0)}), \tag{6.44b}$$

$$\mathbf{q}^{(2)} = \mathbf{q}^{(0)} + \frac{1}{4}\Delta t L(\mathbf{q}^{(1)}) + \frac{1}{4}\Delta t L(\mathbf{q}^{(0)}), \tag{6.44c}$$

$$\mathbf{q}^{(3)} = \mathbf{q}^{(0)} + \frac{1}{6}\Delta t L(\mathbf{q}^{(1)}) + \frac{1}{6}\Delta t L(\mathbf{q}^{(0)}) + \frac{2}{3}\Delta t L(\mathbf{q}^{(2)}), \tag{6.44d}$$

$$\mathbf{q}^{n+1} = \mathbf{q}^{(3)}. \tag{6.44e}$$

Runge Kutta 4b Time Integration- The evolution of the states is as follows:

$$\mathbf{q}^{(0)} = \mathbf{q}^n, \tag{6.45a}$$

$$\mathbf{q}^{(1)} = \mathbf{q}^{(0)} + \frac{1}{2}\Delta t L(\mathbf{q}^{(0)}), \tag{6.45b}$$

$$\mathbf{q}^{(2)} = \mathbf{q}^{(0)} + \frac{1}{2}\Delta t L(\mathbf{q}^{(1)}), \tag{6.45c}$$

$$\mathbf{q}^{(3)} = \mathbf{q}^{(0)} + 1\Delta t L(\mathbf{q}^{(2)}), \tag{6.45d}$$

$$\mathbf{q}^{(4)} = \mathbf{q}^{(0)} + \frac{1}{6}\Delta t L(\mathbf{q}^{(0)}) + \frac{1}{3}\Delta t L(\mathbf{q}^{(1)}) + \frac{1}{3}\Delta t L(\mathbf{q}^{(2)}) + \frac{1}{6}\Delta t L(\mathbf{q}^{(3)}), \tag{6.45e}$$

$$\mathbf{q}^{n+1} = \mathbf{q}^{(4)}. \tag{6.45f}$$

Many other variants of explicit time discretization techniques exist, but we will limit the discussion around the above mentioned ones.

Remark 6.3.4 *The obtained semi-discrete system, governed by set of ODEs, is stiff due to distinct time scales in the problem. We do not investigate this aspect in more detail as there is no consensus on the definition of the stiffness of the ODEs in the non-linear setting. It should also be mentioned that explicit methods are often unacceptable for stiff problems due to limited size of stability regions. For instance, the family of Runge Kutta methods are considered to be inefficient for solving stiff equations. They are inefficient as the step size is controlled by stability rather than the accuracy requirements. However, explicit methods, including family of Runge Kutta methods, are still popularly used for temporal discretization with success; see [57, 90] in the scope of the DFM. Hence, we will also rely on explicit time discretization methods.*

Remark 6.3.5 *We mentioned in Remark 6.3.3 that the first-order Forward Euler method is typically not a stable choice for temporal discretization. Despite this*

fact, we will still resort to the first-order Forward Euler temporal discretization for few numerical experiments in Section 6.5 and assess the properties of the discrete system. The success of the first-order Forward Euler method, in practice, can be attributed to the numerical diffusion inherent in the construction of most of the spatial discretization schemes. As a result of such numerical diffusion, the eigenvalues of the discrete system tend to shift towards the stable regime of the spatially discretized system. We would like to point out that even though a numerical scheme with inherent (artificial) diffusion works well in practice, a numerical scheme with zero to minimal artificial damping is desired to replicate the "true" behaviour of the mathematical model. However, a numerical scheme with (almost) zero numerical diffusion yields a system of ODEs, which is definitely much more challenging to be solved via temporal discretization techniques, in particular via first-order Forward Euler due to the stability constraints. For instance, it is known that AUSM, a spatial discretization method discussed in Section 6.3.1, achieves zero-point interface capturing, i.e., zero numerical smearing. As a consequence, the application of a first-order Forward Euler method for temporal discretization brings out relevant numerical issues. We defer further numerical discussion to Section 6.5.

Remark 6.3.6 *Linear(ized) stability analysis of the discrete system can be misleading. This claim can be supported by the fact that problems (such as numerical instabilities) may occur if the reduction in the step length happens to place eigenvalues of the discrete system outside the absolute stability region due to the shape of the (stability) boundary. As a result, the notion that 'the finer the mesh, the better' may fail. The system eigenvalues are known to shift as per the chosen (spatial and temporal) discretization. Hence, non-linear stability conditions become critical for the (convergence) analysis of the numerical scheme in the presence of shocks or sharp gradients.*

Having discussed about the spatial and temporal discretization along with the aspects of the numerical implementation of the boundary conditions, we now have all the ingredients to delve into the error estimation and certification of a numerical scheme.

6.4   Error estimation and certification

The numerical predictions tend to differ from the exact solutions of the model of interest, even if by a small amount, due to various sources of error. Temporal and spatial discretization errors are the most significant sources of errors

in computational simulations. These discretization errors are expected to approach to zero as the grid is systematically refined. There is a close relationship between discretization error, the underlying mesh and the employed numerical method. For the purposes of verification and validation, it is necessary to have quantitative error bounds on each generated solutions, and to quantify the adequacy of an underlying grid.

Approximate solution methods introduce numerical errors, which should be kept minimal. The numerical errors may not be generated at a certain location in the flow domain, but transported as per the error dynamics. It is, hence, important to understand the generation and evolution of the errors. It is also important to have a mechanism that can account for the errors resulting from the spatial resolution or from the choice of the numerical schemes. The goal is to certify the fluid flow simulations (here in the scope of the DFM) and to characterize errors generated by different numerical schemes in problems of practical interest.

An approach called Error Transport Equation (ETE) [18], which is a fairly recent approach, has come to the fore for discretization error analysis. The ETE approach can have a continuous or a discrete formulation. This approach helps to predict discretization error using the numerical solution on only one grid. Using the ETE approach, error transport analysis can serve as a good criterion for mesh adaptation (local) and for solution improvement through defect correction. The ETE approach has been recently exploited for discretization error analysis in the scope of both unsteady and steady problems, including one-dimensional quasi Euler equations, Navier-Stokes equations, etc [18, 38, 143, 205]. However, such an analysis has not been carried out for the DFM so far. Hence, we strive to quantify the discretization error using several challenging numerical benchmarks in the scope of the DFM.

The use of non-linear error transport equation has been motivated in [18] and we exploit the same (general) finite-volume framework. Consider a partial differential equation (PDE), equivalent to (6.1), of the form:

$$\partial_t u + \underbrace{\partial_x f(u, u_x)}_{\text{evolution operator}} = s(u, u_x). \tag{6.46}$$

Consider a function $\tilde{u}(x, t)$ that approximates $u(x, t)$, but does not exactly satisfy (6.46). The approximation error can now be defined as follows:

$$\epsilon(x, t) = u(x, t) - \tilde{u}(x, t). \tag{6.47}$$

In an analogous manner, the spatial derivative of the approximation error is

given by:
$$\epsilon_x(x,t) = u_x(x,t) - \tilde{u}_x(x,t). \tag{6.48}$$

Using (6.47) and (6.48) in (6.46), we obtain:
$$\partial_t(\epsilon + \tilde{u}) + \partial_x f(\epsilon + \tilde{u}, \epsilon_x + \tilde{u}_x) = s(\epsilon + \tilde{u}, \epsilon_x + \tilde{u}_x). \tag{6.49}$$

Subtracting $\partial_t\tilde{u} + \partial_x f(\tilde{u}, \tilde{u}_x) - s(\tilde{u}, \tilde{u}_x)$ from both sides of (6.49), we have:
$$\partial_t\epsilon + \underbrace{\partial_x(f(\epsilon + \tilde{u}, \epsilon_x + \tilde{u}_x) - f(\tilde{u}, \tilde{u}_x))}_{\text{error evolution operator}} - (s(\epsilon + \tilde{u}, \epsilon_x + \tilde{u}_x) - s(\tilde{u}, \tilde{u}_x)) = -\underbrace{(R(\tilde{u}, \tilde{u}_x))}_{\text{residual}},$$
$$\tag{6.50}$$

where
$$R(\tilde{u}, \tilde{u}_x) = \partial_t\tilde{u} + \partial_x f(\tilde{u}, \tilde{u}_x) - s(\tilde{u}, \tilde{u}_x). \tag{6.51}$$

From now on, we will refer (6.50) as an error transport equation.

Remark 6.4.1 *An error transport equation, as in (6.50), is usually linearized to obtain an error evolution. A linearization of the flux function about $(\tilde{u}, \tilde{u}_x)$ is illustrated below:*

$$f(\tilde{u} + \epsilon, \tilde{u}_x + \epsilon_x) = f(\tilde{u}, \tilde{u}_x) + \partial_u f(\tilde{u}, \tilde{u}_x)\epsilon + \partial_{u_x} f(\tilde{u}, \tilde{u}_x)\epsilon_x + \mathbb{O}(||\epsilon||^2). \tag{6.52}$$

Using (6.52) in (6.50), a linearized transport equation can be obtained as shown below:

$$\partial_t\epsilon + \partial_x(\partial_u f(\tilde{u}, \tilde{u}_x)\epsilon + \partial_{u_x} f(\tilde{u}, \tilde{u}_x)\epsilon_x) - (\partial_u s(\tilde{u}, \tilde{u}_x)\epsilon + \partial_{u_x} s(\tilde{u}, \tilde{u}_x)\epsilon_x) \approx -(R(\tilde{u}, \tilde{u}_x)). \tag{6.53}$$

However, such a linearization is questionable. In [18], it has been shown that the linear error transport approach is not valid in several important cases. Given the shortcomings of the linear error transport approach, we will strive to solve the error evolution (6.50) in the scope of the DFM without resorting to any kind of linearization.

In principle, we neither have an exact solution to (6.46), nor to an error transport equation (6.50). However, similar to conventional approach of discretizing a PDE, the error transport equation can also be discretized. The discretization error can be estimated by concurrently solving the primal governing PDE along with an auxilliary (non-linear) error transport equation. Appropriate choices need to be made to discretize the evolution operator in (6.46), and to discretize the error evolution operator and the residual in (6.50). We will now step by step discuss: *(i)* the choice of discretization of the primal governing PDE, *(ii)* the choice of discretization of the evolution operator in the error transport equation, and *(iii)* the choice of evaluation technique of the error residual.

6.4.1 Discretization of the primal governing PDE

The spatial and temporal discretization of the PDE as in (6.46) has already been discussed extensively in Section 6.3. We also refer the reader to Section 6.3.1 for aspects pertaining to first-order and higher-order of spatial and temporal accuracy. We will omit further discussions about the discretization of the primal governing equation here, and instead focus on the discretization aspects related to the auxilliary error transport equation.

6.4.2 Discretization of the auxilliary error transport equation

The discretization of the auxilliary error transport equation constitutes the choice of discretization of the error evolution operator and the choice of evaluating corresponding residual. Some key aspects need to be kept in mind while developing a (continuous-time and discrete-time) approximation of the error. As explained in [18], the residual should be evaluated at a higher order than the primal equation, and the error transport equation should be discretized at the same or a higher order than that of the primal equation. This technicality is also elaborated in further detail in the sequel.

In the scope of the DFM for flow across a constant geometrical cross-section, the semi-discrete form of the primal equation (6.46) is given by:

$$\partial_t \bar{u} + G(\bar{u}) = 0, \tag{6.54}$$

where $G(\bar{u})$ is a discrete operator representative of the spatially discrete form of $\partial_x f(\bar{u}) - s(\bar{u})$. It should be noted that the dependence of the flux function and the source terms on the spatial derivatives of the solution has been dropped in view of the absence of such terms for the DFM with constant cross-section. The auxilliary error transport equation (6.50) takes the following form:

$$\partial_t \epsilon + \underbrace{\partial_x (f(\epsilon + \bar{u}) - f(\bar{u}))}_{\text{error evolution operator}} - (s(\epsilon + \bar{u}) - s(\bar{u})) = - \underbrace{(R(\bar{u}))}_{\text{residual}}, \tag{6.55}$$

with

$$R(\bar{u}) = \partial_t \bar{u} + \partial_x f(\bar{u}) - s(\bar{u}). \tag{6.56}$$

Equation (6.55) can be equivalently written as follows:

$$\partial_t \epsilon + \partial_x (f(\epsilon + \bar{u})) - \partial_x (f(\bar{u})) - (s(\epsilon + \bar{u}) - s(\bar{u})) = -(R(\bar{u})), \tag{6.57}$$

Using (6.54) in (6.56), we can write the residual $R(\tilde{u})$ as follows:

$$R(\tilde{u}) = -G(\tilde{u}) + \partial_x f(\tilde{u}) - s(\tilde{u}). \tag{6.58}$$

Let us now introduce a discrete operator $\tilde{G}(\tilde{u})$ that is representative of the spatially discrete form of $\partial_x f(\tilde{u}) - s(\tilde{u})$ in (6.58). As a consequence, (6.58) can be further simplified to

$$R(\tilde{u}) = -G(\tilde{u}) + \tilde{G}(\tilde{u}). \tag{6.59}$$

Using (6.59), it is easy to see that if the discrete operators G and $\tilde{G}$ are exactly same, the source of error dynamics will vanish. Furthermore, the error estimate will be asymptotically correct if the order of discretization of the residual is greater than that of the primal governing PDE; see [18] for a technical discussion. Hence, if G is obtained by a first-order spatial discretization method, then $\tilde{G}$ should be computed by employing higher-order spatial discretization technique. To summarize, the semi-discrete form of (6.55) or (6.57) can be written as follows:

$$\partial_t \epsilon + \frac{1}{\Delta x} \left(F_{i+1/2}^{\epsilon}(U_L^{\epsilon}, U_R^{\epsilon}) - F_{i-1/2}^{\epsilon}(U_L^{\epsilon}, U_R^{\epsilon}) \right) - \frac{1}{\Delta x} \left(F_{i+1/2}(U_L, U_R) - F_{i-1/2}(U_L, U_R) \right)$$

$$- S_i^{\epsilon} + S_i = \underbrace{\frac{1}{\Delta x} \left(F_{i+1/2}(U_L, U_R) - F_{i-1/2}(U_L, U_R) \right) - S_i}_{G(\tilde{u})} -$$

$$\underbrace{\frac{1}{\Delta x} \left(F_{i+1/2}^{\epsilon}(U_L^{\epsilon,\star}, U_R^{\epsilon,\star}) - F_{i-1/2}^{\epsilon}(U_L^{\epsilon,\star}, U_R^{\epsilon,\star}) \right) + S_i^{\star}}_{\tilde{G}(\tilde{u})}. \tag{6.60}$$

Here, $\frac{1}{\Delta x} \left(F_{i+1/2}^{\epsilon}(U_L^{\epsilon}, U_R^{\epsilon}) - F_{i-1/2}^{\epsilon}(U_L^{\epsilon}, U_R^{\epsilon}) \right)$ represents the spatially discrete counterpart of $\partial_x(f(\epsilon + \tilde{u}))$. The starred values in (6.60) refer to the ones which are approximated by a higher-order scheme (here, second-order), and the non-starred values refer to the quantities approximated by a first-order scheme. Any spatial discretization methodology, as discussed in Section 6.3.1, can be adopted to compute the numerical fluxes $F_{i+1/2}^{\epsilon}(U_L^{\epsilon}, U_R^{\epsilon})$, etc. While S_i is generally computed based on point-wise cell evaluations, the methodology to discretize $S_i^{\star}$ should be consistent with the order of discretization of the term $\frac{1}{\Delta x} \left(F_{i+1/2}^{\epsilon}(U_L^{\epsilon,\star}, U_R^{\epsilon,\star}) - F_{i-1/2}^{\epsilon}(U_L^{\epsilon,\star}, U_R^{\epsilon,\star}) \right)$. We omit the detailed discussion around the (consistent) approximation of $S_i^{\star}$ as we only consider the problems without source terms while performing numerical experiments in Section 6.5. Such a simplification is pursued to make the discontinuous features more prominent,

and to exclude the role of physical damping in order to (some extent) quantify the amount of numerical viscosity. After spatial discretization, the semi-discrete form, as in (6.60), can be evolved over time by using any of the temporal discretization methodologies which have been discussed in Section 6.3.3. We have now completely set up the framework for error estimation and certification via the ETE approach. Next, we delve into the numerical experiments to quantify the performance of the numerical method.

6.5   Numerical Experiments

Numerical experiments are required to assess the performance of the introduced numerical schemes in terms of the inherent accuracy, stability, etc., and, thus, comment on their strengths and weaknesses. In this section, we test the model and the numerical method on problems with exact reference solutions, and also carry out grid refinement studies. Furthermore, we perform various numerical experiments to test relevant numerical properties using different combinations of spatial and temporal discretization schemes discussed in Sections 6.3.1 and 6.3.3. For instance, we check if the numerical schemes follow Abgrall's principle which states that: "A flow, uniform in pressure and velocity, must remain uniform in the same variables during its time evolution." The impact of various existing limiters on the numerical response is also studied in the scope of few numerical tests. In the pursuit of error estimation and certification, we discretize the primal governing PDE along with the auxilliary error transport equation to provide additional metric for establishing the relative order of merit of the numerical schemes.

The considered test cases are extensively described next.

6.5.1   Shock-tube type test-case

Shock-tube type problems are interesting as they test the ability of a numerical scheme to handle the initial data that is far removed from the state of equilibrium. Generally, a membrane, placed at a certain location in the spatial domain, is ruptured, and this eventually leads to the evolution of waves depending upon the initial states of the system. These waves travel in either directions of the spatial domain depending upon the states of the system, and travel towards the boundary of the spatial domain. Clearly, the role of the boundary conditions is also important. However, in the scope of the following test cases, the end time of the simulation is chosen such that the emanating waves do not reach

the boundaries of the spatial domain under consideration. Hence, the boundary conditions are not specified in the test cases that follow within this sub-section.

Case A

This numerical experiment is inspired from [181]. This is a test case to verify the property of pressure invariance and, hence, check whether the numerical schemes follow Abgrall's principle. Furthermore, such a test case also uncovers the deficiency of numerical methods in capturing the material interface or the contact discontinuity. Moreover, this test case aids to test the model and numerical method with exact reference solutions. Such properties, except for testing the efficiency of the numerical schemes in capturing the (contact) discontinuites, have not yet been theoretically or numerically tested in the literature on the numerical methods in the scope of the DFM.

This test case does not consider the slip between the two phases, i.e., $v_g = v_\ell = v$. Furthermore, the model is considered to be inviscid, and, as a consequence, the role of the gravitational and frictional source terms is neglected. In such a setting, the DFM essentially becomes similar to the model in [181]. In addition to the zero slip and the absence of source terms, this test case uses Tait's equation of state for the phases under consideration, which are given by:

$$\rho_{\ell/g}(p) = \rho_{\ell/g}^0 \left(\frac{(p/p^0) + \eta_{\ell/g}}{1 + \eta_{p/s}} \right)^{1/\gamma_{\ell/g}}, \tag{6.61}$$

where the subscripts ℓ and g refer to the corresponding value for the liquid phase and the gaseous phase, respectively, $\rho_{\ell/g}$ indicate the densities of the phases which are given as a function of pressure p only and, hence, the equation of state is barotropic in nature. Here, p^0 is the reference pressure, $\rho_{\ell/g}^0$ stands for the density of the phases (indicated by the subscript) at the reference pressure, and $\eta_{\ell/g} \geq 0$ and $\gamma_{\ell/g} > 1$ are phase-specific constants. The equation of state, as in (6.61), is different from those in (2.4) and (2.5), which we considered earlier in Section 6.2 for studying the hyperbolicity of the DFM with its corresponding closure laws. Here, we consider the Tait's equation of state for the phases since the exact solutions under such a setting have been reported in [181], and we seek to exploit the exact solutions for benchmarking the numerical schemes. It is worth mentioning that the hyperbolicity of the model with the aforementioned modifications can also be studied. We omit the corresponding discussion, and refer the reader to follow the principles introduced in Section 6.2 for this purpose.

Furthermore, in the pursuit of assessing whether the numerical schemes fol-

low Abgrall's principle, the initial velocity and pressure are considered to be uniform. Also, a pure liquid is considered at the left side of the membrane located at the centre of the spatial domain, i.e., $x_0 = 2$, and a pure gas is considered at the right side of the membrane. In essence, a large density ratio is considered across the membrane separating the two phases. Mathematically, the initial conditions are as follows: $\alpha_{g,L} = 0, p_L = 1, v_L = 100$ and $\alpha_{g,R} = 1, p_R = 1, v_R = 100$, where L (resp., R) in the subscript indicates the states at the left (resp., right) side of the membrane. Under the chosen initial conditions, the resulting numerical solution essentially corresponds to the translation of the (material) interface. The coefficients of the equations of state vary discontinuously from one medium to another across the interface. The spatial domain is assumed to be of length $L = 4$. The spatial step size and the temporal step size are considered to be $\Delta x = 2^{-6}$ and $\Delta t = 2^{-9}$, respectively. For the purposes of numerical simulation of the model, the values of the constants appearing in the Tait's equation of state are chosen as follows: $\rho_\ell^0 = 1, \eta_\ell = 3000, \gamma_\ell = 7, \rho_g^0 = 10^{-3}, \eta_g = 0$ and $\gamma_g = 7/5$.

Before delving into the numerical study, we briefly discuss how to obtain the primitive variables if the conservative variables are known. To this end, $\mathbf{q}$ in (6.1), under the assumption of the zero slip between the phases, takes the following form:

$$\mathbf{q} := \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} \alpha_\ell \rho_\ell \\ \alpha_g \rho_g \\ (\alpha_\ell \rho_\ell + \alpha_g \rho_g) v \end{pmatrix}. \tag{6.62}$$

Using (6.62), we have that:

$$(q_1 + q_2) v = q_3 \text{ or } v = \frac{q_3}{q_1 + q_2}. \tag{6.63}$$

Using (6.62) and (6.61), we can write:

$$q_1 = \alpha_\ell \rho_\ell^0 \left(\frac{(p/p^0) + \eta_\ell}{1 + \eta_\ell} \right)^{1/\gamma_\ell}. \tag{6.64}$$

Similarly, by using (6.62), (6.61) and the fact that $\alpha_g + \alpha_\ell = 1$, we have that:

$$q_2 = \alpha_g \rho_g^0 \left(\frac{(p/p^0) + \eta_g}{1 + \eta_g} \right)^{1/\gamma_g} = (1 - \alpha_\ell) \left(\rho_g^0 \left(\frac{(p/p^0) + \eta_g}{1 + \eta_g} \right)^{1/\gamma_g} \right). \tag{6.65}$$

Equation (6.64) can also be re-written as follows:

$$\alpha_\ell = \frac{q_1}{\rho_\ell^0 \left(\frac{(p/p^0) + \eta_\ell}{1 + \eta_\ell} \right)^{1/\gamma_\ell}}. \tag{6.66}$$

Furthermore, using (6.65) and (6.66), we have that:

$$q_2 = \left(1 - \frac{q_1}{\rho_\ell^0 \left(\frac{(p/p^0)+\eta_\ell}{1+\eta_\ell}\right)^{1/\gamma_\ell}}\right)\left(\rho_g^0 \left(\frac{(p/p^0)+\eta_g}{1+\eta_g}\right)^{1/\gamma_g}\right). \qquad (6.67)$$

Re-arranging (6.67) yields the following expression:

$$\left(\frac{(p/p^0)+\eta_g}{1+\eta_g}\right)^{1/\gamma_g} - \frac{q_1}{\rho_\ell^0 \left(\frac{(p/p^0)+\eta_\ell}{1+\eta_\ell}\right)^{1/\gamma_\ell}}\left(\frac{(p/p^0)+\eta_g}{1+\eta_g}\right)^{1/\gamma_g} = \frac{q_2}{\rho_g^0}, \qquad (6.68)$$

which can be solved for p given q_1 and q_2. This equation is highly non-linear in p, and needs to be solved at every time iteration while converting from conservative to primitive variables, which are needed in the computation of the numerical flux. Having shown the way to obtain the primitive variables from the conservative variables, we now perform the numerical study on this test case.
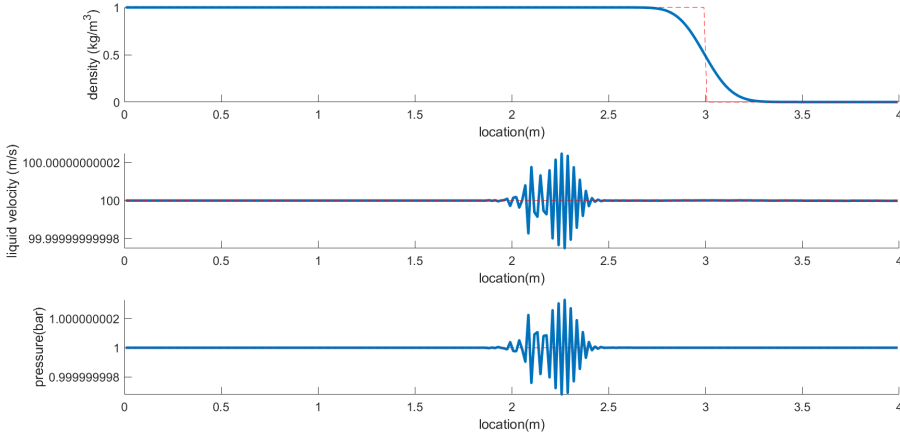

Figure 6.4: Case A: Behaviour of the density, (liquid) velocity and pressure at $t = 0.01$. Here, $\Delta x = 2^{-6}$, $\Delta t = 2^{-9}$ and the first-order FVS scheme and the first-order Forward Euler method has been employed for the discretization. The exact solution is depicted by red-dashed lines and the numerical prediction is shown in solid blue.

Figure 6.4 shows the profile of the density, (liquid/gas) velocity, and the pressure at time $t = 0.01$ when the (first-order) FVS scheme is employed for the spatial discretization and a first-order Forward Euler scheme for the temporal discretization. As per the initial conditions, the velocity and the pressure are supposed to remain uniform for all time instants, i.e., $p = 1$ and $v = 100$ for all $t \in [0, 0.01]$, and the exact solution of the density at $t = 0.01$ is [181]:

$$\rho := \alpha_g \rho_g + \alpha_\ell \rho_\ell = \begin{cases} 1, & x \leq 3, \\ 0, & x > 3. \end{cases}$$

Basically, the contact discontinuity travels from $x = 2$ to $x = 3$ during the considered time-horizon. We can clearly observe that the numerical solution of the density, velocity and pressure is qualitatively indicative of the exact solution. However, some oscillations in the velocity and pressure profile can be observed. Also, the contact discontinuity, visible in the profile of the density, is smeared out, which reveals that the numerical scheme possesses a relatively large amount of numerical viscosity. Given the fact that the velocity and the pressure profile are not exactly uniform, the FVS scheme does not satisfy the Abgrall's principle in an exact sense.

Next, we investigate the role of the higher-order spatial and temporal discretization methods in the (numerical) satisfaction of the Abgrall's principle. Figure 6.5 again shows the numerical solution of the density, velocity and pressure at $t = 0.01$. Here, different combinations of higher-order spatial and temporal discretization methods have been assessed in addition to the discretization setting discussed in the previous paragraph. We employ the Osher limiter and the FVS scheme for a (formally) second-order spatial discretization. Furthermore, we apply three different variants of temporal discretization schemes, namely, the second-order Runge Kutta, the third-order Runge Kutta and the fourth-order Runge Kutta method. From Figure 6.5, we can clearly observe that, among the considered combinations, the second-order spatial discretization along with the fourth-order Runge Kutta method is the most closest to the satisfaction of the Abgrall's principle. Furthermore, we can observe that the nature of the oscillatory perturbations around the uniform state is different for all the considered combinations. This basically hints at inherently different evolution mechanisms of the solution (and the error) under the implementation of various (combinations of) numerical schemes. Moreover, we observe that the higher-order schemes yield (as expected) a less smeared contact discontinuity compared to the first-order schemes. It is interesting to note that the second-order spatial discretization along with the considered choices of higher-order
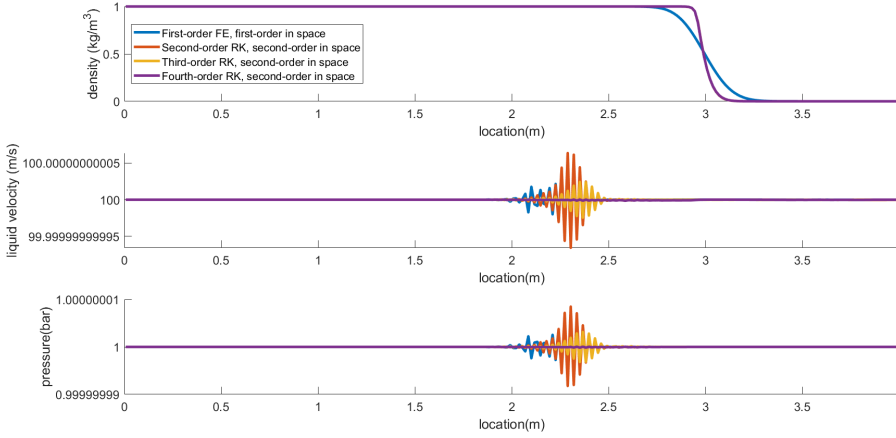
Figure 6.5: Case A: Behaviour of the density, (liquid) velocity and pressure at $t = 0.01$ for different combinations of spatial and temporal discretization schemes with $\Delta x = 2^{-6}$ and $\Delta t = 2^{-9}$. The red and the yellow curves, not visible in the top plot, are basically hidden behind the purple curve.

Runge Kutta methods results in an exactly same evolution of the contact discontinuity; see the top plot in Figure 6.5.

We now construct the numerical fluxes, i.e., perform the spatial discretization, in several different ways as discussed in Section 6.3.1, and assess which scheme most accurately satisfies the Abgrall's principle in a numerical sense. We only employ first-order variants of FVS, VanLeer, AUSM, AUSMV and AUSMD along with a first-order Forward Euler scheme. Figure 6.6 shows the velocity and the pressure profile at $t = 0.01$ for different spatial discretization schemes. We can observe that each of the considered schemes possesses (a different type of) oscillatory perturbations around the uniform state. Similar to an earlier observation, it is clear that the inherent evolution mechanisms of the solutions are different for a certain combination of the spatial and temporal discretization methodology. Furthermore, the propagation of the information towards the left of $x = 2$ is quite visible for the FVS and the VanLeer scheme, but not for the other considered schemes. Among the considered schemes, the AUSMV one clearly outperforms and (almost) satisfies the Abgrall's principle. We can also observe that the AUSM scheme is the worst since the resulting solution deviates the most
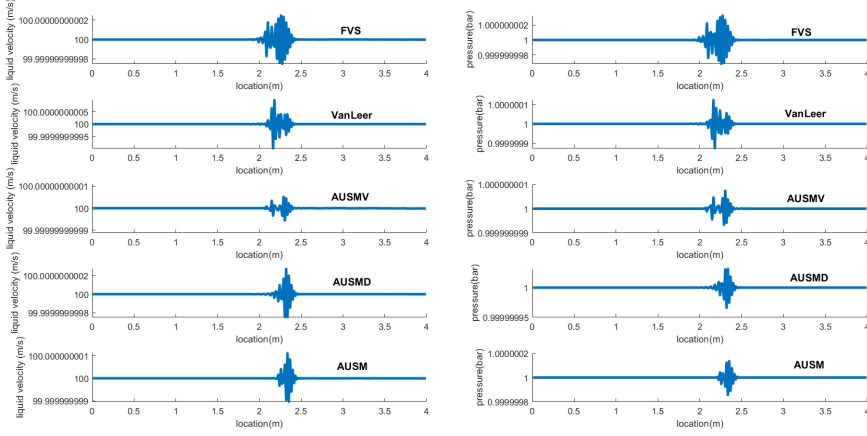
Figure 6.6: Case A: (left) Behaviour of the (liquid) velocity and (right) Behaviour of the pressure at $t = 0.01$ with $\Delta x = 2^{-6}$ and $\Delta t = 2^{-9}$. Here the numerical flux is approximated via several schemes: FVS, VanLeer, AUSMV, AUSMD and AUSM, and the fourth-order Runge Kutta is employed for the temporal discretization.

from the uniform state.

We also consider the role of the limiters in assessing the impact on the numerical satisfaction of the Abgrall's principle. In particular, we consider two different choices of the limiters, i.e., Osher and MinMod generalized limiter among the large assortment of available limiters. Given the fact that the (first-order) AUSMV scheme most closely satisfies the Abgrall's principle, we construct a second-order variant of AUSMV by using the two aforementioned limiters. Furthermore, using the fact that the second-order FVS scheme (with the Osher limiter) along with the fourth-order Runge Kutta (almost) satisfies the Abgrall's principle (see Figure 6.5), we will employ the fourth-order Runge Kutta method for the temporal discretization of the setting under consideration. Figure 6.7 depicts the profile of the density, velocity and pressure at $t = 0.01$ under the implementation of the second-order AUSMV (using the Osher and MinMod generalized limiters) and the fourth-order Runge Kutta method. Clearly, the limiters impact the evolution mechanism. The MinMod generalized limiter is observed to yield much less numerical oscillations around the uniform state. It is worth
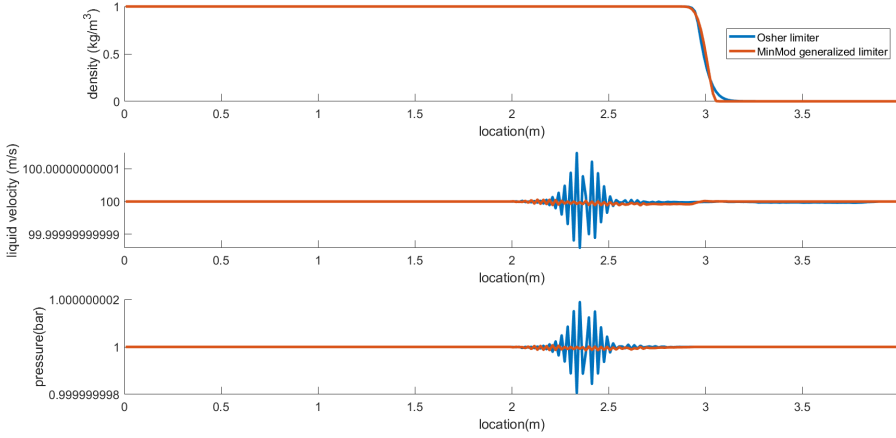
Figure 6.7: Case A: Behaviour of the density, (liquid) velocity and pressure at $t = 0.01$ for two different choices of the limiters, namely, the Osher and the MinMod generalized limiter. Here, $\Delta x = 2^{-6}$, $\Delta t = 2^{-9}$ and the second-order AUSMV and the fourth-order Runge Kutta method has been used for the discretization.

emphasizing that the role of the limiters is problem specific, and by now means should the MinMod generalized limiter be considered superior to the Osher limiter in a generic sense.

Case B

This shock-tube type numerical experiment is inspired from [57]. It is considered here to showcase the potential benefits of the modified model of the speed of sound in the two-phase mixture over the existing surrogate one and, thus, to study the importance of the role of the speed of sound of a two-phase mixture. This test case is also considered in the scope of error estimation and certification. Similar to [57], a 100 m horizontal pipe is considered which is initially separated by a membrane into a left and a right state at $x_0 = 50$ m. Figure 6.8 represents the initial conditions of the numerical experiment. As a consequence of the chosen initial conditions and the underlying nature of the DFM, the numerical solution is composed of three waves: a 1-shock, a 2-contact discontinuity, and a 3-shock, which start emanating upon the rupturing of the membrane. Furthermore, we consider that the two phases are governed by the
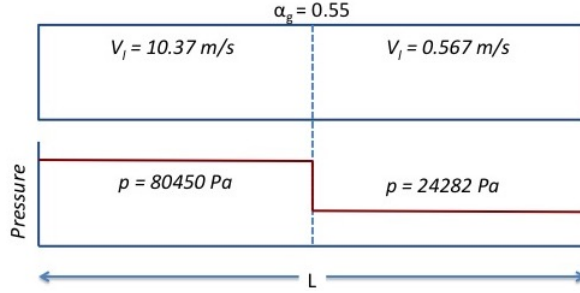
Figure 6.8: Case B: Initial conditions of the experiment.

equations of state as described in (2.5) and (2.4). Moreover, frictional effects are neglected, and the parameters in the Zuber-Findlay slip law are considered to be: $K = 1.07$ and $S = 0.216$. Next, we provide numerical observations in line with the aforementioned objectives of this test case.
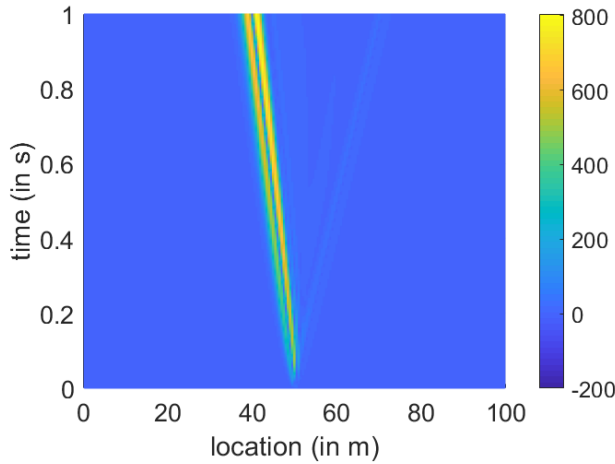

Figure 6.9: Case B: Absolute difference in the pressure computation (in Pa) upon using the existing surrogate and the proposed model for the sound speed of the two-phase mixture.

Sound speed model: We consider that the spatial domain is discretized into 200 FV cells, and that the CFL = 0.8. Furthermore, we use the first-order FVS scheme along with the first-order Forward Euler method for the purposes of the numerical discretization. Figure 6.9 depicts the space-time diagram of the absolute difference in the pressure obtained by employing the two different sound speed models. The numerical differences in the propagation behaviour is clearly evident. It must be pointed out that any other scheme can also be used for the spatial discretization, and the similar analysis can be carried out. The location/profile of the fronts, under two different sound speed models, is observed not to be significantly different in the scope of this test case. The numerical differences are problem specific, and are better observed in another test case discussed in Section 6.5.2. However, it is worth mentioning that the approach where the proposed sound speed model is used is computationally expensive as it involves solving the eigenvalue problem at every time iteration. We defer further discussion to Section 6.5.2 in the scope of the pros and the cons of the proposed sound speed model and the existing surrogate one.
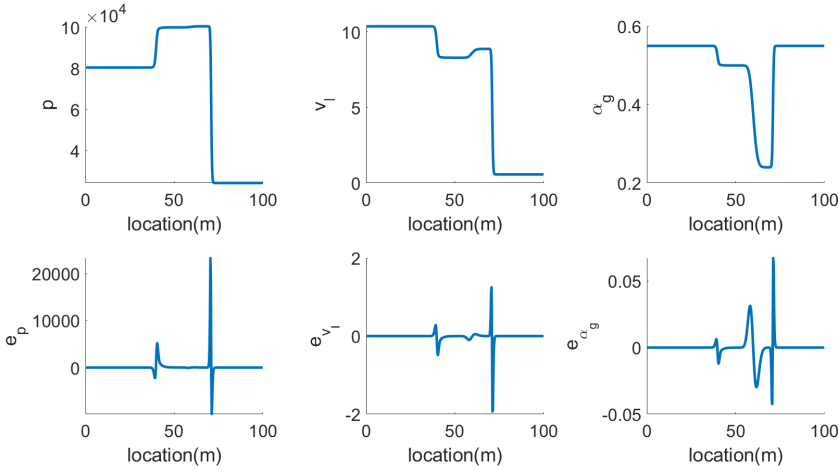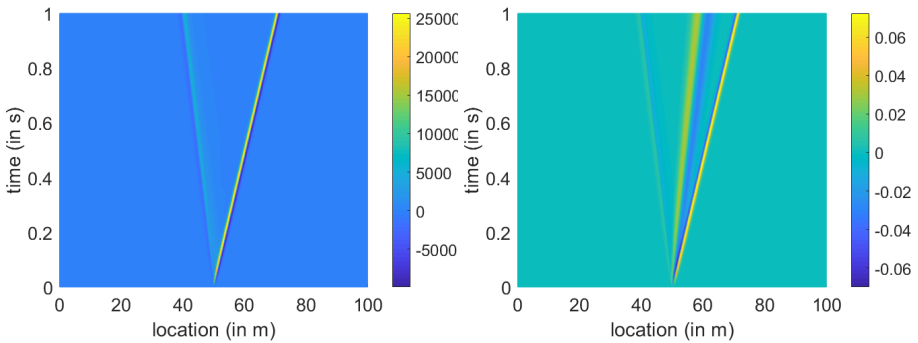

Figure 6.10: Case B: (top row: left to right) Behaviour of the pressure, the liquid velocity and the gas void fraction, and (bottom right: left to right) behaviour of the approximate error in the pressure, the approximate error in the liquid velocity and the approximate error in the gas void fraction. Here, CFL = 0.8 and the computational domain is uniformly divided into 400 FV cells.

Error estimation and certification: In this numerical case study, we divide the
spatial domain into 400 FV cells, consider CFL = 0.8, employ the FVS scheme
for the spatial discretization and the Osher limiter for the higher-order approx-
imation of the numerical fluxes, and use a first-order Forward Euler scheme for
the temporal discretization. Figure 6.10 shows the discrete approximation of
the pressure, the liquid velocity and the gas void fraction along with the corre-
sponding approximate errors at time $t = 1$. It can be observed that the error in
the physical quantities (say, pressure, liquid velocity, gas void fraction, etc.) is
dominated by the errors near the numerically resolved shock. We can see that
the error in the 2-contact discontinuity has minimal impact on the pressure field.
As a consequence, it can be deduced that the smearing of the 2-contact discon-
tinuity should not be of concern if the pressure is the (only) quantity of interest.
Furthermore, we can see that the source of the dominant error in all the quanti-
ties depicted in Figure 6.10 is the behaviour around the 3-shock. Also, the error
in the gas void fraction is almost equally attributed to the features around the
2-contact discontinuity as to those around the 3-shock. Basically, the gradient at
the moving interface dictates the dominant sources of the error in the fields of
interest. Figures 6.11a and 6.11b demonstrate the space-time evolution of the
approximate error in the pressure and the gas void fraction, respectively. The
observed propagation behaviour (and the speed) is closely related to the un-

(a) Approximate error in the pressure (in Pa).   (b) Approximate error in the gas void fraction.

Figure 6.11: Case B: A space-time diagram for the approximate error in the pressure and
gas void fraction. Here, CFL = 0.8 and the computational domain is divided
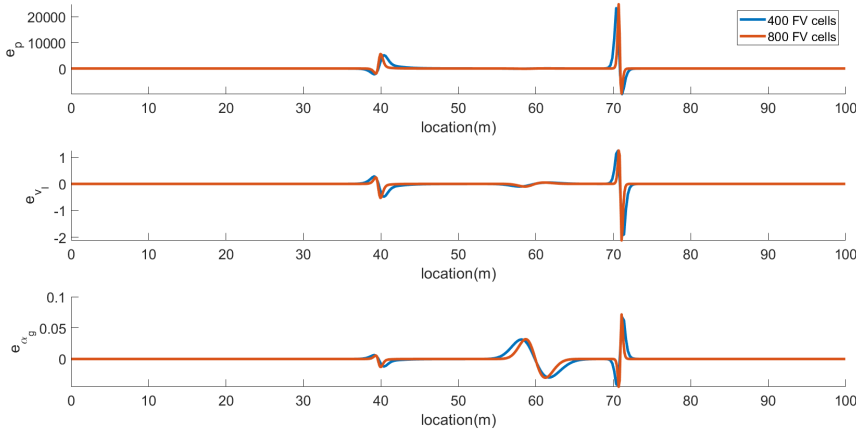into 400 FV cells.

Figure 6.12: Case B: behaviour of (top) approximate error in the pressure, (middle) approximate error in the liquid velocity, (bottom) approximate error in the gas void fraction at $t = 1$ and CFL $= 0.8$ on a mesh with 400 and 800 FV cells.

derlying eigenstructure of the system. Figure 6.12 shows the approximate error in the pressure, the liquid velocity and the gas void fraction at $t = 1$ under the choice of 2 different mesh resolutions. Here, we use 400 and 800 FV cells, and adapt the time step such that the CFL is still equal to 0.8. It can be observed that the behaviour of the approximate error is almost similar for both considered settings. However, two important observations can be made. The first observation is that the spatial width, across which the approximate errors are dominant, is less for the finer mesh compared to that for the coarser mesh. This is related to the fact that using a larger number of FV cells implies a smaller spatial step size and, consequently, less numerical viscosity and a better capability in resolving the sharp gradient feature. The second observation is that the maximum (approximate) error in the quantities of interest almost remains constant with the mesh refinement. This behaviour is unlike the one usually observed when the linear error transport approaches, which results in an unbounded growth of the error under mesh refinement, are employed. This observation reveals the power of the non-linear error transport approach.

We also implement other schemes, namely, AUSM, AUSMV, AUSMD and VanLeer for the spatial discretization instead of employing the FVS scheme. We observe similar behaviour as the one depicted in Figures 6.10 and 6.12 under

the consideration of the AUSM, AUSMV and VanLeer scheme and, hence, we omit the related numerical results. However, the behaviour under the implementation of the AUSMD scheme deserves some discussion. Consider that we only solve a primal governing PDE by using a first-order variant of the AUSMD scheme along with a first-order Forward Euler temporal discretization on the computation domain with 400 FV cells. Figure 6.13 demonstrates the pressure, liquid velocity and the gas void fraction at $t = 1$ for the considered setting. While qualitatively the numerical solution of the quantities of interest are similar to the other numerical schemes (see the first row in Figure 6.10), it can be observed that there are significantly high amount of oscillations post the 2-contact discontinuity. These oscillations are observed for coarser spatial mesh resolutions as well as for lower CFL numbers. It seems that the AUSMD scheme is not tailored to accurately resolve the features that are observed, in particular the 2-contact discontinuity. It is worth mentioning that the AUSMD scheme has been employed for other kind of test cases in the past. However, it has never been tested for the case that we have considered. The observations hint to the need of modifications in the numerical approximation of the flux for resolving
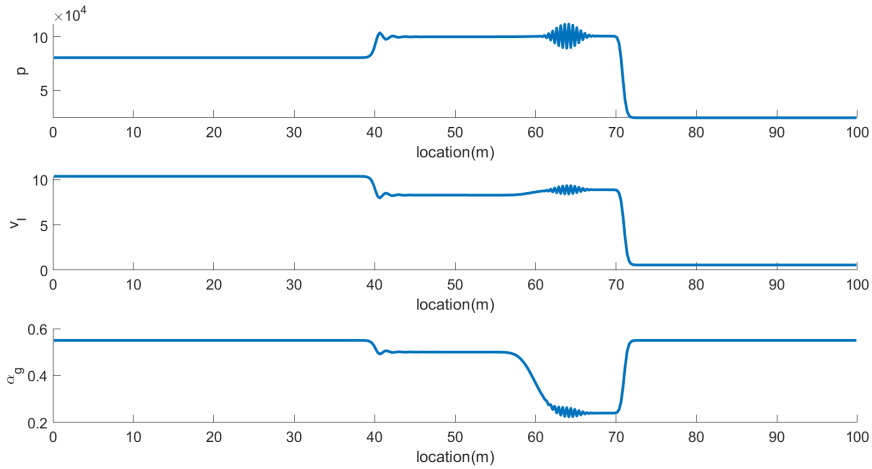

Figure 6.13: Case B: behaviour of (top) the pressure, (middle) the liquid velocity, (bottom) the gas void fraction at $t = 1$ and CFL = 0.8 on a uniform mesh with 400 FV cells. Here, the first-order AUSMD and the first-order Forward Euler method has been employed for discretization.

the aforementioned issues. Furthermore, it should be mentioned that the evaluation of the error residual requires a higher-order approximation. The oscillations, observed in the first-order implementation, tend to aggravate under the higher-order implementation of a numerical scheme and, consequently, pollute the evolution of the numerical solution and the approximate error. Employing the AUSMD scheme for numerically solving the primal governing PDE and the auxiliary error transport equation is, hence, not considered.

6.5.2   Fast Transients

This numerical experiment is inspired from [57]. One of the objectives of this test case is to delve into the relative assessment of the numerical schemes by pursuing grid refinement studies. Another objective is to showcase the importance of considering compressibility effects of both phases in obtaining the model of the speed of sound for the two-phase mixture.

Consider a 1000m long pipe with the following initial conditions:

$$\alpha_g = \begin{cases} 0.01, & x \leq 750, \\ 0.9, & x > 750 \end{cases} \quad , \quad p = 1e5 \quad \forall x \in [0, 1000]. \tag{6.69}$$

Additionally, a zero slip is considered between the two phases, i.e., $v_\ell = v_g = v$. Furthermore, the fluids are initially considered to be stagnant throughout the pipe, i.e., $v = 0$. Moreover, we consider that the two phases are governed by the equations of state as described in (2.5) and (2.4). A pulse in the pressure and fluid velocities is now introduced by increasing the inlet liquid flowrate from 0 to 0.3 kg/s in 0.0025 s, and the pressure at the outlet side of the pipe is kept constant. When these pulses experience the sudden jump in the gas void fraction, the changes in the compressibility, the mixture density and the speed of sound in the two phase mixture produce reflecting waves in the pressure and velocity fields. Finally, the simulation considers frictional effects, and is run over a time horizon of 10 seconds, i.e., $t \in [0, 10]$.

We first assess the numerical differences under the use of the surrogate and the proposed model for the speed of sound of the two-phase mixture. To this end, we consider that the spatial domain is discretized into 200 FV cells, and that the CFL = 0.8. Furthermore, we use the first-order AUSMV scheme and characteristic-based boundary treatment along with the first-order Forward Euler method for the purposes of the numerical discretization. Figure 6.14 depicts the space-time diagram of the absolute difference in the pressure obtained by
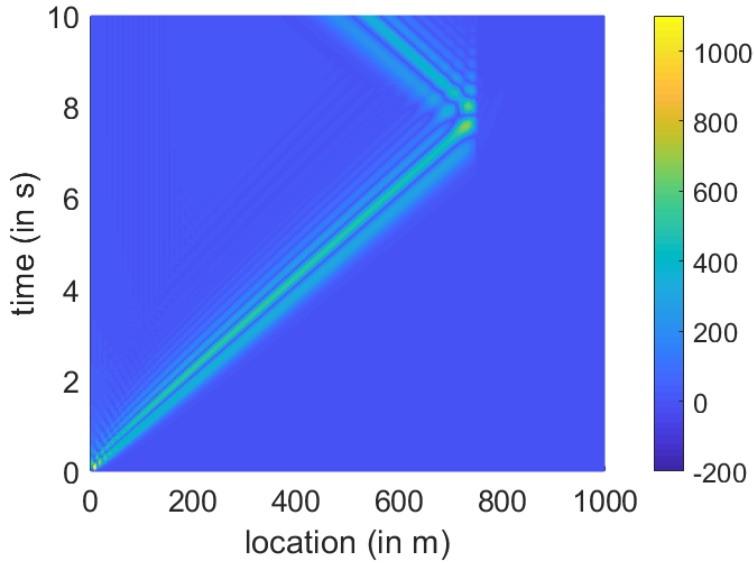
Figure 6.14: Fast Transients Test Case: Absolute difference in the pressure computation (in Pa) upon using the existing surrogate and the proposed model for the sound speed of the two-phase mixture.

employing the two different sound speed models. Clearly, there is a numerical impact. Before we delve further into the numerical differences, it is worth observing that the reflection phenomena at $x = 750$ is clearly visible. A space-time diagram of the absolute difference in the liquid velocity can also be analogously obtained. In order to provide further (visual) insights into the numerical differences, we study the profile of the gas void fraction, liquid velocity and pressure at $t = 10$. From Figure 6.15, we can observe that there is no difference in the evolution of the gas void fraction upon employing different sound speed models. Some differences can however be observed in the profile of the liquid velocity and the pressure. If we carefully look at the middle plot in Figure 6.15, some numerical oscillations can be observed if the numerical fluxes are approximated using the proposed sound speed model. Such oscillations are much more bounded or not observed when the surrogate sound speed model is used for the flux approximation. At a first glance, this might be viewed as a disadvantage of the proposed model over the surrogate one. However, such an
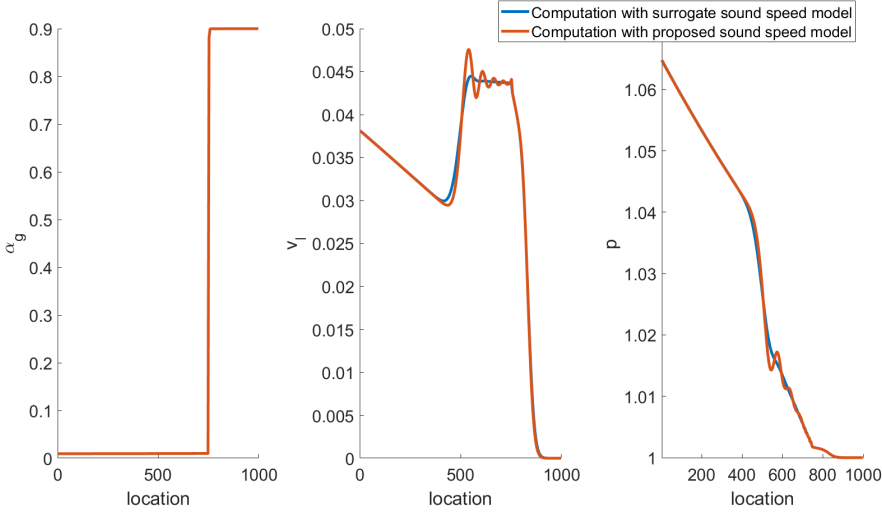
Figure 6.15: Fast Transients Test Case: behaviour of (left) the gas void fraction, (middle) the liquid velocity, (right) the pressure at $t = 10$ and CFL $= 0.8$ on a uniform mesh with 200 FV cells. Here, we have used the first-order AUSMV scheme and characteristic-based boundary treatment along with the first-order Forward Euler method for the numerical discretization.

observation can be supported from the following reasoning. An accurate sound speed model will consequently lead to an accurate evolution of the (moving) discontinuities/fronts. This is also in a way observable in the middle plot in Figure 6.15 where we can see that the front around $x = 500$ is much sharper if the computations are performed with the proposed sound speed model than with the surrogate one. It is well-known that "the sharper the front, the higher is the possibility to run into numerical oscillations". Basically, a more accurate resolution of the evolving front comes at the expense of the oscillations in the numerical predictions. However, the occurrence of the numerical oscillations is problem specific. For instance, in Case B, Section 6.5.1, we do not observe any numerical oscillations even though the proposed approach yields a slightly sharper front than the one that results with the existing surrogate models. To conclude, we can say that the sound speed model has a clear role in terms of the overall viscosity of the numerical scheme, and that the proposed sound speed model is numerically less viscous and eventually devoid of self-dissipating
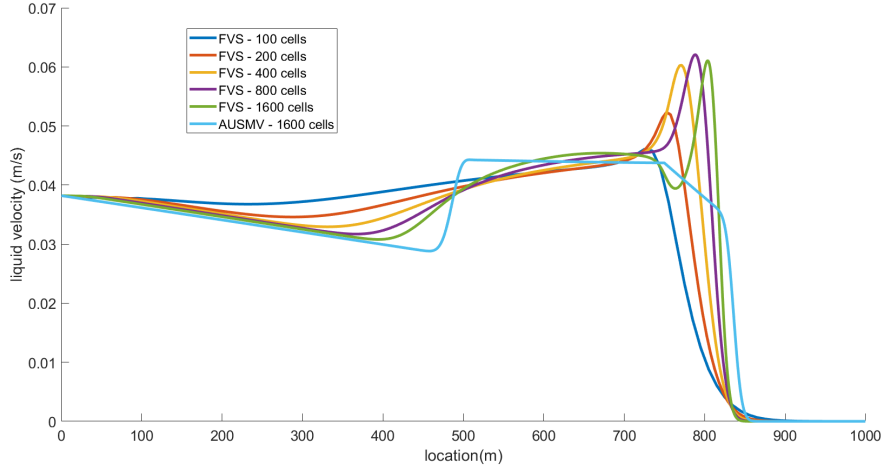
Figure 6.16: Fast Transient Test Case: Behaviour of the liquid velocity at $t = 10$ and CFL
$= 0.8$. Here, surrogate sound speed model has been used for the approximation of the numerical fluxes.

mechanisms that possibly lead to oscillatory numerical behaviour.

Next, in the scope of the grid refinement studies, we will only employ the
FVS scheme and the AUSMV scheme among all the other considered methods
for the approximation of the numerical fluxes since these are representative
enough for the point that we wish to make. We generate a numerical solution
by using the AUSMV scheme on a mesh with 1600 FV cells. Furthermore, we
also perform computations by using the FVS scheme over different spatial mesh
resolutions. In particular, we compute the numerical prediction with the FVS
over the spatial domain with $100, 200, 400, 800$ and 1600 FV cells. From Figure
6.16, we can clearly observe that the FVS and the AUSMV scheme do not yield
similar features at a finer spatial step size.

We recall that the authors in [57] considered the AUSMV scheme at a finer
mesh resolution to yield a reference solution, and used it to benchmark other
numerical schemes. At this point, it is also worth emphasizing that the nu-
merical solutions computed from various numerical schemes are theoretically
expected to converge to the same solution under grid refinements. However,
this is not the case as per the results shown in Figure 6.16. This instigates the

question whether it is correct to consider solutions generated by the AUSMV scheme on a finer grid as the reference solutions. Also, the reasons behind discarding the spiking behaviour of the solutions produced by the FVS scheme, and whether such a spiking behaviour is physical do not exist in the literature. It is observed that the FVS scheme yields such a spiking behaviour only after the front coming from the left side of the spatial domain interacts with the interface at $x = 750$. The spiking behaviour seems unlike an oscillatory behaviour and, hence, it becomes more challenging to ensure the validity of the obtained numerical result. It is, in principle, hard to answer which numerical result is trustworthy. This aspect deserves detailed investigation in the future.

6.6   Conclusions

In this chapter, we have performed the numerical analysis of the Drift Flux Model (DFM). We discussed about the conditional hyperbolicity of the model, and proposed to account for compressibility effects in the model of the speed of sound of the two-phase mixture. We then revisited the existing basic concepts about the (first- and higher-order) spatial and the temporal discretization along with the aspects related to the numerical implementation of the boundary conditions in the scope of the DFM. Furthermore, we exploited the approach of non-linear error transport, and used the acquired discretization knowledge to build a framework for discretizing the primal governing partial differential equation and the associated auxilliary error transport equation. This paved the way to understand the generation and the transport mechanisms of the numerical error.

Finally, we performed a large number of numerical experiments ranging from shock-tube type test cases to fast transient scenarios. In particular, we tested the model and the numerical method with exact reference solutions, and checked whether the numerical schemes satisfy the Abgrall's principle. We found out that the AUSMV scheme does satisfy the Abgrall's principle, and outperforms the other considered schemes, namely, the FVS, VanLeer, AUSM and AUSMD scheme. We also numerically studied the role of the speed of sound of the two-phase mixture, and observed that such a model is highly important from the numerics point of view. Particularly, it impacts the numerical viscosity of the scheme and also yields differences in the location (and the speed) of the evolving fronts. The approximate error, obtained using the approach of non-linear error transport, is observed to be dominant near the captured numerical shock or a front with a large gradient. Also, the maximum approximate error does not

grow unboundedly under mesh refinements. While all the considered schemes (except AUSMD to some extent) converge to the same solution (under grid refinements) in the scope of the shock-tube type test cases, there is still an open question in the context of the scenario that is representative of the propagation of fast transients. We emphasize that the FVS and the AUSMV scheme produce different qualitative features in the scope of the fast transient test case, and it is extremely hard, but important to develop a sound reasoning to validate/discard the observed numerical results. Such a validation would also instigate the fundamental modifications in the existing numerical methods so as to produce physically consistent results.

As far as the relative order of merit of the numerical schemes is concerned, it is largely problem specific. However, out of the considered schemes, we would recommend to use the FVS and the AUSMV scheme, and assess the convergence of these methods in the scope of the test case of interest. The reasons behind discarding the consideration of other schemes are as follows: *(i)* the AUSM scheme does not satisfy the Abgrall's principle, and is also known to be highly oscillatory even in the first-order setting, *(ii)* the FVS and the VanLeer scheme largely produce similar results, and *(iii)* the (first-order) AUSMD scheme is observed to possess the numerical oscillations post the 2– contact discontinuity. Further detailed investigation into the relative merit of the schemes is the subject of future research.

One of the important frameworks developed in this chapter pertains to the numerical estimation of the discretization error. It should be mentioned that the results (obtained results and the ones that can be generated on other test cases) will be highly useful in the scope of certifying a reduced-order model as well. Although we do not employ such a framework in the scope of certifying a reduced-order model in Part II of this dissertation, we stress that a dedicated study is needed to make an improved assessment of the error between the solution obtained by reduced-basis approximations and the (expected) exact response.

Part II: Model order reduction for transport-dominated problems

This part is dedicated to the contributions pertaining to the model order reduction for transport-dominated problems. In particular, we strive to address questions Q6 - Q7 posed in Chapter 1 by fulfilling outlined research objectives O8 - O10. Chapter 7 is an effort in the direction of achieving research objective O8. The advancements in Chapters 8 and 9 form the basis of meeting the research objectives O9 and O10.

Chapter 7

Standard Model Order Reduction for (non-linear) hyperbolic problems

In this chapter, the main goal is to develop reduced-order representations of the two-phase flow models in the scope of reducing the complexity of (conditionally) hyperbolic Managed Pressure Drilling (MPD) systems, in particular under gas influx. We first lay the mathematical foundation for model order reduction of linear and non-linear dynamical systems by revisiting existing basic principles of projection-based approaches. The standard (Galerkin-type) projection-based approach is then applied to the (non-linear) Burgers' equation for insights into hyperbolic problems, and to the Drift Flux Model in the scope of reducing the two-phase flow hydraulic systems. The numerical performance is assessed on some challenging test cases. These numerical experiments put forward the success and limitations of standard projection-based model order reduction techniques.

7.1   Introduction

In today's world, it is often desired to predict the system response of an underlying mathematical model in real-time, or to perform multi-query simulations. On the one hand, multi-query scenarios, as in Uncertainty Quantification, require to solve the model multiple times for different problem settings/parame-

ters. These parameters may represent initial conditions, boundary conditions, sources, geometrical configurations or physical properties. As a consequence, optimizing or improving the system performance requires simulations of many possible realizations of the system. On the other hand, the solution of the system may need to be predicted at a low computational time for several time-critical applications. For instance, in the scope of Managed Pressure drilling (MPD), low-complexity models are needed as a basis for controller design.

Most of the physical phenomena, including the pressure dynamics for MPD applications, are modelled by using (parametrized) partial differential equations (PDEs), which describe the physics of the problem in a continuous and infinite-dimensional space. We know that a closed form analytical or exact solution is generally not available for a wide class of (parametrized) partial differential equations, as also for MPD. As a consequence, high-fidelity numerical methods are usually employed (also see Chapter 6). However, the usage of high-fidelity numerical techniques to solve the PDEs in (near) real-time or in a multi-query setting is often not feasible due to high computational time and storage constraints required. Hence, classical numerical techniques, i.e., high-fidelity models, are not amenable to (near) real-time estimation and control or multi-query simulations. Model order reduction (MOR) has emerged as an important numerical tool to develop efficient numerical methods for multiscale and multiphysics problems by reducing the dimension of the underlying model. It helps to reduce the computational time in dealing with large dynamical systems, for example, during simulation, control, optimization, parameter estimation, inverse modelling, etc. In general, MOR techniques [147, 159] combine high predictive capacity and low complexity, and greatly reduce the computational cost of complex simulations at a price of a minimal and quantifiable loss of accuracy. This usually allows rapid and accurate evaluation of the solutions to (parametrized) evolutionary PDEs. It is worth mentioning that a reduced-order model is not a low-fidelity model since it does not consider reduced physics, but reduced representations of the full problem.

Reduced-order models can be constructed via two approaches. One of them is a classical approach based on the mathematical manipulation of the system equations, while the other approach is data driven. Both approaches focus to retain significant dynamics and preserve key system properties (e.g., the input/output properties, stability, etc.). Depending on the model under consideration and the a-priori information about the physical process, the suitable approach is chosen for constructing a compact reduced-order model in an automated sense. The performance of a reduced-order modelling technique can be quantified in terms of computational speed up, computational accuracy, long

time simulation performance, behaviour under parametric setting, and generic (and automated) applicability of the method for a large class of problems. In this dissertation, we intend to efficiently reduce the dimensionality of non-linear mathematical models. In particular, we intend to reduce the complexity of the MPD system to enable multi-query simulations and to predict the quantities of interest, say pressure, in (near) real-time. We recall that MPD is composed of conditionally hyperbolic mathematical models. Hence, we are interested to reduce the dimensionality of non-linear (conditional) hyperbolic PDEs. Such PDEs are ubiquitous in science and engineering. Several applications encompassing the fields of chemical industry, nuclear industry, drilling industry, etc., also fall within this class.

Projection-based methods are usually used for model-based reduction. In particular, Proper Orthogonal Decomposition (POD) [25, 195]; also see Section 7.4, Reduced Basis Method (RBM) [22, 44], Balanced Truncation methods [15, 193], Krylov based methods [159], Proper Generalized Decomposition [45], etc., have been used by many researchers to reduce the dimensionality of the high-fidelity model. However, standard projection-based methods, such as the POD or the RBM, standalone, are ineffective in case of non-linear problems as they require a projection of the reduced state variables back to the high-dimensional state space in order to evaluate the non-linear terms [44, 52]. In the past, the Discrete Empirical Interpolation method, popularly called as DEIM (a discrete variant of the Empirical Interpolation Method or EIM) [44] and the Empirical Operator Interpolation [52] have been proposed to overcome this limitation in the POD, its variants and other projection-based approaches. Recently, Generalized Empirical Interpolation Method (GEIM) [22, 78], Localized Discrete Empirical Interpolation Methods (LDEIM) [142], etc., have also been proposed to advance the field of non-linear MOR.

In the scope of reducing the complexity of MPD systems under gas influx, single- and/or two-phase flow models need to be reduced. While some work has been done in the past to reduce single-phase flow-type models [8, 112], no work, to the best of our knowledge, has yet been carried out to construct a reduced-order counterpart of the the Two-Fluid Model and the Drift Flux Model (DFM). Hence, the central objective of this chapter is to reduce the two-phase flow models of interest by applying existing standard projection-based approaches. In particular, we aim to assess the performance of the standard projection-based approaches in reducing the dimensionality of the DFM. Such an assessment will help to answer whether we need to investigate and propose efficient, advanced and automated approaches to obtain reduction (of the DFM) of the desired order, which is suitable for (real-time) control methodologies, while still guar-

anteeing the accurate approximation of wave propagation phenomena.

The outline of this chapter is as follows. In Section 7.2, we discuss the generic framework popularly used for performing model order reduction. In Section 7.3, we introduce the notion of Kolmogorov N-width to shed insights on the reducibility of the high-fidelity model. Section 7.4 is devoted to the principles of constructing a reduced-order model using the POD and Galerkin projection. Section 7.5 is dedicated to perform several numerical case studies pertaining to the reduction of the DFM. In this section, we also apply standard projection-based approaches on Burgers' equation. Finally, Section 7.6 ends this chapter with conclusions.

7.2   Generic Framework of Model Order Reduction

We are interested in finding the solution (or the quantity of interest) $u(x, t, \mu) \in \mathscr{L}^2(\Omega)$ for $x \in \Omega$, $t \in [0, T]$, and parameter $\mu \in \mathscr{D} \subset \mathbb{R}^d$ of a general non-linear (hyperbolic) problem of the form:

$$\partial_t u(x, t, \mu) + L_\mu \circ u(x, t, \mu) = 0, \quad (x, t, \mu) \in \Omega \times [0, T] \times \mathscr{D}, \qquad (7.1)$$

with suitable initial and boundary conditions. Here, Ω stands for the spatial domain of interest, $\mathscr{D}$ stands for the parameter domain of the problem, L_μ indicates the differential operator with the subscript μ depicting the parametric dependency, and T indicates the final simulation time.

We assume that we employ classical numerical discretization techniques, namely Finite Volume Methods, to obtain an approximation which lives in a high-dimensional functional space. This approximation, denoted as the "truth" solution of (7.1), is given by $u_n(x, t, \mu) \in W_n$ for a parameter $\mu \in \mathscr{D}$. Here, W_n stands for the high-dimensional subspace where the high-fidelity "truth" solution resides, and the subscript n indicates the dimension of the high-fidelity model. A low-dimensional subspace (comprising POD/RBM ansatz), $W_r \subset W_n$, is defined and used for projecting the "truth" solution. The POD/RBM solution is denoted by $u_r(x, t, \mu) \in W_r$, where r is the dimension of the reduced-order model. The major requirements from the reduced-order framework are: *(i)* Reliable and efficient error control, with the error defined as $||u_n(x, t, \mu) - u_r(x, t, \mu)||$, *(ii)* Choice of W_r, *(iii)* Offline - Online decomposition and *(iv)* Empirical Operator Interpolation [52] to treat non-affine and non-linear problems. The space W_r is expected to be of significantly smaller dimension than W_n. This significant reduction depends on the choice of the training samples (in the parameter domain) and the inherent system dynamics.

The main idea of projection-based MOR is often based on the offline and online decomposition. The methodology requires that all system operators, matrices and vectors, coming from the discretization of the problem, must be affinely dependent so that many complex computations can be allocated in a so-called (computationally intensive) offline phase and a rapid solution (time- or parameter-dependent) can be computed efficiently in an online phase. Moreover, a reduced basis, consisting of suitable solution samples, which demand heavy computational cost in the "truth" space, is precomputed in the expensive offline phase. Hence, the computational complexity of the offline phase depends on the high-dimensional "truth" space. Overall, it is favorable to pay the price of an expensive offline phase for real-time or multi-query online evaluations. The methodology pays off when a large number of parametric evaluations are required. High-fidelity solutions, which are computed in the offline stage, serve as basis functions for the Galerkin projection, which is performed in the online phase. The computational complexity of the online phase is independent of the dimension of the space in which the high-fidelity solutions reside.

The difference between the projection-based reduced-basis approximation and the exact solution comprises of two different kind of errors incurred during two separate approximation phases. The original (analytical) problem, defined by PDEs, is posed on some infinite-dimensional functional spaces. Implementation of projection-based approaches requires a finite-dimensional space. The first approximation phase comprises of numerical approximation of an infinite-dimensional model. This numerical approximation poses the original continuous problem in a discrete sense such that the solution lives on a (high) finite-dimensional space. The error committed in this approximation phase is due to the choice of the high-fidelity numerical solver. The second approximation phase aims at building a reduced-order representation of the high-dimensional space-time-parametric manifold. The error in this phase solely depends on how well a reduced-order model captures the solution in the "truth" space. The total accrued error should be within an acceptable tolerance, which in general depends on the kind of application one is interested in.

POD forms a reduced basis by linearly combining the snapshots corresponding to different time instants and/or parametric values. In contrast, the RBM selects a time instant and/or a value of the parameter in an iterative manner using an error estimator (greedy algorithm [146]) and then uses the span of the corresponding snapshot to build the reduced basis. Both the POD and the RBM lack the ability to treat a non-linear and a non-affine problem efficiently. Furthermore, both the POD and the RBM are accurate in representing the data used to generate them. POD/RBM also work for parameters they have not been

trained for as long as the span of training parameter solutions approximates the entire solution manifold of interest well enough.

Traditional projection-based methods, such as the POD and the RBM, possess good approximation properties for fast decay of Kolmogorov N-width (see Section 7.3 for a discussion on Kolmogorov N-width). The downside of these projection-based methods is (problem dependent) enormous offline complexity and that they fail for the slow decay of Kolmogorov N-width. To sum up, the success of these projection-based methods depends on the regularity (in the sense of continuity and differentiability) of the solution map.

7.3   Kolmogorov N-width

The notion of Kolmogorov N-width is a classical concept in approximation theory. It quantifies how well a compact set, in some high-dimensional functional space (say, a Banach space), can be approximated in the corresponding norm by a linear space and, therefore, can be used as a benchmark for the effectiveness of a model reduction strategy [144]. Kolmogorov N-width can be explained as follows.

Let X be a normed linear space, S be a subset of X, and X_N be some N-dimensional subspace of X. The deviation of S from X_N is :

$$E(S; X_N) = \sup_{u \in S} \ \inf_{v_N \in X_N} ||u - v_N||_X. \qquad (7.2)$$

The Kolmogorov N-width of S in X_N is given by:

$$d_N(S, X_N) = \inf_{X_N} \ \sup_{u \in S} \ \inf_{v_N \in X_N} ||u - v_N||_X. \qquad (7.3)$$

This quantity measures the extent to which S may be approximated by a finite-dimensional subspace. The decay behaviour of d_N indicates the number of POD or RBM modes that need to be considered to develop a reduced-order model. In other words, the faster the decay, the lower is the required dimension of the linear subspace.

It is well known that the Kolmogorov N-width decays exponentially fast for certain linear and coercive parametrized problems [30, 135]. This is generally the case for (parametrized) elliptic and parabolic partial differential equations. In the scope of hyperbolic partial differential equations, as in advection equation, linear transport problem, etc., the Kolmogorov N-width decays

slowly [77, 135]. Clearly, the linear reducibility in the scope of transport-dominated problems is hindered. It should also be mentioned that this hindrance is not limited to specific discontinuous initial conditions, but also holds for (smooth) continuous initial conditions.

For instance, consider the one-dimensional linear advection equation:

$$\partial_t u(x,t) + c\partial_x u(x,u) = 0, \quad (x,t) \in \Omega \times [0,T],$$
$$u(x,0) = u_0(x),$$

(7.4)

where c is a constant speed of propagation and $u_0(x)$ is the initial condition.

The general solution of such one-dimensional linear transport equation is well-known, and is given by:

$$u(x,t) = u_0(x - ct).$$

(7.5)

Consider N_s snapshots of the solution (7.5) at time instants $t = 0, \Delta t, 2\Delta t, \ldots, (N_s - 2)\Delta t, T$, with $T = (N_s - 1)\Delta t$ and construct a snapshot matrix. Further consider a non-smooth initial data that is localized with compact support. Then, it is easy to see that the singular values of the snapshot matrix do not decay at all. Hence, projection-based methods will not be of any use in reducing the dimensionality of the linear advection equation, which is a simplistic example that is representative of a travelling wave phenomena.

At this stage, we would also like to recall that the two-phase flow models of interest possess such wave-propagation phenomena; see Chapter 6.

7.4   Projection-Based Approaches

While several projection-based approaches exist, we will only focus on one of them, namely the Proper Orthogonal Decomposition (POD). The POD is representative enough of the performance of projection-based approaches in the scope of dealing with the MOR of hyperbolic PDEs or transport-dominated problems. In this section, we first discuss about the POD, and then we delve into the Galerkin projection methodology. In unison, these two core concepts help to lay the foundation of the POD-Galerkin methodology.

Proper Orthogonal Decomposition
The POD is a well-known and mostly used MOR technique. It is known in the scientific community by several names, such as Principal Component Analysis (PCA), Karhunen Loeve Decomposition and method of empirical orthogonal

eigenfunctions [25, 195]. This method was introduced by Lumley to study turbulence. The POD offers orders of magnitude reduction in dimensionality for a wide range of applications such as in image processing, signal processing, data compression, oceanography, fluid mechanics, optimal control, and many others. Several POD variants, such as Balanced Proper Orthogonal Decomposition [153], have also been proposed by several researchers. In this chapter, we will not delve into any variants of the POD and will restrict ourselves to the consideration of the standard POD only.

The POD is a snapshot-based MOR technique and is typically used for time-dependent problems. This technique can be viewed as a mathematical tool to calculate the basis vectors from an ensemble of realizations of a high-fidelity numerical simulation. The POD looks for a linear space, which captures the dominant dynamics of the problem and identifies the most energetic modes in a dynamical system. The main idea behind this method is to represent a given (large) data set by a smaller orthonormal basis which is optimal in a least-squares sense, which is mathematically defined next.

Consider the snapshot matrix $Y = \{y_1(t) \ y_2(t) \ \dots \ y_m(t)\}$ as the trajectory of the states of the system over the time interval $[0, T]$, where $y_i(t) \in \mathbb{R}^n$ for $i = 1, 2, \dots, m$. Here, n is indicative of the number of finite-elements or finite-volume cells of the discretized computational domain and m indicates the number of time samples in $[0, T]$. Consider the least squares distance defined as follows:

$$\| Y - P_{W_r} Y \|_{\mathbb{R}^n}^2 := \sum_{j=1}^{m} \| y_j(t) - P_{W_r} y_j(t) \|_{\mathbb{R}^n}^2, \tag{7.6}$$

where $P_{W_r} : W_n \to W_r$ is the orthogonal projection operator onto W_r. The POD seeks to find P_{W_r} such that the least squares distance, defined in (7.6), is minimized in the following sense:

$$\min_{\tilde{\psi}_1, \dots, \tilde{\psi}_r \in \mathbb{R}^n} \sum_{j=1}^{m} \| y_j - \sum_{i=1}^{r} \langle y_j, \tilde{\psi}_i \rangle_{\mathbb{R}^n} \tilde{\psi}_i \|_{\mathbb{R}^n}^2 \quad s.t. \ \langle \tilde{\psi}_i, \tilde{\psi}_j \rangle_{\mathbb{R}^n} = \delta_{ij} \ \forall \ 1 \le i, j \le r, \tag{7.7}$$

where δ_{ij} is a Kronecker symbol that satisfies $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$. The minimization problem governed by (7.7) can be reformulated as:

$$\max_{\tilde{\psi}_1, \dots, \tilde{\psi}_r \in \mathbb{R}^n} \sum_{i=1}^{r} \sum_{j=1}^{m} | \langle y_j, \tilde{\psi}_i \rangle_{\mathbb{R}^n} |^2 \quad s.t. \ \langle \tilde{\psi}_i, \tilde{\psi}_j \rangle_{\mathbb{R}^n} = \delta_{ij} \ \forall \ 1 \le i, j \le r. \tag{7.8}$$

For any $r \in \{1, \dots, d\}$ with $d = \min(n, m)$, the solution to (7.7) or (7.8) is known to be given by the first r columns of ψ (see Theorem 1.1, [196]), i.e.,

$$P_{W_r} = [\psi_1 \ \psi_2 \ \dots \ \psi_{r-1} \ \psi_r], \tag{7.9}$$

where ψ is the orthogonal matrix obtained by performing singular value decomposition (SVD) on the snapshot matrix Y which is given by $Y = \psi \Sigma \phi^T$. Here, $\psi = [\psi_1 \ \ldots \ \psi_n]$, which is composed of the left singular vectors, is a $n \times n$ orthogonal matrix, $\phi = [\phi_1 \ \ldots \ \phi_m]$, which is composed of the right singular vectors, is a $m \times m$ orthogonal matrix, and Σ is a $n \times m$ diagonal matrix that contains the information of the singular values. Each column of ψ is a mode structure ordered from the most important to the least important and the columns of ϕ demonstrate the evolution of the modes with time. The vectors ψ_i (as in (7.9)), where $i = 1, 2, \ldots, r$, are referred to as the POD modes in the sequel.

The choice of r is of central importance while applying the POD. In general, no a-priori rules are available for the choice of r. The choice of r could be based on heuristic considerations. For instance, the information content of the bases, given by:

$$E(r) = \frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{d} \lambda_i}, \tag{7.10}$$

could be used as a measure to determine the dimension of the reduced-order model. The optimal choice of r is not the central theme in the scope of this dissertation. We will use the decay behaviour of the singular values to identify the (near-optimal/acceptable) dimension of the reduced-order model.

One of the important properties of the POD is that

$$\sum_{i=1}^{r} \sum_{j=1}^{m} |\langle y_j, \tilde{\psi}_i \rangle_{\mathbb{R}^n}|^2 = \sum_{i=1}^{r} \sigma_i^2, \tag{7.11}$$

where σ_i is the singular value obtained after performing SVD on the snapshot matrix Y. Using this property, the error in the computation of $\| Y - P_{W_r} Y \|_{\mathbb{R}^n}^2$, due to truncation, can be shown to be equal to sum of the squares of the singular values corresponding to the neglected POD modes; see [196].

In the above discussion, the SVD approach has been used to find the solution to the optimization problem (7.7) or (7.8). The approach of eigenvalue decomposition can also be utilized to find the POD modes or the POD basis. The eigenvalue decomposition-based approach is usually referred to as the method of snapshots in the literature [196]. Furthermore, the aforementioned discussion on the POD method delved into finding the POD modes based on the natural inner product on the Euclidean space. In principle, the POD modes can be computed with respect to a weighted inner product. For details, we refer the reader to [196] for the POD method with a weighted inner product. Without delving into rigorous technical details, we now present the algorithmic framework for the computation of the POD basis via the approaches that are also

utilized in this chapter and the subsequent chapters of this dissertation. The following algorithm is a mere adaptation of the algorithm proposed in [196]:

Data: Snapshot matrix $Y \in \mathbb{R}^{m \times n}$, reduced-dimension r ($\leq d$), symmetric and positive definite weight matrix $W \in \mathbb{R}^{m \times m}$, flag and solver;

Result: POD modes ψ_i;

if *flag==0* **then**

    Singular Value Decomposition based approach

    Determine $\bar{Y} = W^{1/2} Y$;

    Compute singular value decomposition $[\bar{\psi}, \Sigma, \bar{\phi}] = \text{svd}(\bar{Y})$;

    Set $\psi_i = W^{-1/2} \bar{\psi}_{.,i}$ for $i = 1, 2, \ldots, r$;

else

    Eigenvalue Decomposition based approach

    if *solver==1* **then**

        Determine $\bar{Y} = W^{1/2} Y$;

        Set $R = \bar{Y} \bar{Y}^T$;

        Compute eigenvalue decomposition $[\tilde{\psi}, M] = \text{eig}(R)$;

        Set $\psi_i = W^{-1/2} \tilde{\psi}_{.,i}$ for $i = 1, 2, \ldots, r$;

    else

        Determine $K = Y^T W Y$;

        Compute eigenvalue decomposition $[\tilde{\phi}, N] = \text{eig}(K)$;

        Set $\psi_i = Y \tilde{\phi}_{.,i} / \sqrt{\lambda_i}$ and $\lambda_i = N_{ii}$ for $i = 1, 2, \ldots, r$;

    end

end

Algorithm 1: Basis computation via Proper Orthogonal Decomposition.

To summarize, the basis (also called POD modes) is constructed via the singular value decomposition based approach if the counter "flag" is equal to zero, and via the eigenvalue decomposition based approach if the counter "flag" is not equal to zero. And, the counter "solver" is equal to one if $n < m$, and is not equal to one for $n > m$. Furthermore, the weight matrix W, which is an input to the algorithm, is known and can be computed depending on the inner product deemed appropriate by the user in the scope of the problem of interest. If the natural inner product is considered, the weight matrix is given by an identity matrix $I_{m \times m}$ of size $m \times m$. If the POD modes are computed with respect to an $\mathcal{L}^2$ inner product, then the weight matrix W is given by: $W = \Delta x \times I_{m \times m}$, where Δx is the spatial step size. In case the POD modes are computed with respect to an $\mathcal{H}^1$ inner product, the weight matrix W is composed of the so-called mass matrix and stiffness matrix. The mass matrix is given by $\Delta x \times I_{m \times m}$ and the stiffness matrix is computed by performing spatial discretization on the operator

$-\frac{\partial^2}{\partial x^2}(\cdot)$. It should be noted that the form of the mass and the stiffness matrices are independent of the nature of the particular problem under consideration. Algorithm 1 can be employed to compute the POD modes by either pursuing a singular value decomposition or an eigenvalue decomposition based approach.

Remark 7.4.1 *Usually, pursuing either a singular value decomposition based approach or an eigenvalue decomposition based approach yields the same results. However, some numerical differences may be observed depending on the size and the nature of the matrices, and the ability of such solvers in the programming language (for, e.g., MATLAB) employed to perform the computations.*

Having constructed the POD modes, we are now ready to project the high-dimensional system of equations to construct a reduced-order model by employing a (Petrov) Galerkin projection technique.

Remark 7.4.2 *Remark 7.4.3 We focus on reproducing the results of the time-dependent equation via the standard MOR framework. This reproduction step is essential before attempting to develop a parametric reduced-order model as we cannot hope to have an effective low-complexity reduced-order model if the numerical approach does not fare well in the reproduction step.*

Galerkin Projection

We assume that the PDE is not parameter dependent, and, as a consequence, (7.1) can be re-written as:

$$\partial_t u(x, t) + L \circ u(x, t) = 0, \quad (x, t) \in \Omega \times [0, T]. \tag{7.12}$$

Consider that the solution is approximated by a linear combination of the POD modes as follows:

$$u \approx P_{W_r}^T u_{red}(t). \tag{7.13}$$

If we substitute the approximation (7.13) into (7.12), then we obtain:

$$\partial_t (P_{W_r}^T u_{red}) + L \circ (P_{W_r}^T u_{red}) + \underbrace{R}_{residual} = 0, \tag{7.14}$$

where the (additional) residual term R is introduced to take account of the errors incurred due to the approximation considered in (7.13), and the argument of u_{red} has been omitted for the sake of simplicity. Invoking that the projector P_{W_r} is time-independent, (7.14) can be re-written as follows:

$$P_{W_r}^T \partial_t (u_{red}) + L \circ (P_{W_r}^T u_{red}) + \underbrace{R}_{residual} = 0. \tag{7.15}$$

The unknowns in (7.15) are $u_{red}(t)$ and R. We now formulate the problem in a way that we eliminate the role of (unknown) R and that we have to only find $u_{red}(t)$. To this end, we constrain the residual R to be orthogonal to a subspace $\mathcal{V}_r$ defined by a test projector $P_{V_r}^T$, i.e.,

$$P_{V_r} R = 0. \tag{7.16}$$

Left multiplying (7.15) by P_{V_r}, we get:

$$P_{V_r} P_{W_r}^T \partial_t(u_{red}) + P_{V_r} L \circ (P_{W_r}^T u_{red}) + P_{V_r} R = 0. \tag{7.17}$$

Using (7.16), the third term in (7.17) can be put to zero. We can now write (7.17) as follows:

$$P_{V_r} P_{W_r}^T \partial_t(u_{red}) + P_{V_r} L \circ (P_{W_r}^T u_{red}) = 0. \tag{7.18}$$

A projection methodology where $P_{V_r} \neq P_{W_r}$ is usually referred to as a Petrov-Galerkin projection technique, and a methodology where $P_{V_r} = P_{W_r}$ is referred to as a Galerkin projection technique. Employing any of these two projection methodologies could be pursued in principle, however, we restrict our study to the consideration of Galerkin projection technique in the scope of this dissertation. Invoking $P_{V_r} = P_{W_r}$, (7.18) can be re-written as follows:

$$P_{W_r} P_{W_r}^T \partial_t(u_{red}) + P_{W_r} L \circ (P_{W_r}^T u_{red}) = 0. \tag{7.19}$$

If P_{W_r} is composed of orthonormal modes, we can say that $P_{W_r} P_{W_r}^T = I$. Hence, (7.19) can be simplified to:

$$\partial_t(u_{red}) + P_{W_r} L \circ (P_{W_r}^T u_{red}) = 0. \tag{7.20}$$

Remark 7.4.4 *The basis computed by using the method of snapshots, as elaborated in Algorithm 1, may not be necessarily orthonormal due to limited numerical accuracy. In order to have the property $P_{W_r}^T P_{W_r} = I$, the modes computed from Algorithm 1 should be orthonormalized. Such orthonormalization can be done by using modified Gram-Schmidt re-orthogonalization. Generally, doing such an orthogonalization twice is enough.*

Using the Galerkin projection methodology, we have converted the high-fidelity model (7.12) to a reduced-order model (7.20), where the unknowns u_{red} are the time-dependent coefficients associated to the POD modes.

The whole procedure of the POD-Galerkin approach can be split into an offline phase and an online phase. The offline phase is dedicated to the generation of the POD basis functions and all the operators (time/parameter-independent), and the online phase involves the computation of time-dependent coefficients.

Remark 7.4.5 *Despite the nice offline and online spliting of the computations, the POD-Galerkin method is not computationally appealing for non-linear problems of scientific interest. For instance, in (7.20), the operator L acts on the quantity $P_{W_r} u_{red}$ whose dimension is equivalent to that of the high-fidelity model. This leads to the reduced-order models with similar computational costs as the original high-fidelity models and, therefore, no significant cost savings are achieved. It is clear that the projection alone is not sufficient to reduce the costs of computing a reduced-order model. Hence, an Empirical Operator Interpolation [52], a hyper-reduction recipe, needs to be employed in conjunction with the POD-Galerkin method. Although such a hyper-reduction recipe is also crucial in the scope of developing reduced-order models of MPD systems, we will only employ the POD-Galerkin method to put forward key challenges to reduce MPD systems. The challenges (see Section 7.5) are not attributed to the missing hyper-reduction recipe, but are associated to a fundamental barrier of projection-based approaches for the treatment of transport-dominated type problems.*

Having laid the mathematical foundation of the POD-Galerkin method, we can now employ it in developing reduced-order models for MPD systems, in particular the DFM which is employed for hydraulic modelling in the annulus part of the wellbore. The methodology can also be applied to develop a reduced-order model of Burgers' equation.

7.5   Numerical Test-Beds

In this section, we numerically test the POD-Galerkin method and show its performance. We focus on the problems in the field of fluid dynamics. A series of numerical experiments are presented to assess the accuracy of the reduced-order model and, thus, illustrate the advantages or disadvantages of the (standard) projection-based MOR framework. Firstly, we model reduce Burgers' equation, which is a well-known benchmark model in the field of fluid dynamics. Then, the behaviour of the reduced-order model is assessed using the DFM. We restrict to numerical test cases in one-dimensional setting and also use a uniform discretization across space unless mentioned otherwise. The initial data and boundary conditions are also specified for the completeness of the problem. We will employ the Finite Volume (FV) method for the discretization of the problem at hand.

Before delving into the numerical assessment, we introduce the performance metric to assess the ability of the POD-Galerkin method (and MOR techniques in

general). We quantify the performance of any MOR approach by computing the reduced-order modelling (ROM) error. The considered metrics are discussed next.

- We consider $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time (absolute) ROM error and define it in the following manner (for a reduced-dimension r):

$$e_{rom} = \sqrt{\Delta t \sum_{k=1}^{N_T+1} \Delta x \sum_{i=1}^{N_x} | u_{i,k} - (P_{W_r}^T u_{red})_{i,k} |^2}, \qquad (7.21)$$

where Δt is the time-step, Δx is the spatial step, N_T is the number of time-steps and N_x is the number of Finite Volume elements. And, $u_{i,k}$ stands for u at $x = x_i$ and $t = t_k$ (similarly for $(P_{W_r}^T u_{red})_{i,k}$). Herewith, (7.21) expresses the error which is the difference between the FV-based numerical solution u and the reconstructed solution $P_{W_r}^T u_{red}$ obtained by lifting the standard reduced-order solution u_{red} to the high-dimensional problem space.

- We consider $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time (relative) ROM error and define it as follows:

$$e_{rom}^{rel(\mathscr{L}^2, \mathscr{L}^2)} = \frac{\sqrt{\Delta t \sum_{k=1}^{N_T+1} \Delta x \sum_{i=1}^{N_x} | u_{i,k} - (P_{W_r}^T u_{red})_{i,k} |^2}}{\sqrt{\Delta t \sum_{k=1}^{N_T+1} \Delta x \sum_{i=1}^{N_x} | u_{i,k} |^2}}, \qquad (7.22)$$

where the meaning of the notations is as explained earlier.

- We consider $\mathscr{L}^2$ in space and $\mathscr{L}^\infty$ in time (relative) ROM error, and define it as follows:

$$e_{rom}^{rel(\mathscr{L}^2, \mathscr{L}^\infty)} = \frac{\max\limits_{k=\{1,2,\dots,N_T+1\}} \left(\Delta x \sum_{i=1}^{N_x} | u_{i,k} - (P_{W_r}^T u_{red})_{i,k} |^2 \right)}{\max\limits_{k=\{1,2,\dots,N_T+1\}} \left(\Delta x \sum_{i=1}^{N_x} | u_{i,k} |^2 \right)}. \qquad (7.23)$$

7.5.1 Burgers' Equation

There are several reasons which motivate the study of Burgers' Equation. Firstly, it is a simple analog of the Euler equations and the two-phase flow models, and it mimics the non-linear wave equation where each point on the wave-front can propagate with different speed. Secondly, it possesses coalescence

of characteristics and the formation of discontinuous solutions similar to shock waves/rarefaction waves in general fluid mechanics problem. The mathematical representation of Burgers' equation is given below:

$$\begin{cases} y_t + (\frac{1}{2}y^2)_x = 0, & (x,t) \in [0,L] \times [0,T], \\ y(t,0) = y(t,L), & t \in [0,T], \\ y(0,x) = y_0(x), & x \in [0,L]. \end{cases} \tag{7.24}$$

The model in (7.24) considers (constant) periodic boundary conditions and smooth or non-smooth initial data dictated by the profile of $y_0(x)$. Furthermore, we consider the spatial domain to be of length $L = 12$. The spatial domain is discretized using a spatial step size of 0.002. We use an upwind FV scheme for the spatial discretization and first-order Forward Euler for the time-stepping. We take 8000 steps in time, i.e., $t \in [0,4]$ with a timestep of 0.0005.

Smooth initial data

It is well-known that even if we start with smooth initial data, Burgers' equation tends to develop sharp gradient (discontinuous) features. Consider that the initial data $y_0(x)$ is given by a (smooth) Gaussian profile as shown below:

$$y_0(x) = 2e^{-\frac{(x-5)^2}{1e^{-1}}}. \tag{7.25}$$

This initial data leads to the appearance of a shock after a finite time $t < 4$ and, hence, the travelling discontinuous feature is also present in the response of the solution, which further challenges the projection-based MOR approaches. Furthermore, we assume $g = 0$, i.e., zero periodic boundary conditions are considered.

We perform high-fidelity FV computations on the model (7.24) and obtain a snapshot matrix. In the pursuit of developing a reduced-order model, we then opt to use the SVD based approach for the computation of the basis functions and, hence, we set the "flag", in Algorithm 1, to zero. We compute the POD modes by employing the POD both with respect to an $\mathcal{L}^2$ inner product and an $\mathcal{H}^1$ inner product, and then employ a Galerkin projection methodology. Figure 7.1 depicts the ROM error, defined in (7.22), versus the number of the POD modes for the two mentioned different ways of basis computations. We can clearly see that the decay behaviour of the ROM error is not so steep. The error decay is gradual owing to the travelling nature of the problem and also due to the (significant) changes in the shape profile of the solution, which changes
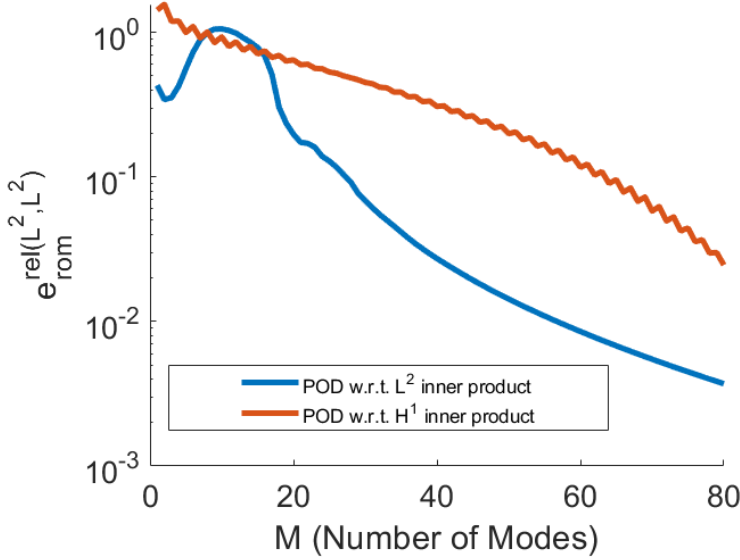
Figure 7.1: ROM error for the case with smooth initial data.

from a Gaussian to a triangular shape (not depicted here). It is observed that the ROM errors, for the POD method with respect to an $\mathscr{L}^2$ inner product, could even increase with an increment in the basis size. It can be argued that this could occur as a result of insufficiently many basis functions. Furthermore, the ROM error for the POD with respect to an $\mathscr{L}^2$ inner product shows a better decay behaviour than the POD with respect to an $\mathscr{H}^1$ inner product. From the ROM error decay for the case where the POD modes are computed with respect to an $\mathscr{H}^1$ inner product, it is evident that the dimension of the reduced-order model needs to be relatively high even to obtain a 1 percent relative error.

Non-smooth initial data

We now consider an initial data of the following form:

$$y_0(x) = \begin{cases} x - 2, & 2 \le x \le 4, \\ \frac{(x-5)}{2}, & 5 \le x \le 7, \\ 0, & \text{otherwise.} \end{cases} \tag{7.26}$$
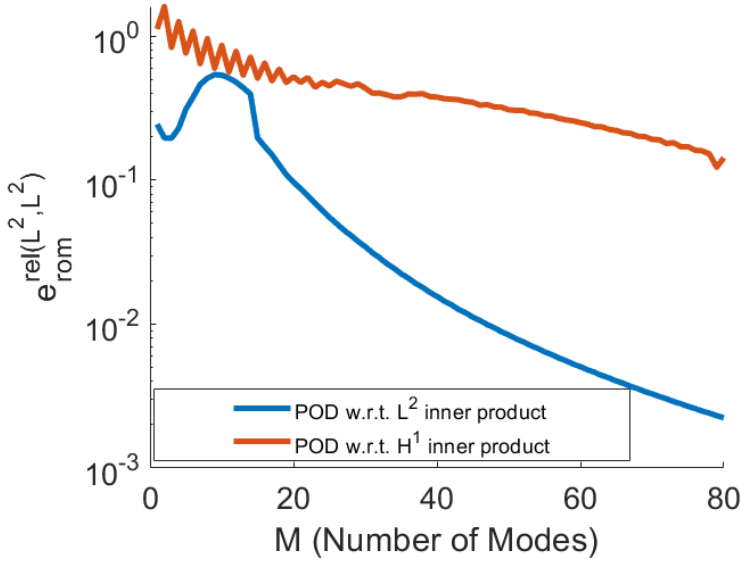
Figure 7.2: ROM error for the case with non-smooth initial data.

 The rest of the settings are kept similar to the numerical test with the smooth initial data. For the setting under consideration, Figure 7.2 depicts the ROM error, defined in (7.22), versus the number of the POD modes for the two mentioned different ways of basis computations. Again, the decay of the ROM error with the increasing number of the POD modes is quite gradual. The qualitative behaviour of the ROM error upon employing the POD with respect to an $\mathcal{L}^2$ inner product seems similar to the corresponding behaviour observed for the case with the smooth initial data. However, such an observation, in general, does not hold and happens to be a sheer coincidence. Furthermore, if we observe the behaviour of the ROM error for the POD method with respect to an $\mathcal{H}^1$ inner product, we can see that it is significantly different than the corresponding behaviour with the smooth initial data. The behaviour of the error decay for the POD method with respect to an $\mathcal{H}^1$ inner product is much worse, and this is potentially owing to the presence of two travelling discontinuous features and the interaction of the head of the left wavefront with the tail of the right wavefront. To conclude, the possible reduction in the dimension of the model does not turn

out to be significant enough if the standard projection-based approaches are employed.

7.5.2   Drift Flux Model

Many flow and transport models can be described by a set of PDEs representing conservation of mass, momentum and energy as a function of pressure and temperature, and reconciled by the equations of state. The DFM, a system of multiscale non-linear hyperbolic PDEs, is one such model. As mentioned in the Part I of this dissertation, the DFM is typically used for hydraulic modelling in the scope of the MPD. In view of real-time predictions and multi-query simulations with such model, we aim to develop a reduced-order model of the DFM. We also recall that capturing wave propagation phenomena, induced by slow or fast transients, is a central challenge both for the high-fidelity and the reduced-order hydraulic models. The DFM, governed by (6.1) and (2.3) - (2.6), exhibits the presence of multiple travelling waves across both directions of the spatial domain. In addition, the model has complexly distributed non-linearities, which are further aggravated due to the numerical discretization reasons, such as the use of the limiters; see Chapter 6 for physical and (high-fidelity) numerical insights about the DFM.

    We use the Flux Vector Splitting (FVS) and the Advection Upstream Splitting (AUSM) scheme for the computation of the numerical fluxes and rely on corresponding first-order computations; see Section 6.3.1 for details about the spatial discretization. For the sake of simplicity, we also employ a first-order Forward Euler scheme to perform the temporal discretization. Since we employ explicit time discretization, any temporal step size cannot be used. The choice of the spatial and the temporal step size is linked via the CFL condition [57]:

$$\Delta t = CFL \frac{\Delta x}{max(|\lambda_1|, |\lambda_2|, |\lambda_3|)}. \tag{7.27}$$

Given the spatial step size Δx and the (constant) CFL, the temporal step size Δt can be computed by using (7.27). Different choices of the spatial step size are considered, and their values will be provided within the respective test cases which we are going to discuss next.

Case 1

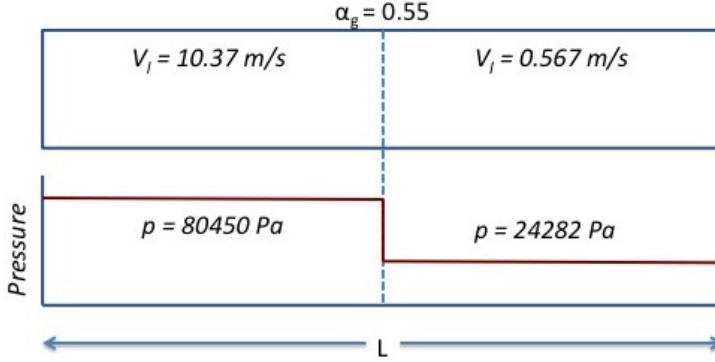We consider a representative multi-phase shock tube test case, inspired from

Figure 7.3: Representative multi-phase shock tube test case.

[57, 90]. The authors in [57, 90] employed this test case in the scope of the high-fidelity numerical computation of the DFM. Similar to [57], the geometry and the initial conditions are depicted in Figure 7.3. The motivation behind the consideration of such shock-tube type test cases has been explained in Section 6.5. We also refer the reader to Case B, Section 6.5.1 for the details regarding the operating conditions of this test case. We recall that the simulation run-time is chosen such that the travelling waves do not reach the end walls of the bounded spatial domain. It can be said that this is a test case where the effect of the boundary conditions do not play any prominent role and, hence, no boundary conditions are specified here. To this end, we perform the numerical simulations over the temporal domain $t \in [0, 1]$. Under the considered setting, the numerical solution of the DFM is composed of a 1−shock, 2−contact discontinuity, and a 3−shock. As a consequence, we will also refer this test case as the S1-S3 case in the sequel.

We employ (only) the FVS scheme for the computation of the numerical fluxes. Furthermore, similar to [57], we use the rough estimate of $max(|\lambda_1|, |\lambda_2|, |\lambda_3|)$ for computing the temporal step size given the spatial step size and the CFL. We consider $max(|\lambda_1|, |\lambda_2|, |\lambda_3|) = 1000$ for computational purposes. We consider CFL = 1 for the computations. Furthermore, we consider different choices of spatial step sizes, i.e., $\Delta x = 0.5$, $\Delta x = 0.25$, $\Delta x = 0.125$ and $\Delta x = 5 \times 10^{-2}$, to obtain insights about the performance of the reduced-order model developed by employing the POD-Galerkin method.

Using the principles laid down in Chapter 6, the high-fidelity computation can be performed. This yields the knowledge of the variables $Q_1 = \alpha_\ell \rho_\ell, Q_2 = $
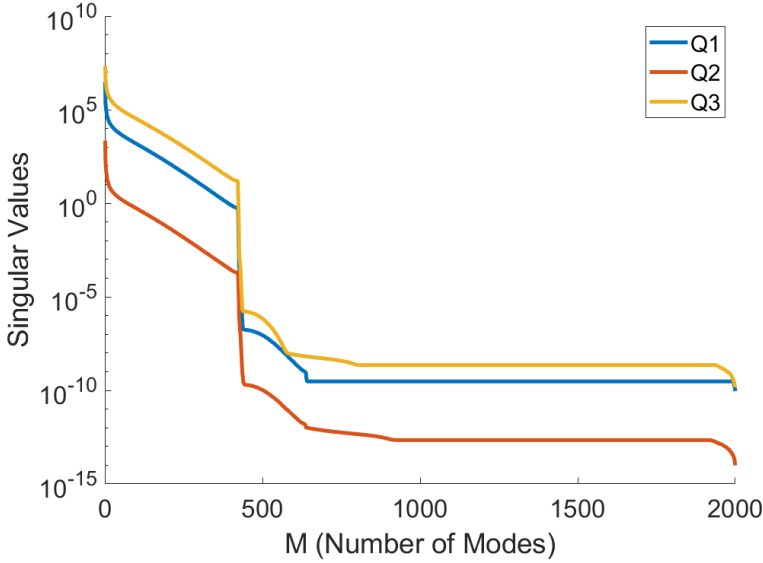
Figure 7.4: behaviour of the decay of the singular values for the transported quantities in the scope of the DFM.

$\alpha_g \rho_g$ and $Q_3 = \alpha_\ell \rho_\ell v_\ell + \alpha_g \rho_g v_g$ over the spatial mesh points for different time instants. Using this, three snapshot matrices, corresponding to each of the three quantities, can be constructed. The SVD can now be performed on each of these snapshot matrices to build the basis (POD modes) during the offline phase of the POD-Galerkin method. Figure 7.4 demonstrates the behaviour of the singular values for the transported quantities Q_1, Q_2 and Q_3 of the DFM. The singular values are useful to determine the extent to which a high-fidelity system can be well approximated by the one of lower dimension. It should be noted that these singular values are not invariant and are a function of the chosen initial condition(s). Moreover, the singular value spectrum is a function of the simulation run-time as the problem under consideration is of transient nature.

Remark 7.5.1 *A practical computation of the SVD may run into several issues if a single SVD is performed on the snapshot matrix, which is composed by the augmentation of Q_1, Q_2 and Q_3. Unknowns, Q_1 (or Q_2) and Q_3 represent different physical quantities and the numerical value of one quantity dominates the other*

*one by several orders of magnitude. In such cases, the contribution of the sub-
dominant variable to the computation of the basis may turn out to be marginal
or almost invisible. The relevant variables may be unnecessarily suppressed with-
out proper scaling. The information from the sub-dominant variables must be
preserved and, hence, SVD is performed separately on each of the three snapshot
matrices Q_1, Q_2 and Q_3. This also holds if the eigenvalue decomposition based
approach is employed for the computation of the POD modes.*

For this test case, we employ the eigenvalue decomposition based approach
for the computation of the POD modes corresponding to each of the three phys-
ical quantities Q_1, Q_2 and Q_3. Furthermore, we consider the POD method with
respect to an $\mathscr{H}^1$ inner product. Based on the number of finite volume cells
and the number of time samples, the "solver", in Algorithm 1, can be identified.
The "solver" is not equal to 1 in the numerical experiments that follow. The
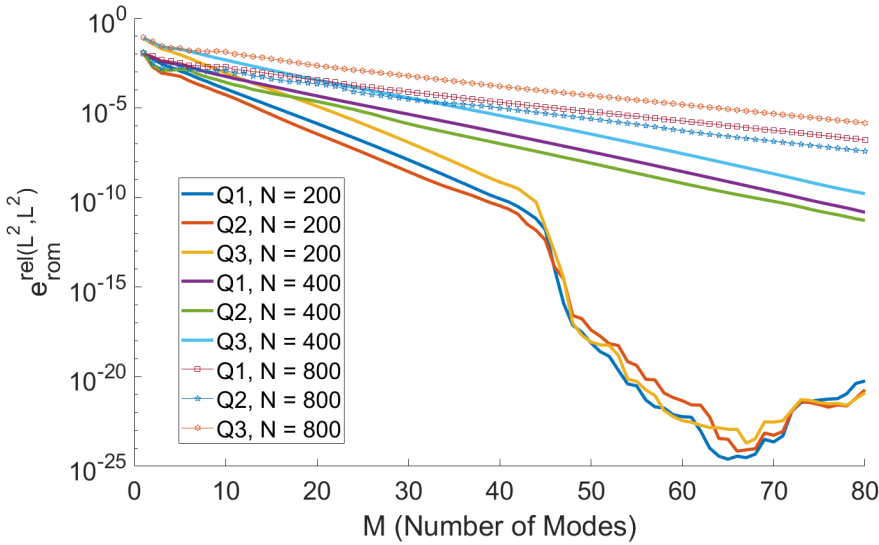implementation of Algorithm 1 returns us the POD modes.

Figure 7.5: S1-S3 Case: behaviour of the ROM error (for Q_1, Q_2 and Q_3) versus the
number of the POD modes for varying number, N, of finite volume cells.
Here, $N = 200$, $N = 400$ and $N = 800$ correspond to the spatial sizes $\Delta x = 0.5$,
$\Delta x = 0.25$ and $\Delta x = 0.125$, respectively.

Figure 7.5 depicts the behaviour of the ROM error, defined in (7.22), versus the number of the POD modes for the three different physical quantities Q_1, Q_2 and Q_3 over different spatial mesh resolutions, i.e., $\Delta x = 0.5$, $\Delta x = 0.25$ and $\Delta x = 0.125$. Figure 7.6 represents the behaviour of the ROM error when the spatial step size $\Delta x = 5 \times 10^{-2}$. A larger spatial step size, for instance $\Delta x = 0.5$, seems to give an impression of an easy reducibility of the problem; see Figure 7.5. However, this is not the case since the choice of $\Delta x = 0.5$ does not correspond to a mesh-independent solution, and the features, such as shocks, are smeared by a large amount due to the coarseness level of the mesh. From Figure 7.5 and 7.6, it can be observed that the finer the mesh, the larger is the incurred relative ROM error. This behaviour is attributed to the fact that the propagating fronts become sharper (and, hence, less smeared) with refining the mesh, and other features are also better resolved. If we observe Figure 7.6 closely, we can see that the ROM error decays steeply for an initial increment in the number of the
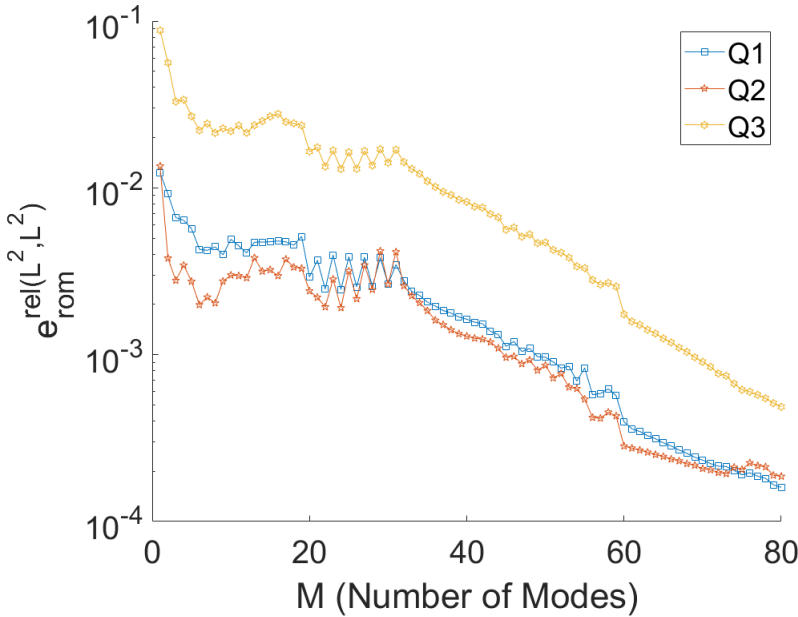

Figure 7.6: S1-S3 Case: behaviour of the ROM error (for Q_1, Q_2 and Q_3) versus the number of the POD modes when the spatial step size $\Delta x = 5 \times 10^{-2}$.

POD modes. This behaviour can be attributed to a large energy content in the (dominant) macroscopic part of the response which can be easily approximated. It is important to mention that the shock-type features, which are localized features, do not necessarily contain a large energy content and, hence, may not always be the most dominant feature. This is already evident from a sharp error decay for the first few POD modes. The ROM error is seen to stagnate around 10^{-2} after the first few POD modes. Such a stagnant behaviour of the ROM error with the increasing number of the POD modes can be due to the fact that the shock front is propagating with a (solution-dependent) speed across the domain. In other words, the travelling nature of the problem becomes dominant and renders the standard projection-based approaches to be ineffective. We already observed such behaviour earlier in the scope of Burgers' equation (with non-smooth initial data) where the ROM error for the POD method with respect to an $\mathcal{H}^1$ inner product depicted a nearly stagnant behaviour with the increasing number of the POD modes; see Figure 7.2. After a stagnation level over a certain number of the POD modes, the ROM error, in Figure 7.6, again shows decay. However, the error decays gradually. This decay could be linked to the sub-macroscopic features in the response of the solution. One observa-
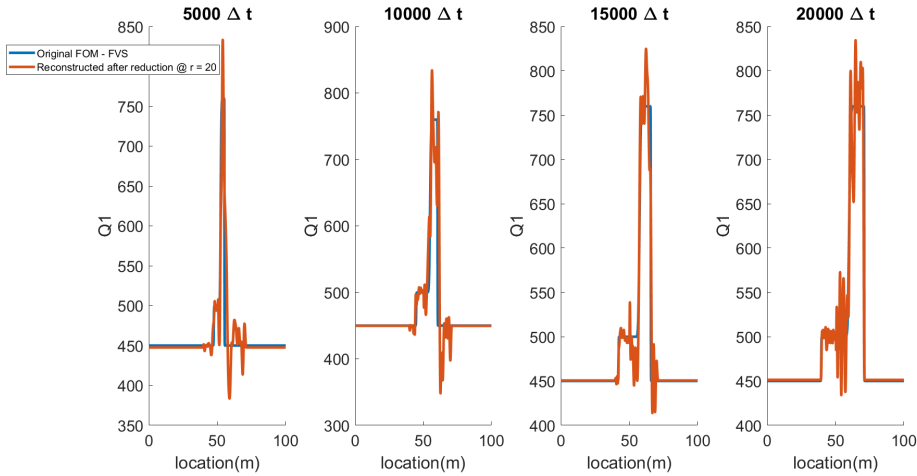

Figure 7.7: Evolution of Q_1 over equidistant time samples. Here, blue colour indicates the high-fidelity solution, and the red colour indicates the (reconstructed) solution of the reduced-order model with reduced dimension $r = 20$.

tion which holds across all the numerical experiments considered so far is that the standard projection-based approaches fail to sufficiently reduce the dimensionality of the high-fidelity model, particularly, due to the transport-dominated effects.

Figure 7.7 demonstrates the evolution of Q_1, obtained from the high-fidelity and the reduced-order model computation, at 4 equidistant time instants, i.e., at $5000\Delta t$, $10000\Delta t$, $15000\Delta t$ and $20000\Delta t$ with $\Delta x = 5 \times 10^{-2}$. The dimension of the reduced-order model has been considered to be equal to 20, which is a number in the region where the ROM error stagnates in Figure 7.6. We can clearly see that a (relatively) low number of modes yield a response which gives rise to high frequency oscillations pre- and post-shock (or around a sharp gradient). Although the oscillations are bounded, and the locations of the fronts are relatively well captured, the magnitude of the oscillations, as observed for Q_1 in Figure 7.7, are not acceptable. The oscillations in Q_1, along with those in Q_2 and Q_3 (not shown) render a large perturbation in the computation of, for instance, the pressure. We recall that the pressure needs to be kept bounded within a small acceptable regime, and any (significant) perturbations outside this regime could be catastrophic; also see Chapter 2.

Case 2

This test case is taken from [90] which dealt with the high-fidelity numerical approximation of the DFM. It also belongs to the shock-tube type test case. However, unlike the previous test case, i.e., Case 1, the initial conditions are chosen so that the solution is composed of a 1−rarefaction, a 2−contact discontinuity, and a 3−rarefaction wave. We will, hence, refer this test case as the R1-R3 case in the sequel. As in Case 1, a 100 m horizontal pipe is considered which is initially separated into a left and right state at $x_0 = 50$ m. Following [90], the initial conditions of this test case are as follows: $\alpha_{g,L} = 0.35$, $v_{\ell,L} = 1.868$ m/s, $p_L = 192170$ Pa, $\alpha_{g,R} = 0.3$, $v_{\ell,R} = 14.47$ m/s, $p_R = 196690$ Pa, where the notations carry the same meaning as introduced in Chapter 6. We again do not specify the boundary conditions as we will choose the final simulation run-time such that the travelling waves do not reach the boundary of the spatial domain. We perform numerical simulation over the following finite-time horizon: $t \in [0, 1]$. We consider the spatial step size $\Delta x = 5 \times 10^{-2}$. As before, we consider $max(|\lambda_1|, |\lambda_2|, |\lambda_3|) = 1000$ and CFL $= 1$ for computational purposes, and use (7.27) to obtain the temporal step-size given the spatial step-size. Furthermore, the computation of the POD modes is performed as in Case 1.

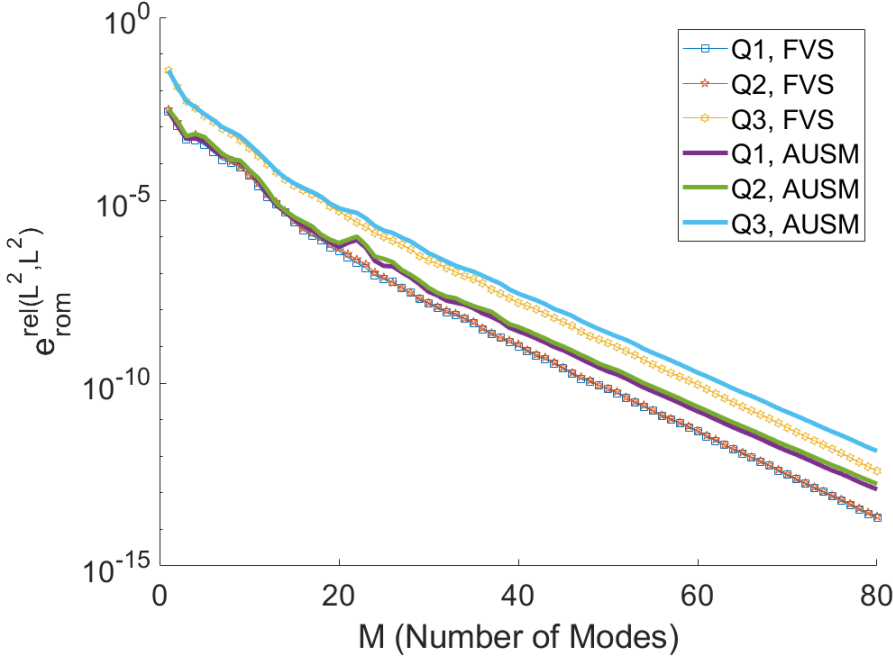Figure 7.8 represents the behaviour of the ROM error, defined in (7.22), ver-

Figure 7.8: R1-R3 case: behaviour of the ROM error (for Q_1, Q_2 and Q_3) versus the number of the POD modes when the spatial step size $\Delta x = 5 \times 10^{-2}$.

sus the number of the POD modes. The ROM errors, corresponding to each of the three physical quantities Q_1, Q_2 and Q_3, are represented in the same plot. Furthermore, the behaviour of the ROM error under the application of different numerical flux approximation schemes, namely the FVS and the AUSM. The implementation of the AUSM results in a larger ROM error than that incurred upon the implementation of the FVS. This is reasonable (and expected) since the AUSM scheme is able to resolve the sharp gradient features more accurately than the FVS scheme; also see [57] and Chapter 6 for the insights about the numerical dissipation mechanisms of these schemes. Overall, the ROM error, in this test case, shows a steeper decay with the increasing number of the POD modes. Such a behaviour can be understood from the following (possible) reasoning. In this case, the only sharp front is due to the presence of the contact discontinuity. However, as dictated by the eigenstructure of the problem, the
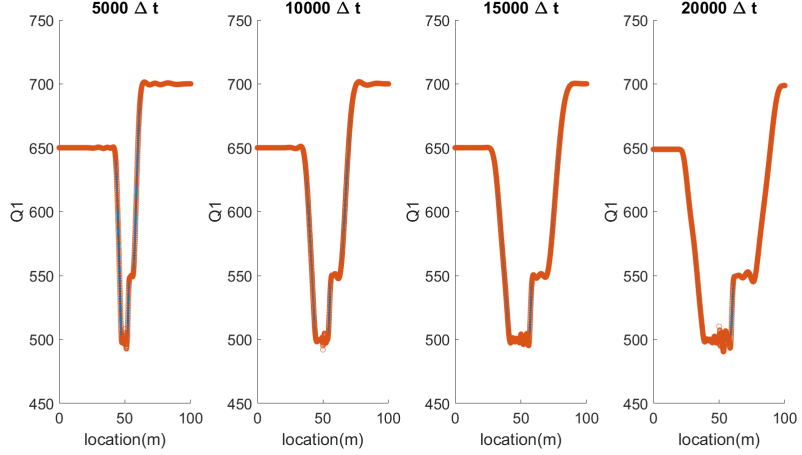
Figure 7.9: Evolution of Q_1 over equidistant time samples. Here, blue colour indicates the high-fidelity solution, and the red circular marker indicates the (reconstructed) solution of the reduced-order model with reduced dimension $r = 15$. And, FVS scheme has been employed for numerical flux approximation.

front evolves at a slow speed across the spatial domain. This is also evident from Figure 7.9, where the second front, i.e., the contact discontinuity, is slowly shifting across the spatial domain. This behaviour is unlike the S1-S3 case, where the (dominant) shock fronts propagate with a high speed. It is easier to approximate a feature that is almost stationary than a (quick) travelling feature. In this R1-R3 case, the macroscopic features happen to be dominant over the propagation aspect of the contact-discontinuity front. In addition, these macroscopic features are easy to reduce as evident from the evolution response shown in Figure 7.9.

Figure 7.9 illustrates the evolution of Q_1, obtained from the high-fidelity and the reduced-order model computation, at 4 equidistant time instants, i.e., at $5000\Delta t$, $10000\Delta t$, $15000\Delta t$ and $20000\Delta t$ with $\Delta x = 5 \times 10^{-2}$. The dimension of the reduced-order model has been considered to be equal to 15, which is a number corresponding to which a relative error $e_{rom}^{rel(\mathscr{L}^2,\mathscr{L}^2)}$ in Figure 7.8 is 10^{-5}. The reduced-order model solution of Q_1 shows almost a perfect match with its high-fidelity counterpart. We can observe some oscillations, but, unlike the previous case, they are not too strong to turn catastrophic.

7.6   Conclusions

In this chapter, we have implemented standard projection-based model order reduction (MOR) approaches to obtain reduced-order models of Burgers' equation and a two-phase flow model, namely, the Drift Flux Model (DFM). We have specifically illustrated the performance of the POD-Galerkin methodology in the scope of the MOR of hyperbolic partial differential equations (PDEs). It is observed that the POD-Galerkin methodology, which solely relies on approximation using linear subspaces, is not appealing for the MOR of hyperbolic PDEs due to an inherent transport-dominated nature of the problem, for instance, due to moving shocks. Although the dimension of the high-fidelity model of the problem of interest is reduced by invoking the standard projection-based MOR approach, the reduced-order model is still required to be high-dimensional to achieve an acceptable level of computational accuracy. Furthermore, (high frequency) oscillations, due to Gibbs' phenomena, are observed around the location of the sharp gradients. As a consequence, the obtained reduced-order models are still not amenable for the purposes of (near) real-time estimation and control, and multi-query simulation purposes. Clearly, this instigates further assessment of the extent to which the studied models can be reduced even further to make it usable for drilling automation. In the next chapters, we delve into assessing the reducibility of transport-dominated problems and devise/propose new MOR framework in order to mitigate the issues observed with the existing standard projection-based MOR approach.

Chapter 8

Method of Freezing combined with Reduced Basis Approximations

This chapter focuses on model complexity reduction for transport-dominated dynamical systems, in which the moving wavefronts hamper the performance of existing reduced-basis models. Moving discontinuous features, such as shocks, are representative characteristics of a wave propagation phenomena, which needs to be accurately resolved in the scope of modelling Managed Pressure Drilling (MPD) systems and several other scientific applications. The location of discontinuous fronts and the speed of propagation varies for different parameters of interest and/or different states of the system. Furthermore, the fronts also interact with each other while evolving across the spatial domain, which induces topological changes in the solution. These evolving and localized discontinuous features affect the performance of the standard projection-based model order reduction (MOR) approaches. While standard MOR approaches do not exploit the underlying hyperbolic structure/physics/dynamics of the problem, we strive to utilize such system behaviour in devising a novel MOR framework, in particular for problems with moving discontinuities. To this end, we employ the "Method of Freezing"/Symmetry Reduction as a pre-processing ingredient before invoking the standard projection-based MOR approaches. In the pursuit to develop a MOR framework for multiple moving, interacting and possibly merging discontinuities, we make a fundamental modification

in the formulation of the "Method of Freezing" by proposing a flux re-distribution approach in order to curb the numerical instabilities induced by wavefront collisions. Finally, we investigate the numerical performance of the combined approach of the "Method of Freezing" and the standard projection-based reduced-basis approximations on few model problems, which embody the challenge faced in the model reduction of MPD systems.

8.1   Introduction

We recall that we are interested in developing a Model Order Reduction (MOR) framework for MPD systems. To this end, it is essential to develop a reduced-order counterpart of single- and two-phase hydraulic models. Also, recall that these hydraulic models are governed by (conditionally) hyperbolic partial differential equations (PDEs). In the previous chapter, we applied existing projection-based methodology on one of the two-phase flow models, i.e., the Drift Flux Model. We observed that the standard projection-based MOR approaches do not yield accurate and stable reduced-order models of lowest dimensionality, thereby further improvements are required in this context. Furthermore, an efficient MOR framework is required to be able to deal with wavefront movements and topological changes induced due to wavefront interactions. This motivates us to investigate and propose efficient, advanced and automated approaches to obtain (stable) reduction of the desired order, while still guaranteeing the accurate approximation of wave propagation phenomena.

A lot of interest has emerged within the scientific community to tackle the challenges of the reduced-order representation of transport-dominated problems. Recently developed/proposed approaches can be classified as follows: *(i)* (data-based and model-based) time and space-dependent coordinate transformation/symmetry reduction framework [28, 31, 73, 131, 134, 150, 154, 155], *(ii)* optimal-transport based framework [55, 86, 151], *(iii)* interpolation/dictionary/tracking framework [172, 202, 203], *(iv)* adaptive and stabilization strategies [29, 36, 141], and *(v)* deep learning/neural network concepts [83, 102].

While several approaches, as mentioned in the previous paragraph, have been developed, we pursue the idea of expressing modes in a co-moving system, i.e., coordinate transformation/symmetry reduction, to build (and improve) a framework for model order reduction of transport-dominated problems. Such a framework consists of, firstly, an important pre-processing step, then, a reduction step and, finally, post-processing of the quantity of interest. The additional pre-processing framework, in the considered approach, is suitable for transport-

dominated problems as it helps to obtain a faster decay of the singular values by transforming a diagonal structure of the solution in space-time to a rectangular structure in space-time. As a consequence, such a framework aids to obtain a lowest possible dimensional representation.

Various strategies, that utilize the aforementioned novel pre-processing step in the MOR framework, have been proposed to reduce the dimensionality of transport-dominated problems. These include: the "Method of Freezing" in conjunction with non-linear reduced basis approximations [134], the Shifted POD [28, 150], the Approximated Lax Pair (ALP) approach [73, 74], etc. The Shifted POD is a data-based MOR approach. The "Method of Freezing" (in conjunction with reduced basis approximations) and the ALP approach use the underlying PDE model to achieve a reduced-order model representation. In the ALP approach, the basis is constructed by computing the eigenfunctions of a linear Schrodinger operator associated with an initial condition. The basis is then propagated in a way that it remains an eigenbasis of the Schrodinger operator at every time instant. However, no theoretical proof of choosing such an ansatz has been presented. Moreover, the ALP approach relies on self-adjointness of the operator, which cannot always be guaranteed, particularly for hyperbolic PDEs/transport-dominated problems. The "Method of Freezing" is based on a Lie group representing the translation of the solution. This method aims to pre-process a solution manifold and separate the dynamics in the group direction from the dynamics in the remaining directions of the phase space. In the pursuit of developing a model-based reduction framework and in the wake of the potential shortcomings of the ALP approach, we focus on the "Method of Freezing" to develop a MOR framework for transport-dominated problems.

The concept of freezing a single wavefront (or a pulse), developed in [26], has been applied for parabolic and hyperbolic problems in [134, 152]. The authors in [152] employed the method for capturing similarity solutions (in multi-dimensional setting), and not for developing a reduced-order model. The only work which exploits the "Method of Freezing" in the scope of model order reduction is [134]. A setting, considered in [134], falls in the realm of hyperbolic problems, where the advection is predominantly along a single direction. Hence, the problem can be dealt with by establishing a single co-moving frame. Furthermore, [134] considers the imposition of periodic boundary conditions. In addition, the equivariance of the discrete operator is assumed to hold. This equivariance property is exploited in [134] to obtain an efficient offline and online decomposition by invoking the POD-Galerkin methodology in conjunction with the Empirical Operator Interpolation. However, the existing MOR framework, as in [134], which builds on the "Method of Freezing" (for a sin-

gle front), lacks consideration of dealing with multiple propagating fronts and non-periodic boundary conditions. In the scope of the MPD, it is by now well understood that multiple wavefronts are propagating across a one-dimensional spatial domain; also see Chapters 6 and 7. Furthermore, the importance of imposing non-periodic boundary conditions is self evident in the physical context; also see Chapter 6. The application motivated requirements along with the current state-of-the-art of the "Method of Freezing" necessitate to extend the MOR framework developed in [134].

The "Method of Freezing", first developed in [26], has been extended in [27] to facilitate dealing with multiple fronts/pulses. However, such a concept of freezing multiple fronts/pulses has not been extended/applied to hyperbolic problems, particularly in the scope of model order reduction. A straightforward adaptation of the method for the hyperbolic setting results in stability issues even in the high-fidelity computations of the transformed model problem; see Section 8.2.3 for further details. The observations can be supported via the following reasoning. Firstly, the authors in [27] developed and numerically analyzed the method in the scope of parabolic and semi-parabolic problems. It is well-known that the solutions of parabolic problems are well-behaved (in terms of regularity) compared to the solutions of hyperbolic problems. Secondly, the collision or merging of pulses, treated in [27], is technically different from the collision or merging of (discontinuous) wavefronts. Furthermore, according to [27], the method of creating multiple frames of reference is quite robust with respect to the choice of partition functions. However, this robustness is challenged when such a method is applied in the hyperbolic setting (see Sections 8.2.3 and 8.4). We recall that we are eventually interested to build a reduced-order model (unlike [27]), and that we seek for the features, obtained from a transformed model problem, such that they guarantee an efficient reduced-order representation. Hence, an approach which may be appropriate to handle strong pulse interactions may not perform satisfactorily under strong front interactions, particularly from the model order reduction point of view. The above discussion puts forward the need of fundamental modifications to extend the use of the "Method of Freezing" for a larger class of hyperbolic problems.

The aim of this chapter is to transform the model problem via the "Method of Freezing" and then reduce it by using the standard POD-Galerkin methodology. The combined approach of the "Method of Freezing" and the POD-Galerkin methodology provides a complete novel MOR framework for dealing with the non-linear transport-dominated problems.

It is worth mentioning that most of the recent literature, primarily built

on the concept of time-dependent shifts, have just focused on the construction of basis functions, and not on the construction of the reduced-order model; see [150, 151]. Moreover, most of the existing literature have been built upon a single co-moving frame construct; for instance, see [31]. Also, recently, the complete framework, starting from the generation of the basis functions upto constructing a fully reduced-order model, has been developed by using the Shifted POD approach in [28] in the scope of dealing with multiple moving waves. However, the ability of the aforementioned framework in dealing with interacting and merging discontinuous fronts is still subject to further research/investigation. In this chapter, we not only demonstrate the generation of basis functions, but also illustrate the use of these generated basis in constructing a reduced-order model for non-linear transport-dominated problems with multiple moving, interacting and possibly merging waves.

The main contributions of this chapter are as follows: *(i)* we apply the existing principles of the "Method of Freezing" and assess the efficiency of the approach in the scope of hyperbolic problems with multiple evolving and interacting (discontinuous) wavefronts, *(ii)* we make a fundamental modification to the "Method of Freezing" by proposing novel flux re-distribution approaches in order to curb the additional travelling structures and/or numerical instabilities that are induced during the interactions between the evolving fronts, *(iii)* we modify the ansatz, responsible for yielding transformed variables, to account for non-periodic boundary effects in the formulation of the "Method of Freezing", *(iv)* we invoke the POD-Galerkin methodology to reduce the transformed model problem that results after the application of the "Method of Freezing", and *(v)* we assess the (combined) "Method of Freezing" and projection-based methodology on several numerical test-beds that are representative of the interaction between the wavefronts propagating along the same or in the opposite direction.

This chapter is organized as follows. Section 8.2 is dedicated to the mathematical formulation of the "Method of Freezing" in the scope of hyperbolic problems. It also consists of aspects related to the numerical discretization of the transformed model problem. Furthermore, it puts forward the issues pertaining to an efficient offline and online decomposition, and also discusses about the design and the choice of the formulation related variables/functions. In Section 8.3, we propose two methodologies to deal with discontinuous wavefront interactions. We then discuss combining the approach of the "Method of Freezing" with reduced basis approximations (POD-Galerkin methodology) to develop a reduced-order model in Section 8.4. We perform several challenging numerical experiments in Section 8.5 to assess the potential of the proposed approach,

and also compare its performance with that of the standard projection-based approach. Finally, Section 8.6 ends the chapter with conclusions.

8.2 Framework for freezing multiple wavefronts

We use the "Method of Freezing" [26, 27] as a pre-processing ingredient in our MOR framework. The "Method of Freezing" is also known by the name of symmetry reduction framework in some literature [154, 155]. This method is inspired by geometric mechanics and was originally developed to study relative equilibria of evolution equations. The "Method of Freezing" approach intuitively seems powerful in exploiting the underlying structure and, thereby, facilitating a much more effective model reduction. Concerning the underlying structure, the transport-dominated problems, for e.g., advection equation, Burgers' equation, etc., often possess translation invariance. This property can be used to reduce the dimensionality of the problem. The "Method of Freezing", exploiting this property, maps all symmetry-related solutions to a single class of solutions. This method helps to decompose the original dynamics into shape dynamics and
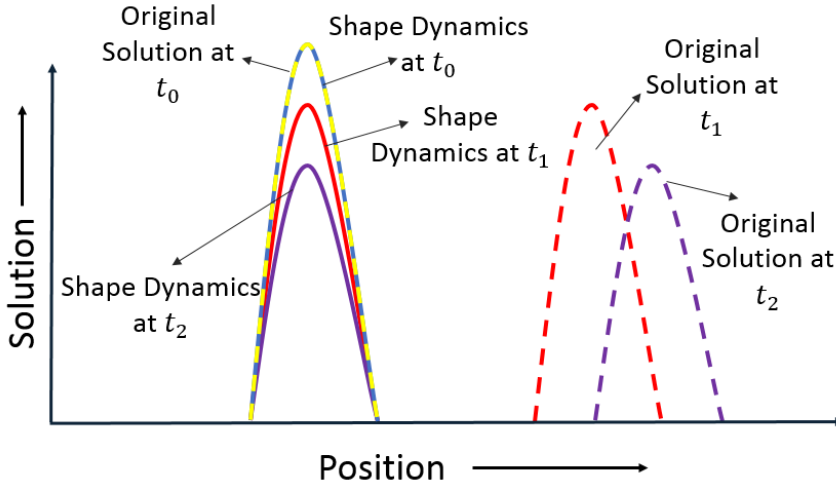

Figure 8.1: Concept Illustration: Here, dashed lines represent the evolution of original dynamics and solid lines represent the evolution of shape dynamics. Original and shape dynamics co-incide at $t_0 = 0$.

travelling dynamics. We refer to Figure 8.1, which shows a schematic description of the evolution of shape dynamics, which are essentially frozen around the initial configuration, and the original (and reconstructed) dynamics.

The dimensionality of the reduced-order model is generally large due to the travelling nature of the solution. The travelling part of the solution is exactly the component we aim to filter out from the solution dynamics. First, we formulate the transformed model problem, and solve for shape dynamics and travelling dynamics by using numerical discretization methods. We then invoke standard projection-based MOR on the shape dynamics, which is devoid of the translational feature of the original realization. The aim is to approximate the shape dynamics with a minimal number of modes/basis functions. This is crucial since there is a one-to-one correlation between the number of modes/basis functions and the resulting dimension of the reduced-order model. The reduced shape dynamics are then acted upon by a (time-dependent) discrete shift operator, which is built using the knowledge of the travelling dynamics, to give an approximate realization of the quantity of interest (QOI) in the reduced space. An inverse group action can then be applied to lift the QOI/solution from the reduced space to the high-fidelity space. We will call this lifted function as the "reconstructed solution". Finally, we can analyze the errors accrued during the numerical stages of the whole MOR framework. Figure 8.2 illustrates the whole workflow in a schematic manner.

Remark 8.2.1 *It should be mentioned that the consideration of non-periodic boundary conditions also imposes the need of devising multiple co-moving frames. For*
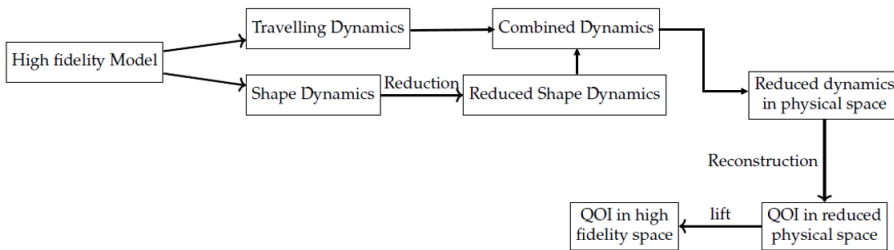

Figure 8.2: Flow diagram of the "Method of Freezing" in conjunction with reduced-basis approximations.

instance, the inflow boundary conditions may result in the development of a front, which propagates across the spatial domain as the time evolves. Other fronts may already be present within the spatial domain due to the choice of the initial conditions. Moreover, extra fronts could also develop due to the reflection at the boundaries. The evolution of these several fronts can ultimately lead to the interaction between the propagating fronts. An evident requirement is to have a MOR framework to deal with the multiple evolving and interacting fronts in the internal part of the spatial domain, i.e., away from the boundary. Hence, in this section, we only focus on developing multiple co-moving frames to deal with the multiple co-existing and interacting wavefronts in the internal part of the spatial domain, and restrict ourselves to the consideration of periodic boundary conditions.

Remark 8.2.2 *We focus on reproducing the results of the time-dependent equation via the standard and proposed MOR framework. This reproduction step is essential before attempting to develop a parametric reduced-order model as we cannot hope to have an effective low-complexity reduced-order model if the numerical approach does not fare well in the reproduction step.*

8.2.1   Standard (infinite-dimensional) formulation

Let us consider a scalar PDE, which can admit multiple pulses/wavefronts in its solution profile, of the following form:

$$\partial_t u + f(u, \partial_x u) = 0, \tag{8.1}$$

where $x \in \Omega$ represents the spatial domain and $t \in [0, t_f]$ represents the time domain with t_f denoting the final simulation time, $u \in H^1$ is the solution we are interested in, and $f : H^1 \times H^1 \to H^{1*}$ is the associated flux function. Here H^1 denotes the Sobolev space of functions that also possess a weak derivative, and H^{1*} denotes the dual space of H^1.

Remark 8.2.3 *It is worth mentioning that the solutions of hyperbolic problems are non-classical and do not generally lie in H^1. The derivation of the "Method of Freezing" [26, 27], originally developed and numerically analyzed for parabolic problems, assumes that the solutions are smooth enough. It is hard to do numerical analysis and show any properties, in the scope of the "Method of Freezing (and also in general), for non-classical solutions.*

We seek to establish multiple co-moving frames corresponding to individual fronts present in the solution. The original solution can be expressed as a combination, i.e., a simple superposition, of solutions in each co-moving frame and

can be written as follows:

$$u(x,t) = \sum_{j=1}^{N} v_j(\xi_j(x,t),t), \text{with } \xi_j(x,t) = x - g_j(t), \qquad (8.2)$$

where v_j stands for the shape dynamics, g_j stands for the travelling dynamics, and N refers to the number of co-moving frames in consideration. Moreover, $\xi_j = x - g_j$ denotes the transformed space coordinate.

Remark 8.2.4 *One could also think of a weighted combination of local solutions in (8.2). However, in such a case, one would need to devise a rationale for computing the weights.*

Let us now derive the dynamics that govern v_j and g_j such that $u(x,t)$ defined in (8.2) is a solution of (8.1). Hereto, we first express $\partial_t u$ using (8.2), leading to

$$\partial_t u = \sum_{j=1}^{N} [\partial_t v_j(x - g_j(t),t) - \partial_{\xi_j} v_j(x - g_j(t),t)\partial_t g_j]. \qquad (8.3)$$

Using the decomposition illustrated in (8.2), applying the concept of chain rule and using (8.1), the temporal derivative of the solution, i.e., $\partial_t u$, can also be expressed as follows:

$$\partial_t u = -f(\sum_{k=1}^{N} v_k(x - g_k(t),t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(x - g_k(t),t)). \qquad (8.4)$$

We now invoke the concept of time-dependent partition of unity to couple local functions in a non-local manner. Using this concept, (8.4) can be re-written as follows:

$$\partial_t u = -\sum_{j=1}^{N} [\frac{\phi_j(x - g_j(t))}{\sum_{k=1}^{N} \phi_k(x - g_k(t))} f(\sum_{k=1}^{N} v_k(x - g_k(t),t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(x - g_k(t),t))], \quad (8.5)$$

where $\phi_i, i \in \{1,2,...,N\}$ are the functions constituting a partition of unity.

The right-hand sides of (8.3) and (8.5) can be equated to impose the governing

dynamics in (8.1). This yields:

$$\sum_{j=1}^{N} [\partial_t v_j(x - g_j(t), t) - \partial_{\xi_j} v_j(x - g_j(t), t) \partial_t g_j] =$$

$$- \sum_{j=1}^{N} [\frac{\phi_j(x - g_j(t))}{\sum_{k=1}^{N} \phi_k(x - g_k(t))} f(\sum_{k=1}^{N} v_k(x - g_k(t), t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(x - g_k(t), t))]. \quad (8.6)$$

Considering that the individual components in the left-hand side and the right-hand side of (8.6) are equal for $j \in \{1, 2, ..., N\}$, we obtain

$$\partial_t v_j(x - g_j(t), t) - \partial_{\xi_j} v_j(x - g_j(t), t) \partial_t g_j(t) =$$

$$-[\frac{\phi_j(x - g_j(t))}{\sum_{k=1}^{N} \phi_k(x - g_k(t))} f(\sum_{k=1}^{N} v_k(x - g_k(t), t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(x - g_k(t), t))]. \quad (8.7)$$

Remark 8.2.5 *While (8.6) always holds, we have made a design choice that even the individual components, which constitute the sum in the left-hand and the right-hand side of (8.6), are equal. This design choice, leading to (8.7), is just a sufficient condition and is not a necessary condition to formulate the governing shape an travelling dynamics.*

Defining $\xi_j := x - g_j$, and re-arranging some terms in (8.7), we obtain

$$\partial_t v_j(\xi_j, t) = \partial_{\xi_j} v_j(\xi_j, t) \partial_t g_j(t) - \frac{\phi_j(\xi_j + g_j(t) - g_j(t))}{\sum_{k=1}^{N} \phi_k(\xi_j + g_j(t) - g_k(t))} \times$$

$$f(\sum_{k=1}^{N} v_k(\xi_j + g_j(t) - g_k(t), t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(\xi_j + g_j(t) - g_k(t), t)). \quad (8.8)$$

Equation (8.8) can be compactly written as follows:

$$\partial_t v_j(\xi_j, t) = \partial_{\xi_j} v_j(\xi_j, t) \partial_t g_j(t) - \frac{\phi_j(\xi_j)}{\sum_{k=1}^{N} \phi_k(\xi_{kj}^g)} f(\sum_{k=1}^{N} v_k(\xi_{kj}^g, t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(\xi_{kj}^g, t)),$$

$$(8.9)$$

where $\xi_{kj}^g := \xi_j - g_k(t) + g_j(t)$. Equation (8.9) expresses the shape dynamics. The group component $g_j(t)$, representing the travelling dynamics, satisfies:

$$\partial_t g_j(t) = \mu_j, \quad (8.10)$$

where μ_j is still unknown. In other words, the notion of group component adds additional degrees of freedom μ_j and, thereby, introduces under-determinacy. As a consequence, algebraic constraints or so-called phase conditions are required to have sufficient equations for the unknowns in the formulation. The phase conditions aid to determine μ_j, which upon using (8.10) provides the knowledge of the group/travelling component g_j. This implies that the phase conditions are linked to the propagation speed of the fronts. It is known that the hyperbolic PDEs are unforgiving when it comes to the propagation of spatial and temporal truncation errors because the information flows with characteristic velocities. Thus, the phase conditions should comply with the underlying physics and yield correct transport speed(s) and direction(s). To this end, the phase conditions are devised such that they decompose a space- and time-dependent solution into a time-dependent group orbit and a spatial profile that demonstrates minimum variation over time. Two variants for designing the phase conditions exist in the literature. One of them is called the orthogonal phase condition and the other is known as the fixed phase condition; see [26,27]. We resort to use the orthogonal phase conditions throughout this chapter. These are mathematically written as follows:

$$0 = < \partial_{\xi_j} v_j, \left(-\frac{\phi_j(\xi_j)}{\sum_{k=1}^{N} \phi_k(\xi_{kj}^g)} f(\sum_{k=1}^{N} v_k(\xi_{kj}^g, t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(\xi_{kj}^g, t)) + \mu_j \partial_{\xi_j} v_j \right) >_{\mathscr{L}^2},$$

(8.11)

where, $< .,. >_{\mathscr{L}^2}$ denotes the $\mathscr{L}^2$ inner product.

Alternatively, the propagation speeds (or μ_j) can be prescribed a-priori for some physical problems. However, a-priori prescription of these speeds will not always be feasible and, hence, an automated methodology, such as the notion of the phase conditions, is needed to determine the propagation speed and, consequently, the shifts.

Remark 8.2.6 *The notion of phase conditions for determining the propagation speeds (and the shifts) is in contrast to the Shifted POD approach [150], where the authors use peak or front tracking, threshold search or singular value maximization for determining the shifts from the snapshot/simulation data. However, the "Method of Freezing" does not require such a-posteriori analysis of the simulation data to extract the information of the shape and velocities of the fronts present in the original solution.*

To summarize, the three key ingredients in the approach of the "Method of Freezing" are: *(i)* shape dynamics governed by (8.9), *(ii)* travelling dynamics/-group component governed by (8.10), and *(iii)* phase conditions governed by

(8.11). The "Method of Freezing" transforms an original PDE (8.1) into a system of Partial Differential Algebraic Equations (PDAEs) given by (8.9),(8.10) with (8.11). Now, rather than seeking for the solution u via a numerical scheme, we seek to solve for the shape dynamics v_j and travelling dynamics g_j by using a suitable numerical method.

8.2.2   Discrete Framework

In this section, we provide brief insights about the spatial and temporal discretization methods employed for the high-fidelity computation of the transformed model problem governed by (8.9),(8.10) and (8.11). The resulting system of PDAEs to be discretized, for all $j \in \{1,2,\dots,N\}$, is summarized below:

$$\partial_t v_j(\xi_j,t) = \overbrace{\partial_{\xi_j} v_j(\xi_j,t)}^{H_{j,1}} \mu_j(t) - \overbrace{\frac{\phi_j(\xi_j)}{\sum_{k=1}^{M} \phi_k(\xi_{kj}^g)}}^{Q_j^g(\xi_j,t)} f(\overbrace{\sum_{k=1}^{N} v_k(\xi_{kj}^g,t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(\xi_{kj}^g,t)}^{H_{j,0}^{shifted}}), \quad (8.12)$$

$$0 = <\partial_{\xi_j} v_j, (-\frac{\phi_j(\xi_j)}{\sum_{k=1}^{N} \phi_k(\xi_{kj}^g)} f(\sum_{k=1}^{N} v_k(\xi_{kj}^g,t), \sum_{k=1}^{N} \partial_{\xi_k} v_k(\xi_{kj}^g,t)) + \mu_j \partial_{\xi_j} v_j) >_{\mathscr{L}^2}, \quad (8.13)$$

$$\partial_t g_j = \mu_j(t). \quad (8.14)$$

We first perform the spatial discretization to obtain a semi-discretized form of the system of PDAEs and then perform the temporal discretization. We take the Finite Volume (FV) approach to perform the spatial discretization. We employ the Lax-Friedrichs scheme [105] or the Upwind scheme [105] to compute the flux numerically. The spatial discretization of the PDAE system yields:

$$\partial_t v_j = -(Q_j^g H_{j,0}^{shifted}) + (H_{j,1})\mu_j, \quad (8.15)$$

$$\left(H_{j,1}^T H_{j,1}\right)\mu_j = \left(H_{j,1}^T Q_j^g H_{j,0}^{shifted}\right), \quad (8.16)$$

$$\partial_t g_j = \mu_j(t), \quad (8.17)$$

where $Q_j^g, H_{j,0}^{shifted}$, and $H_{j,1}$ are discrete operators accounting for the spatial discretization of the quantities as indicated in (8.12), and $H_{j,1}^T$ denotes the transpose of the discrete operator $H_{j,1}$.

Remark 8.2.7 *The Lax-Friedrichs scheme is known to be monotonicity preserving (and, hence, "Total Variation Diminishing" as well) under its standard application to scalar conservation laws. This property ascertains that there will be no numerical oscillations in the solution. It should, however, be realized that the spatial discretization of the system of PDAEs (8.12) - (8.14) results in FV discrete operators, which are no longer standard FV discrete operators. Hence, the application of the Lax-Friedrichs scheme calls for further numerical analysis and to obtain the conditions under which the monotonicity can be guaranteed for the system under consideration.*

The semi-discretized system (8.15) - (8.17) can now be solved over time by employing a suitable method for temporal discretization. We restrict ourselves to the usage of an explicit first-order Forward Euler scheme for the temporal discretization. We could also consider an implicit Euler scheme or higher-order explicit Runge Kutta schemes for the temporal discretization.

Remark 8.2.8 *We avoid exhaustive discussion about the numerical implementation of the spatial and temporal discretization as the considered methods are usually applied and well documented in the literature [105].*

We have now obtained a fully discrete representation of the transformed model problem. Before using the resulting knowledge in conjunction with the reduced-basis approximations, we delve into the discussion about the rationale behind the design of the partition functions ϕ_i and the choice of the number of co-moving frames N.

8.2.3 Design of partition functions and choice of number of frames

The partition functions form the building block in decomposing the solution into the shape and travelling dynamics; see (8.5). In principle, the partition functions ϕ_j constitute a partition of unity if $\sum_{k=1}^{N} \phi_k(\cdot) = 1$. It is not necessary to ensure $\sum_{k=1}^{N} \phi_k(\cdot) = 1$ for the formulation of the "Method of Freezing" to be mathematically correct. However, it is necessary to ensure that $\sum_{k=1}^{N} \phi_k(\cdot) \neq 0$. The number of partition functions is decided by a-priori decision on the number of co-moving frames N, which need to be considered in the scope of the problem at hand. One of the possible choices of these partition functions is illustrated in Figure 8.3, where we assume $N = 3$. Other possible choices of the partition
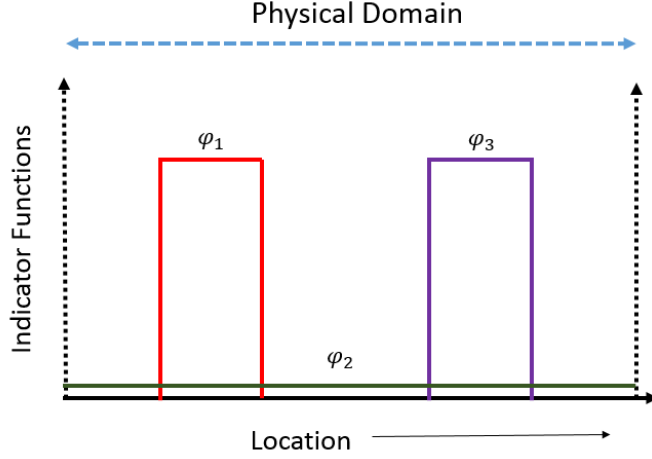
Figure 8.3: Possible choice of partition functions.

functions are illustrated in Figure 8.4, where we again consider $N = 3$. The partition functions can also have a smooth profile, unlike that of Figures 8.3 and 8.4. Theoretically, there is a one-to-one link between the decomposition of the solution field and the number of structures (fronts) with different transport velocities. However, in principle additional frames are required to obtain desired outcomes, particularly from the model order reduction point of view, which demands that not more than one translating front should be present in each of the decomposed components v_j. Different types of partition functions result in different qualitative and quantitative behaviour of the shape dynamics v_j. It should also be mentioned that it is necessary to have the overlap between the constituting partition functions ϕ_i to avoid the numerical breakdown. Furthermore, these partition functions also impact the monotonicity behaviour of the numerical scheme, such as the Lax-Friedrichs scheme.

There does not exist any strong theory that can help us in optimally choosing the partition functions. To obtain some insights about the high-fidelity behaviour of the transformed model problem, we now construct a test case. We will use it to showcase the challenges (and importance) of the design of partition functions and the choice of number of frames. The test case falls in the realm of scalar conservation laws. A few popular (non-linear) scalar conservation laws are Burgers' equation, Korteweg de Vries equation, etc. While the
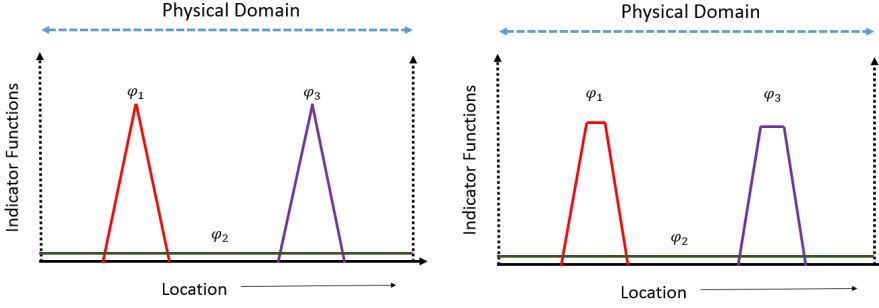
Figure 8.4: Possible choice of partition functions.

solitons obtained in the Korteweg de Vries equation can pass each other without interacting non-linearly, the wavefronts in Burgers' equation generally interact non-linearly. In the scope of developing a MOR framework to deal with moving, interacting and possibly merging discontinuities, Burgers' equation is, hence, suitable. In the sequel, we will use Burgers' equation to test the robustness of the "Method of Freezing".

The Burgers' Equation can be written as follows:

$$\partial_t u + \partial_x(\frac{u^2}{2}) = 0, \quad x \in [0, L], \quad t \in [0, t_f]. \tag{8.18}$$

We impose periodic boundary conditions, $u(0, t) = u(L, t)$, where L is the length of the spatial domain, and the initial condition for the simulation is illustrated in Figure 8.5. The ensuing discussion is based on this test case.

Clearly, the initial condition is characterized by two Gaussian pulses. These Gaussian pulses are centred around different spatial points, and also have different amplitude (and sign). It is well-known that the Gaussian pulse will develop into a shock front after a finite period of time. Given the fact that two different fronts (structures) develop, we strive to decompose the solution field into two local constituents. We begin by using two smooth partition functions (with non-compact support) to decompose the solution field. We then numerically solve the original dynamics in (8.18) and the associated system of PDAEs: (8.12), (8.13), (8.14) with $f(u, \partial_x u) = \partial_x(\frac{u^2}{2})$. The reconstructed dynamics and the original dynamics match perfectly. However, when the two wavefronts, propagating from opposite directions, travel towards each other, we witness the contributions of both propagating wavefronts in the shape dynamics corresponding
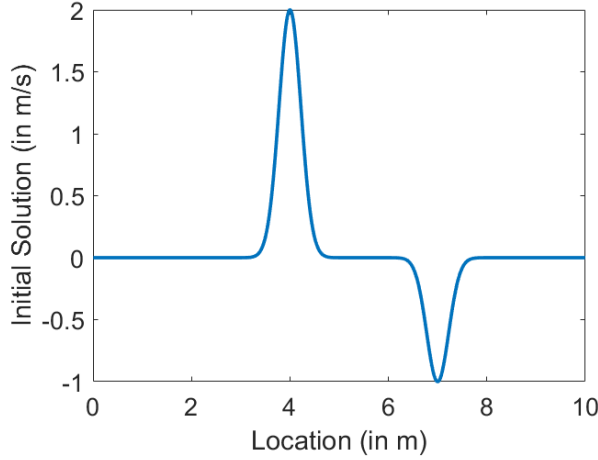
Figure 8.5: Initial condition $u(x, t = 0)$ for a test case with two opposite moving wave-
          fronts.

to each of the two (local) co-moving frames. Such contributions can be math-
ematically attributed to the second term in the right-hand side of (8.12). It is
observed that the contribution from the wavefront in the left frame[1] (resp., right
frame[1]) to the resulting shape dynamics of the right frame (resp., left frame) re-
sults in a shock of different height in the right frame (resp., left frame). Further-
more, an additional (discontinuous) front of a certain height with a non-zero
speed is observed in both co-moving frames. The height of this additional dis-
continuity is linked to the difference between the height of the already existing
(discontinuous) front in the co-moving frame and the "true" height of the cor-
responding (discontinuous) front in the original solution of (8.18). Moreover,
the shape profile of the already existing (discontinuous) front undergoes signif-
icant changes. Such features, in particular the additional travelling structure in
the co-moving frame, will have an impact on the decay of the singular values,
which are obtained after performing the singular value decomposition on the
snapshot matrix of the shape dynamics v_j in each co-moving frame. Such an
observation is not something we envisioned initially from the "Method of Freez-
ing" as the main idea of this strategy is to factor out the travelling dynamics

---

[1] The co-moving frame localized around the left (resp., right) Gaussian pulse, depicted in Figure
8.5, is referred to as the left (resp., right) frame in the subsequent discussion.

from the profile of the shape dynamics. As there still exist travelling structures in the profile of the shape dynamics, we seek for a remedy and adopt an idea of using additional frames of reference along with different shapes of the partition functions. This idea is explained below using the same test case.

We make the support of the partition functions as small as possible. In order to do so, we use non-smooth and compact indicator functions as partition functions. Furthermore, we define an additional partition function, which takes a constant value that is non-zero everywhere in the spatial domain. Such a choice of the partition functions is schematically depicted in Figure 8.3. Henceforth, we consider three co-moving frames, i.e., $N = 3$, for such a test case. The constant partition function, i.e., ϕ_2 in Figure 8.3, would correspond to the background solution which does not contain any moving discontinuities in the "true" physical space. Hence, we manually set the speed (and the shift) of the co-moving frame, corresponding to the constant partition function, to zero. The automated shift determination is done for two of the three co-moving frames by using the notion of the phase conditions as discussed earlier in Section 8.2.1. High-fidelity numerical experiments of the system of PDAEs, with aforementioned choice of the partition functions and the number of frames, still yielded *(i)* extra structures in the profile of the shape dynamics, and *(ii)* numerical oscillations/stability issues. The extra structures include moving/stationary, but non-differentiable continuous features[2] in the shape dynamics. It is important to mention that the approximation of such moving/stationary, non-differentiable features is as difficult (maybe more difficult) as the approximation of the moving fronts, which is already a serious concern (see Chapter 7). It is, hence, necessary to obtain shape dynamics that are devoid of such non-differentiable features. Furthermore, the numerical oscillations/stability issues are observed when the compactly supported partition functions start overlapping in the "true" physical space. A low-dimensional representation, which is our end goal, should not suffer from the spurious behaviour and extra/unwanted structures and, hence, to begin with, the high-fidelity representation of the shape dynamics v_j should be devoid of such features. The observations, so far, hint to the need of fundamental modifications to the framework of the "Method of Freezing". To this end, we need to propose measures to enhance the ability of the framework in dealing with dis-

---

[2]The support of the partition functions can be tuned. It is observed that the shape dynamics may no more constitute of the non-differentiable features. However, the support of the partition functions needs to be changed appropriately depending on the direction of the propagating wavefronts and the problem under consideration. To conclude, the choice of the support of the constituting partition functions to avoid non-differentiable features is to-date a kind of a trial and error exercise and does not have sound mathematical background yet.

continuous wavefront interactions. These mitigating measures are elaborated in the next section.

8.3 Flux re-distribution to deal with discontinuous wavefront interactions

The need of fundamental modifications to the framework of the "Method of Freezing" has been discussed in Section 8.2.3. In this section, we aim to improve the performance of the "Method of Freezing" in dealing with merging (discontinuous) wavefronts. We propose two flux re-distribution approaches in order to curb the numerical instabilities and extra/unwanted structures induced by wavefront interactions and collisions, and to ensure that the obtained shape dynamics are easily reducible. We introduce the proposed approaches in this section and, eventually, discuss the merits and demerits of the propositions in Section 8.4 via numerical case studies.

8.3.1 Proposition 1

This approach can be seen as a modification of the design choice made in (8.7). Figure 8.6 illustrates the idea of flux re-distribution in a schematic manner. The partition functions, referred to as indicator functions in Figure 8.6, introduced into the formalism in Section 8.2, serve as a tool to identify the region of (strong) wave interaction in this proposed approach. The solid lines, shown in red and purple, in Figure 8.6 represent the initial configuration of the partition functions, which will be the same in the "true" physical space as well as in the transformed space. These partition functions will propagate across the spatial domain as the time evolves. The two associated partition functions, represented in physical space by dotted lines, start to overlap after few time steps. We detect the region of overlap and, as a preliminary step, we partition the overlap region into two equal halves. We then allocate the flux corresponding to the left half of the overlap region to the left frame. Similarly, we allocate the flux corresponding to the right half of the overlap region to the right frame. In order to define the approach mathematically, we consider the set

$$x_{pf} = \text{supp } \phi_1 \ \cap \ \text{supp } \phi_2, \tag{8.19}$$

where ϕ_1 and ϕ_2 represent the partition functions associated to the left and the right frame, respectively, and x_{pf} is a collection of the discrete spatial co-
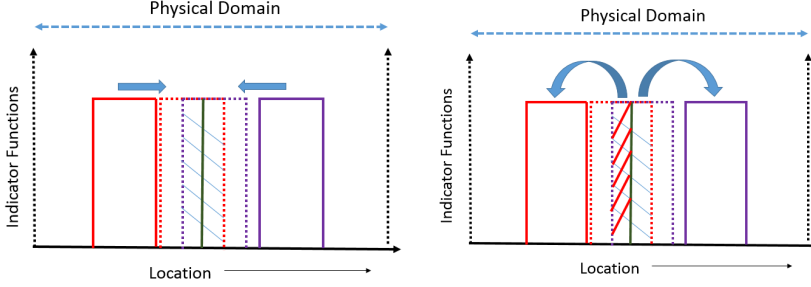
Figure 8.6: An illustration of the proposed flux re-distribution approach (in physical space).

ordinates where the two partition functions ϕ_1 and ϕ_2 overlap. If $x_{pf} = \emptyset$, this implies that the partition functions do not intersect/overlap. We recall the definition $Q_j^g(\xi_j, t) = \frac{\phi_j(\xi_j)}{\sum_{k=1}^{M} \phi_k(\xi_{kj}^g)}$. We propose to modify this definition of $Q_j^g(\xi_j, t)$ if $x_{pf} \neq \emptyset$. We first partition the set, constituting the discrete collection of spatial coordinates, in the following way:

- If the cardinality of the set x_{pf} is an even number l, then we split the elements of the set into two equal parts: $x_{pf1} = [x_1, x_2, ..., x_{\frac{l}{2}}]$, $x_{pf2} = [x_{\frac{l}{2}+1}, x_{\frac{l}{2}+2}, ..., x_l]$ such that $x_{pf1} \cup x_{pf2} = x_{pf}$ holds.

- If the cardinality of the set x_{pf} is an odd number l, then we allocate one extra element to either set x_{pf1} or x_{pf2}. Suppose that we allocate an extra element to the set x_{pf1}. We then have $x_{pf1} = [x_1, x_2, ..., x_{\frac{(l+1)}{2}}]$, and $x_{pf2} = [x_{\frac{(l+1)}{2}+1}, x_{\frac{(l+1)}{2}+2}, ..., x_l]$, such that $x_{pf1} \cup x_{pf2} = x_{pf}$ holds.

Based on the aforementioned partitioning, we now define Q_j^g for $j = 1, 2$ as follows:

$$Q_1^g(\cdot, t) = \begin{cases} 1 \ \forall \ (\cdot) \in x_{pf1}, \\ 0 \ \forall \ (\cdot) \in x_{pf2}. \end{cases} \tag{8.20}$$

Similarly,

$$Q_2^g(\cdot, t) = \begin{cases} 0 \ \forall \ (\cdot) \in x_{pf1}, \\ 1 \ \forall \ (\cdot) \in x_{pf2}. \end{cases} \tag{8.21}$$

In the above discussion, i.e., in (8.20) and (8.21), $j = 1$ will attribute to the left frame and $j = 2$ will attribute to the right frame.

Remark 8.3.1 *The discrete way of defining Q_j^g is ideal from the point of view of numerical implementation.*

$Q_1^g(\cdot, t)$ in (8.20) and $Q_2^g(, t)$ in (8.21) can be determined if x_{pf1} and x_{pf2} are known. Given the partition functions, the sets x_{pf1} and x_{pf2} can be easily be obtained using (8.19).

The proposed approach needs the detection of the region of overlap between the partition functions. It is easier to compute the region of overlap between two compactly supported square (or indicator) functions than to compute it for smooth functions with non-compact support. It is known and also can be intuitively understood that the smooth functions with non-compact support can intersect at multiple points in the physical space, thereby demanding the solution of a non-linear equation at every time iteration. In view of the additional computational burden in detecting the region of overlap between the smooth partition functions, we use non-smooth indicator functions/partition functions in the numerical case studies pursued in this chapter.

All the concepts, except the modified definition of Q_j^g, are the same as those discussed extensively in Section 8.2. Using the new definition of Q_j^g along with all the notations and ingredients introduced in 8.2, we can assess the performance of the proposed approach.

Next, we discuss another proposition to deal with wavefront interactions/-collisions.

8.3.2   Proposition 2

It is apparent that the design of the partition functions is challenging and the numerical results are sensitive to the choice of the partition functions; see Section 8.2.3 for an in-depth discussion about the effect of the choice of the partition functions. The numerical sensitivity can be viewed from the viewpoint of the "Method of Freezing" in terms of the stability issues. Alternatively, the sensitivity can also be viewed from the model order reduction viewpoint, particularly, in terms of the decay of singular values corresponding to the snapshot matrix of each of the shape dynamics. The following proposed measure seeks to address the following questions:

- Can we perform the decomposition without invoking the concept of partition functions?

- Can we avoid the need of additional frames, which do not have a sound theoretical basis, during the decomposition?

Let us again consider a scalar PDE of the following form:

$$\partial_t u + f(u, \partial_x u) = 0. \tag{8.22}$$

The original solution, u, is now decomposed as follows:

$$u(x, t) := u_1(x, t) + u_2(x, t) = \sum_{j=1}^{N=2} v_j(x - g_j(t), t), \tag{8.23}$$

where it has been assumed that the solution u is decomposed into two parts as we know that the solution to (8.18) will have two discontinuous fronts due to the choice of the initial conditions as shown in Figure 8.5. Here, the v_j's are supposedly devoid of the travelling nature, and each of the decomposed part $u_1(x, t)$ and $u_2(x, t)$ constitute not more than one travelling front. Also, the features in the decomposed parts should be such that the travelling dynamics is more dominant than the corresponding shape dynamics. This is important in view of the effectivity of the model reduction strategy.

Remark 8.3.2 *The number of decomposed parts, as in (8.23), are determined by the know-how of the (possible) number of discontinuous features in the solution. In principle, (8.23) and the framework that follows, in the scope of this proposition, can be generalized. However, we limit ourselves to a specific setting in accordance with the decomposition (8.23).*

The temporal derivative of the solution u can now be expressed as follows:

$$\partial_t u = \sum_{j=1}^{N=2} \left(\partial_t v_j(x - g_j(t), t) - \partial_{\xi_j} v_j(x - g_j(t), t) \partial_t g_j \right), \quad \xi_j = x - g_j. \tag{8.24}$$

Using the decomposition illustrated in (8.23), applying the concept of chain rule and using (8.22), the temporal derivative $\partial_t u$ of the solution can also be expressed as follows:

$$\partial_t u = \partial_t u_1 + \partial_t u_2 = -f\left(\sum_{k=1}^{N=2} v_k(x - g_k(t), t), \sum_{k=1}^{N=2} \partial_{\xi_k} v_k(x - g_k(t), t) \right). \tag{8.25}$$

The flux function f is perceived to be composed of four terms, h^{11}, h^{12}, h^{21} and h^{22}. Mathematically, the flux function is equivalently expressed as follows:

$$f\left(\sum_{k=1}^{N=2} v_k(x - g_k(t), t), \sum_{k=1}^{N=2} \partial_{\xi_k} v_k(x - g_k(t), t) \right) = h^{11} + h^{12} + h^{21} + h^{22}, \tag{8.26}$$

where the terms h^{lm} refer to the contribution of the lth local component to the mth frame. Using (8.25) and (8.26), and defining $\partial_t u_1$ and $\partial_t u_2$ as shown below:

$$\partial_t u_1 := \partial_t v_1(x - g_1(t), t) - \partial_{\xi_1} v_1(x - g_1(t), t) \partial_t g_1 = -(h^{11} + h^{21}), \qquad (8.27)$$

$$\partial_t u_2 := \partial_t v_2(x - g_2(t), t) - \partial_{\xi_2} v_2(x - g_2(t), t) \partial_t g_2 = -(h^{12} + h^{22}), \qquad (8.28)$$

we obtain the decomposed components that satisfy the scalar PDE (8.22).

For Burgers' equation, the flux function f can be written as follows:

$$f(u, \partial_x u) = u\partial_x u = (u_1 + u_2)(\partial_x u_1 + \partial_x u_2) = \underbrace{u_1\partial_x u_1}_{h^{11}} + \underbrace{u_1\partial_x u_2}_{h^{12}} + \underbrace{u_2\partial_x u_1}_{h^{21}} + \underbrace{u_2\partial_x u_2}_{h^{22}}, \quad (8.29)$$

or

$$f(u, \partial_x u) = u\partial_x u = (u_1 + u_2)(\partial_x u_1 + \partial_x u_2) = \underbrace{u_1\partial_x u_1}_{h^{11}} + \underbrace{u_1\partial_x u_2}_{h^{21}} + \underbrace{u_2\partial_x u_1}_{h^{12}} + \underbrace{u_2\partial_x u_2}_{h^{22}}. \quad (8.30)$$

Using (8.27), (8.28) and (8.29), the dynamics of the evolution of the decomposed components can be expressed as follows:

$$\partial_t u_1 + u_1\partial_x u_1 + u_2\partial_x u_1 = 0, \qquad (8.31a)$$

$$\partial_t u_2 + u_2\partial_x u_2 + u_1\partial_x u_2 = 0. \qquad (8.31b)$$

This approach does not fall into the category of strict hyperbolic systems as the eigenvalues are real and equal rather than being real and distinct.

Alternatively, using (8.27), (8.28) and (8.30), the dynamics of the evolution of the decomposed components can be expressed as follows:

$$\partial_t u_1 + u_1\partial_x u_1 + u_1\partial_x u_2 = 0, \qquad (8.32a)$$

$$\partial_t u_2 + u_2\partial_x u_2 + u_2\partial_x u_1 = 0. \qquad (8.32b)$$

This approach falls in the category of strict hyperbolic systems, except for the case where $u_1 = u_2$.

Both approaches, as in (8.31) and (8.32), give rise to different local speeds. In principle, any of these two approaches can be utilized in the scope of constructing a co-moving frame by using the knowledge of the local speed. However, the choice of the discretization method of the aforementioned two approaches is critical. The discretization methodology becomes furthermore important due to the presence of the non-conservative terms. The coupled Burgers' system, as in (8.31) and (8.32), have been discretized in the past and, hence, we do not delve into the details of the discretization of the obtained system(s).

Remark 8.3.3 *Recall that we envision to obtain the decomposed parts $u_1(x,t)$ and $u_2(x,t)$ such that they do not constitute more than one travelling front. However, contrary to our aim, the numerical results, obtained on discretizing (8.31), demonstrate that one of the decomposed parts $u_2(x,t)$ consists of two discontinuous fronts; see at the top-right of Figure 8.7. We implement several numerical schemes to solve (8.31) and (8.32) in the high-fidelity sense. However, none of the schemes yield (only) one discontinuous feature in each of the two decomposed parts.*

It is well-known that the solution of Burgers' equation inherits sawtooth wave profile as its solution features after a finite-time, irrespective of the initial conditions. One such sawtooth wave profile for a specific choice of initial conditions is depicted at the top-left in Figure 8.7. The decomposed solution, obtained after performing spatial and temporal discretization on (8.31), looks as illustrated at the top-right in Figure 8.7. We can clearly observe that
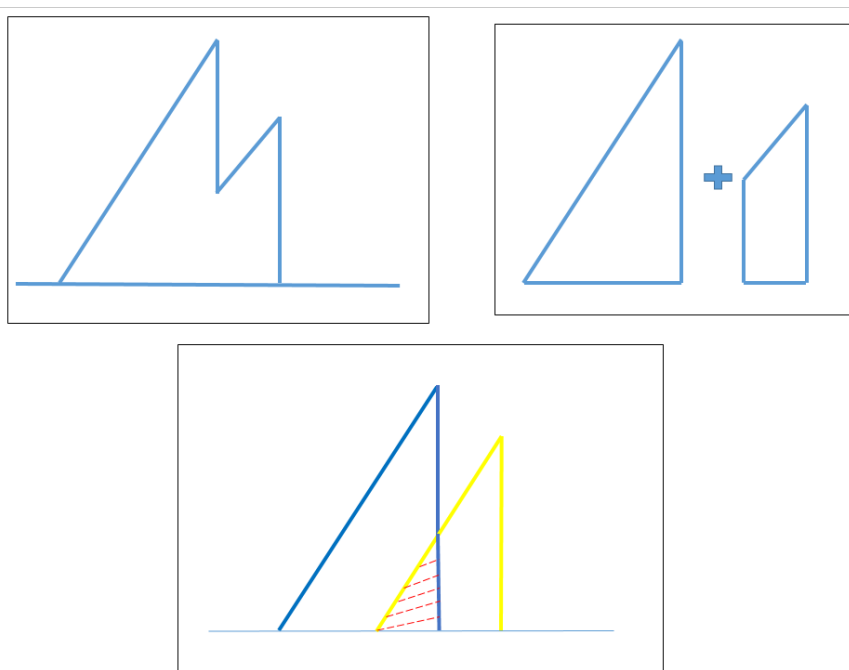
Figure 8.7: A schematic illustrating the idea in proposition 2.

a straight line connects the two evolving discontinuous fronts until the solution corresponding to the second decomposed part is eventually eaten up. We now suppose that we are primarily interested in the solution regime until the second decomposed part, i.e., the one to the right, vanishes. We put to use the afore-mentioned observation and seek to post-process the solutions obtained after the discretization. We only post-process the solution corresponding to the second decomposed part $u_2(x, t)$, and do not perform any post-processing on the first decomposed part $u_1(x, t)$. The post-processed solution can be written as follows:

$$u(x, t) = u_{11}(x, t) + u_{22}(x, t) - u_{33}(x, t), \tag{8.33}$$

 where all the three decomposed parts, i.e., $u_{11}(x, t), u_{22}(x, t)$ and $u_{33}(x, t)$ have a travelling nature, but constitute only one evolving front. Using the fact that we do not post-process the first decomposed part, we have that $u_{11}(x, t) = u_1(x, t)$. The other two decomposed parts $u_{22}(x, t)$ and $u_{33}(x, t)$ are obtained from $u_2(x, t)$ by following the steps as mentioned next. We first identify the slope of the straight line connecting the two sharp gradient features in the second decom-posed part $u_2(x, t)$; see at the top-right in Figure 8.7. We then extrapolate the straight line backwards and find the spatial coordinate x_m at which it passes the x-axis and allocate the extra region, i.e., a triangular region shown in red-dashed lines in Figure 8.7, to $u_{22}(x, t)$. It should be noted that the solution $u_2(x, t)$ is also a part of $u_{22}(x, t)$. Basically, the region under the yellow curve in Figure 8.7 is the decomposed part $u_{22}(x, t)$. It is observable that a small triangu-lar region, i.e., the red-dashed part in Figure 8.7, is accounted both in $u_{11}(x, t)$ and $u_{22}(x, t)$. Hence, this doubly counted region should be subtracted once and the solution corresponding to this overlapped region is interpreted as $u_{33}(x, t)$.

To summarize, the approach renders triangular features in each decomposed part. Such a decomposition is attractive from the view point of model order re-duction. The proposed approach helps to avoid multiple discontinuities in each decomposed part. Moreover, it should be noted that the post-processing is per-formed outside the temporal integration routine and, hence, is not intrusive. One of the limitations, however, is that the post-processing framework needs to know the time instant when the discontinuous front in $u_1(x, t)$ starts interact-ing with the tail of $u_2(x, t)$. Also, such an approach does not seem so easy to generalize in order to deal with both scalar and system of conservation laws.

Remark 8.3.4 *We would like to emphasize that the post-processing is not an ar-tificial way to solve the problem, and actually has a physical meaning. The post-processing can be interpreted as a flux re-distribution. One of the future research directions would be to come up with a more natural decomposition framework that*

automatically gives rise to the nice features without any need of post-processing. An automatic retrieval of the decomposition will also be attractive in the scope of obtaining a MOR framework and would thus mitigate the challenges which the current state-of-the-art approach to post-process poses.

Next, we combine the pre-processing ingredient, so-called "Method of Freezing", with standard reduced-basis approximation techniques to develop a complete MOR framework.

8.4 Combining Method of Freezing and reduced-basis approximations

In this section, we invoke the POD-Galerkin methodology on the transformed model problem governed by the system of PDAEs (8.12) - (8.14), and develop a reduced-order model.

Employing the first-order Lax-Friedrichs scheme for the spatial discretization of (8.12), (8.13) and (8.14), and the first-order Forward Euler scheme for the temporal discretization of (8.15), (8.16) and (8.17) yields a numerical approximation of the shape dynamics v_j for all $j = \{1, 2, \ldots, N\}$ at several spatial locations and time instants. The obtained approximation of the shape dynamics lives in a high-fidelity space, say $V_{j,h}$, where h in the subscript indicates the dimension of the high-fidelity model, and j stands for the co-moving frame under consideration. Having computed the high-fidelity approximation of the transformed model problem, we now construct a snapshot matrix of the shape dynamics corresponding to each co-moving frame. We then apply the method of snapshots (with respect to the natural inner product) on each of the computed snapshot matrices to obtain the basis functions or the POD-modes. We refer the reader to Algorithm 1, Chapter 7 for the technical aspects pertaining to POD/method of snapshots. Using the computed POD modes, we finally apply the Galerkin projection methodology, discussed in Chapter 7, as briefly described next.

We reduce the system of PDAEs (8.12) - (8.14) via Galerkin projection onto $V_{j,r} \subseteq V_{j,h}$, where $V_{j,r}$ is a r-dimensional reduced space spanned by the functions obtained from a truncated singular value decomposition of the snapshot matrix of the shape dynamics v_j corresponding to each co-moving frame. Upon considering the projection operator $P_{V_{j,r}} : V_{j,h} \to V_{j,r}$, the governing dynamics of

the reduced model for $j = \{1, 2, \ldots, N\}$ are given by:

$$\partial_t v_{j,r} = \mu_{j,r} P_{V_{j,r}} \partial_{\xi_j} (P_{V_{j,r}}^T v_{j,r}) - P_{V_{j,r}} \frac{\phi_j(\xi_j)}{\sum_{k=1}^M \phi_k(\xi_{kj}^g)} f\Big(\sum_{k=1}^N P_{V_{k,r}}^T v_{k,r}, \sum_{k=1}^N \partial_{\xi_k} (P_{V_{k,r}}^T v_{k,r}) \Big), \quad (8.34)$$

$$0 = <\partial_{\xi_j} (P_{V_{j,r}}^T v_{j,r}), \Big(-\frac{\phi_j(\xi_j)}{\sum_{k=1}^N \phi_k(\xi_{kj}^g)} f\Big(\sum_{k=1}^N P_{V_{k,r}}^T v_{k,r}, \sum_{k=1}^N \partial_{\xi_k} (P_{V_{k,r}}^T v_{k,r}) \Big) + \mu_{j,r} \partial_{\xi_j} (P_{V_{j,r}}^T v_{j,r}) \Big) >_{\mathscr{L}^2},$$
$$(8.35)$$

$$\partial_t g_{j,r} = \mu_{j,r}, \quad (8.36)$$

where $v_{j,r}, \mu_{j,r}$ and $g_{j,r}$ can be computed by again invoking the Lax-Friedrichs (or any other) scheme for the spatial discretization and the first-order Forward Euler (or its variants) for the temporal discretization. The reduced-solution $v_{j,r}$ can be lifted back to the high-dimensional space, and the travelling dynamics $g_{j,r}$ can be embedded to eventually obtain the solution to the PDE of interest.

Remark 8.4.1 *As mentioned in Remark 7.4.5, an effective offline-online decomposition is an important ingredient of reduced-basis approximations. However, its discussion is not in the scope of this work. Although we do not pursue this aspect in further detail, it is worth emphasizing that the operator*

$$\frac{\phi_j(\xi_j)}{\sum_{k=1}^N \phi_k(\xi_j - g_k + g_j)} \cdot f\Big(\sum_{k=1}^N v_k(\xi_{kj}^g, t), \sum_{k=1}^N v_{k,\xi}(\xi_{kj}^g, t) \Big) \neq f(v_j, v_{j,\xi}), \quad \forall j \in \{1, 2, \ldots, N\},$$
$$(8.37)$$

is not equivariant[3] due to the term $\frac{\phi_j(\xi_j)}{\sum_{k=1}^N \phi_k(\xi_j - g_k + g_j)}$. Recall from Section 8.2.1 that the decomposition, as in the "Method of Freezing" framework, relies on the concept of the partition of unity. However, as shown in (8.37), the partition functions hamper the equivariance property in the setting where multiple wavefronts evolve across the spatial domain. As a consequence, an effective offline-online decomposition becomes even more challenging, and it will be the subject of the research in the future.

---

[3]An equivariant operator is defined as follows:

$$g^{-1} \circ L \circ (g.v) = L \circ (v),$$

where L is any differential operator, $(g.v)$ represents the group action on the shape dynamics, g^{-1} stands for the inverse group action. In other words, the equivariance implies that a differential operator acting on the shifted solution is exactly equivalent to applying the differential operator on the original dynamics $(g.v)$ followed by shifting the resulting solution.

Remark 8.4.2 *For the reduced scheme governed by* (8.34) - (8.36)*, we need to compute partition functions to be able to obtain $v_{j,r}$ at every iteration while evolving the reduced-order model over time.*

Remark 8.4.3 *The pre-processing step, leading to the computation of the shape dynamics, helps to construct the POD modes which capture the underlying dynamical structure of a time-dependent problem.*

Remark 8.4.4 *Here, we only discussed about the construction of the reduced-order model starting with the general form of the transformed model problem governed by the system of PDAEs* (8.12) - (8.14)*. The (general) form of these PDAEs automatically accounts for the modifications due to Proposition 1, which has been introduced in 8.3.1. However, the idea of Proposition 2, introduced in 8.3.2, has a slightly different framework, and the corresponding reduced-order model needs to be separately constructed. We opt not to delve into the related discussion as the reduced-order model can be easily written mathematically by employing the POD-Galerkin methodology.*

Having discussed the mathematical construction of a reduced-order model with the proposed MOR framework, we now present numerical case studies dealing with several challenging scenarios encompassing different types of discontinuous wavefront interactions.

8.5   Numerical case studies

In this section, we numerically test the proposed MOR framework built by combining the approach of the "Method of Freezing" and the POD-Galerkin methodology, and also compare its performance with the standard (standalone) POD-Galerkin methodology. We focus on the problems in the field of fluid dynamics, in particular Burgers' equation. A series of numerical experiments are presented to verify and test the accuracy of the reduced-order model and, thus, illustrate the ability of the proposed MOR framework in dealing with multiple moving, interacting and possibly merging (discontinuous) fronts. We restrict to numerical experiments in a one-dimensional setting. Furthermore, we use a uniform discretization across space and time with spatial step size $\Delta x = 0.01$ and temporal step size $\Delta t = 0.0005$. Moreover, we employ the Lax-Friedrichs scheme for the spatial discretization, and the first-order Forward Euler for the temporal discretization. In the following numerical experiments, we consider only periodic boundary conditions. The initial data, required for the completeness of
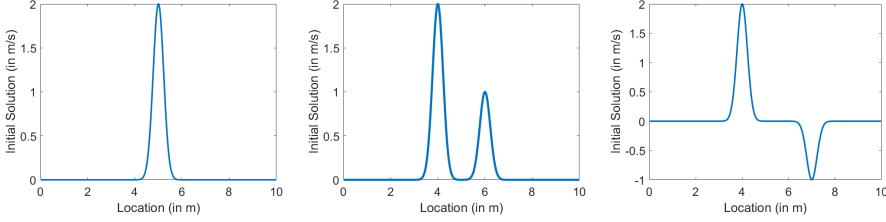
Figure 8.8: Initial conditions for several test cases.

the problem, is considered of the following form. Figure 8.8 (left) is an initial condition similar to the one used in [152, 154]. This experiment only involves setting up a single co-moving frame. Figure 8.8 (middle and right) are the additional initial conditions devised for specific reasons. In particular, Figure 8.8 (middle) helps to set up a scenario where two fronts are propagating along the same direction. However, the left front propagates with a higher speed and soon catches up with a slower moving front to its right, thereby undergoing interactions and, consequently, evolving as per the governing dynamics. Figure 8.8 (right) sets up a scenario where the two fronts propagate in opposite directions and, ultimately, interact. The test cases, corresponding to Figure 8.8 (middle and right), require the need of establishing multiple co-moving frames.

8.5.1   Single wavefront setting

In this section, we demonstrate the numerical results for the test case illustrated in Figure 8.8 (left). We solve the system of PDAEs (8.12) - (8.14) with $f(u, \partial_x u) = \partial_x(\frac{u^2}{2})$ and $N = 1$, and collect the snapshot data of the shape dynamics until $t = 2$, $t = 4$ and $t = 6$, i.e., for $4000, 8000$ and 12000 timesteps. We also collect the snapshot data of the original dynamics obtained by solving Burgers' equation (8.18) until $t = 2$, $t = 4$ and $t = 6$. We build the snapshot matrices from the collected snapshot data. We then employ the SVD approach and perform POD (with respect to the natural inner product) on these snapshot matrices to compute the POD modes and the associated singular values. Figure 8.9 depicts the behaviour of the decay of the computed singular values. The large difference between the decay behaviour of the singular values corresponding to the original dynamics (depicted in dashed dotted lines) and those corresponding to the shape dynamics (shown by using solid lines) demonstrates that most of the bases are actually required to capture the travelling dynamics. It also indi-
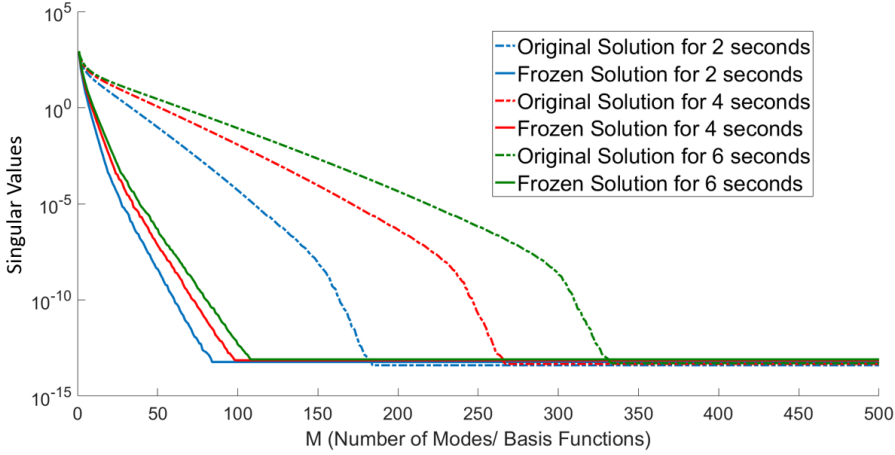
Figure 8.9: Behaviour of Singular Value Decay for test case with initial conditions as illustrated in Figure 8.8 (left).

cates that very few bases functions are required to capture the shape dynamics. This figure, thereby, shows the power of the pre-processing framework, i.e., the "Method of Freezing", in facilitating more effective model reduction.

8.5.2   Multiple wavefront setting

In this section, we discuss the test cases with initial conditions as illustrated in Figure 8.8 (middle) and Figure 8.8 (right).

We first discuss the test case with initial conditions as illustrated in Figure 8.8 (right). In view of the choice of the initial conditions and the discussion in Section 8.2.3, we choose to employ three co-moving frames, i.e., $N = 3$. We also use the flux re-distribution approach introduced in Section 8.3.1, and solve the system of PDAEs (8.12) - (8.14) with $f(u, \partial_x u) = \partial_x(\frac{u^2}{2})$. Contrary to the observations, discussed earlier in Section 8.2.3, we do not witness numerical instabilities as soon as the partition functions start overlapping in the "true" physical space. However, numerical instabilities in the high-fidelity computations of the system of PDAEs (8.12) - (8.14) arise just when the two fronts in the "true" physical space align themselves. In other words, we observe (high-frequency) numerical oscillations when the two fronts propagating from oppo-

site directions merge into each other and evolve as a single front. To summarize, the flux re-distribution approach, proposed in Section 8.3.1, helps to delay the onset of stability issues, but cannot deal with the scenario where the (discontinuous) fronts merge and evolve as a single entity. In view of this observation, the "Method of Freezing" cannot be used as a pre-processing framework after the wavefronts merge. Hence, we will limit our subsequent computations to time $t < t^*$, where t^* indicates the time instant when the fronts fully merge.

We recall that we seek to develop a reduced-order model as explained in Section 8.4. To this end, we collect the snapshot data of the shape dynamics in each co-moving frame until the point of collision of the two wavefronts. We then perform POD with respect to the natural inner product on these snapshot matrices by employing the SVD approach to compute the POD modes and the associated singular values. We also solve Burgers' equation (8.18) together with the chosen initial conditions, and build the snapshot matrix from the computed solution u. Then, as discussed in Chapter 7, we perform the SVD with respect to the natural inner product to compute the POD modes and the associated singular values. Although multiple wavefronts propagate across the spatial domain due to the considered initial condition, we seek to observe the role of the choice of the number of co-moving frames. To this end, we collect the snapshot data of the shape dynamics, use the SVD (with respect to the natural inner product) to compute the POD modes and the associated singular values for the setting where a single co-moving frame is considered, i.e., $N = 1$, instead of setting up multiple co-moving frames. Multiple ways can be adopted to build such a single co-moving frame for problems with multiple (discontinuous) wavefronts propagating with different speeds. Several possibilities of building a single co-moving frame are as follows: *(i)* the phase conditions can be devised to develop a co-moving frame corresponding to the slow evolving wavefront, *(ii)* the phase conditions can be tailored to develop a co-moving frame corresponding to the fast evolving counterpart, and *(iii)* the phase conditions can be left free without actually tailoring them for a specific evolution. However, either of the aforementioned ways of constructing only a single co-moving frame do not factor out the travelling component completely. One of the wavefronts, out of the resulting two due to the choice of the initial conditions, either slows down or speeds up. Without tailoring the co-moving frame to evolve as per the slower wavefront or as per the faster wavefront, we let the phase conditions themselves decide the frame speed and, consequently, the shifts. Next, we discuss the numerical observations in the scope of the possible reduction in the dimension of the high-fidelity model both for $N = 1$ and $N = 3$, and also compare with the reduction offered by the standard (standalone) POD-Galerkin methodology.
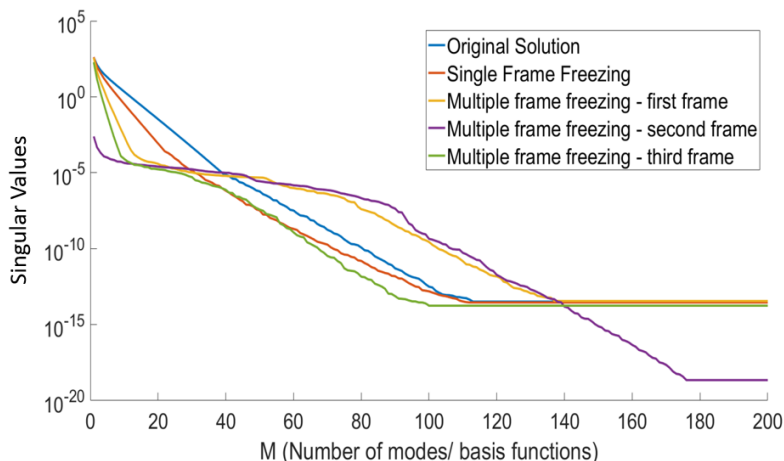
Figure 8.10: Behaviour of Singular Value Decay for test case with initial conditions as illustrated in Figure 8.8 (right).

Figure 8.10 depicts the behaviour of the decay of the singular values, computed for the different settings described in the previous paragraph, versus the number of the POD modes. We observe that the decay of the singular values, obtained with the choice of a single co-moving frame, is better than the decay observed for the standard setting where the "Method of Freezing" is not employed before invoking the POD-Galerkin methodology. Figure 8.10 also reveals that the shape dynamics, in the context where multiple co-moving frames are set up, can be represented by a relatively small number of basis functions compared to using only a single co-moving frame. However, it can also be observed that after a certain threshold accuracy level (associated to singular values of the order of 10^{-5} in Figure 8.10), the shape dynamics corresponding to co-moving frames 1 and 2 require large number of basis functions for an efficient representation. We observe that the decay of the singular values, corresponding to the shape dynamics in the co-moving frames (particularly frame 1 and 2), stagnates around certain (but different) levels, which are well above the zero machine precision level. This at first may seem to convey a poor performance of the proposed approach compared to the decay observed with the standard projection-based MOR framework. We can, however, argue against this view-

point by using the following reasoning. We know that the high-fidelity solution is accurate up to a certain level that is consistent with the order of the discretization error, and we cannot expect the reduced-order representation to be any better than its high-fidelity counterpart. The stagnation in the decay of the singular values, corresponding to the snapshot data of the shape dynamics, can also be interpreted from a different perspective. It is well-known that expressing the mode decomposition in different velocity frames, other than the "true" co-moving frame, requires many modes to describe the dynamics reasonably well. This instigates to investigate whether the high frequency modes are well resolved by the phase conditions. We investigated and found out that the high frequency modes are well resolved in their respective frames and, hence, an incorrect frame speed cannot be the reason behind the stagnation in the decay of the singular values. The stagnation in the decay behaviour of the singular values can potentially be attributed to the presence of sub-scale moving/stationary, non-differentiable continuous features. These type of features have been observed to be macroscopically present in the shape dynamics corresponding to each co-moving frame; see Section 8.2.3. Tuning the partition functions, as discussed in Section 8.2.3, probably makes the aforementioned features less apparent on the macroscopic scale, but does not completely remove the unwanted moving/stationary, non-differentiable continuous features.

Figure 8.11 depicts the absolute ($\mathscr{L}^2$ in space and $\mathscr{L}^1$ in time) error, computed between the high-fidelity "truth" solution and the reduced solution both with and without the "Method of Freezing", versus the number of the POD modes. The benefits of the POD-Galerkin approximation combined with the "Method of Freezing" are quite evident as we see orders of magnitude improvement in the approximations, though until a certain threshold accuracy level. It is also observed that the MOR approximation error behaves somewhat similar to the decay profile of the singular values.

Remark 8.5.1 *One could opt for different norms in space or time, as per the physical reasoning, while computing the MOR approximation error.*

We now briefly discuss the test case with initial conditions as illustrated in Figure 8.8 (middle). The naive flux re-distribution approach, introduced in Section 8.3.1, does not work for the numerical setting under consideration, and yields unwanted features in the high-fidelity computation of the shape dynamics in each co-moving frame. These unwanted features are similar to the features discussed in Section 8.2.3 in the scope of the case study on Burgers' equation with the initial condition illustrated in 8.5. As a consequence of the aforementioned features, we cannot hope to develop a reduced-order model
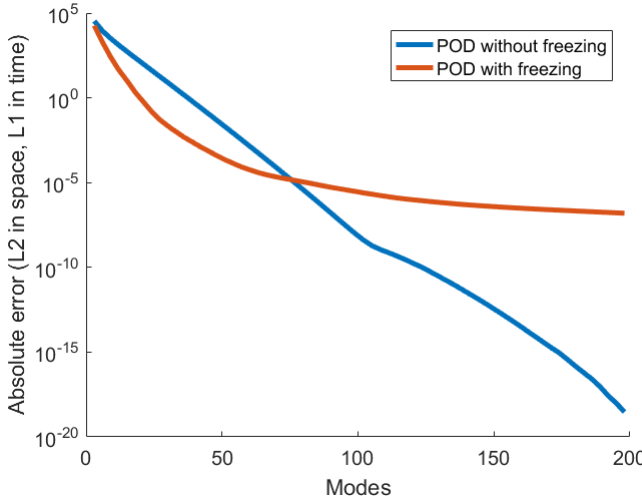
Figure 8.11: Test case with initial conditions as illustrated in Figure 8.8 (right): Absolute ($\mathscr{L}^2$ in space and $\mathscr{L}^1$ in time) error, computed between the high-fidelity "truth" solution and the reduced solution both with and without the "Method of Freezing", versus the number of the POD modes.

with significantly smaller dimension. We also recall that although the approach (based on Proposition 1) has been shown to work well for the setting where the wavefronts propagate in opposite direction, it does not fare well once the point of singularity is attained, i.e., the fronts merge. These limitations pose the need of developing a more generic framework. To this end, Proposition 2, discussed in 8.3.2, is assessed. We particularly consider a numerical experiment in the scope of dealing with the setting where the wavefronts propagate along the same direction and undergo strong/weak interactions. We consider the scenario with the initial conditions as illustrated in Figure 8.8 (middle). We compute the high-fidelity numerical solution of the system of PDAEs (8.12) - (8.14) with $f(u, \partial_x u) = \partial_x(\frac{u^2}{2})$ and $N = 3$ until $t = 3.5$. Following the same principles as before, we compute the singular values corresponding to original dynamics and the shape dynamics in each co-moving frame. Figure 8.12 depicts the behaviour of the decay of the singular values. It demonstrates the benefits of the framework of the "Method of Freezing" as we can clearly observe that the singular values of the sum of all the shape dynamics decay faster than the decay
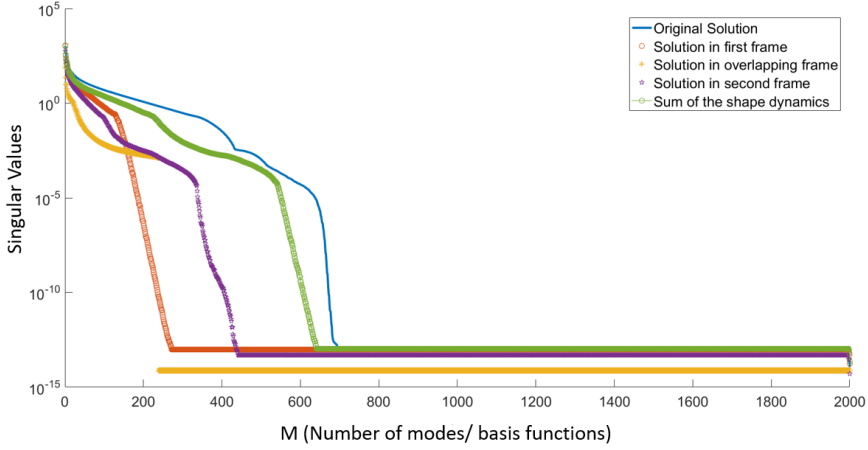
Figure 8.12: Behaviour of Singular Value Decay for test case with initial conditions as illustrated in Figure 8.8 (middle).

of the singular values corresponding to the snapshot matrix constituted of the original (travelling) dynamics. Considering the decay behaviour of the singular values of the sum of all the shape dynamics is a conservative way to quantify the performance of the proposed approach. The actual improvements offered will be better than the standard POD-Galerkin approach by an even better factor. Although the decay behaviour seems promising, it is worth mentioning that the approach, in the scope of Proposition 2, is quite specific to the Burgers' case, and the possibility to generalize it is subject to question/further research.

8.6   Conclusions

In this chapter, we made investigations in the direction to develop a novel MOR framework in order to deal with multiple moving, interacting and possibly merging discontinuous wavefronts. The proposed MOR framework relies on two concepts: *(i)* the "Method of Freezing", and *(ii)* the POD-Galerkin methodology. The "Method of Freezing" is used as a pre-processing ingredient that provides transformed model variables that are (supposedly) easy to reduce. This method can also yield unwanted features and the numerical instabilities in

the high fidelity computations of the transformed problem under wavefront interactions and collisions. This renders the computations of no use in the scope of developing a reduced-order model. In order to mitigate the aforementioned issues, we propose a flux re-distribution principle and modify the standard setting of the "Method of Freezing". The proposed MOR framework, which accounts for the flux re-distribution as introduced in Sections 8.3.1 and 8.3.2, do perform better than the standard POD-Galerkin methodology. However, none of the proposed approaches for flux re-distribution are able to deal with the scenario where the wavefronts merge and evolve as a single front. Also, the developed framework seems to work for specific test cases, and lack an easy generalization to be able to encompass a relatively larger class of hyperbolic problems admitting wavefront propagation, interaction and collision.

It is worth recalling that we tried to decompose the solution into several parts even for the setting where the fronts merge and evolve as a single entity. One may question the need of setting up multiple co-moving frames given the fact that we know how the original solution to Burgers' equation evolves after the wavefronts merge and that a single co-moving frame might be sufficient. Also, the more the decomposed parts, the larger is the system of Partial Differential Algebraic Equations (PDAEs) that need to be solved. Solving a system larger than necessary can also be questioned from the viewpoint of the requirement of saving computational resources. One of the proposed idea, henceforth, could be to switch from the setting of a multiple co-moving frame to that of a single co-moving frame upon merging of the wavefronts. However, such an approach calls for an a-priori knowhow about the evolution of the original dynamics. This is not always feasible for hyperbolic partial differential equations, which are very sensitive to the choice of the initial and boundary conditions. Moreover, unlike the case of Burgers' equation, the wavefronts propagating from opposite directions, for instance, will not always merge to become a single wavefront, particularly in the scope of the system of conservation laws. Various physical scenarios are possible where instead of merging and evolving as a single entity, the fronts would actually continue propagating with different strengths and topology after passing through the region of strong interaction. The need for an adaptive/automated framework to devise the number of co-moving frames (hence, the (high-fidelity) system size) is quite evident from the above discussion.

Furthermore, we recall that the mathematical framework and the numerical case studies, discussed in this chapter, impose the restriction of periodicity of boundary conditions. In view of implementing the pursued MOR framework for real industrial scenarios, such as the drilling applications, it is essential to ex-

tend the functionality such that the effects of non-periodic boundary conditions can also be accounted for.

To conclude, there still does not exist a MOR framework that is suitable and generic enough to deal with the scenarios where the (discontinuous) wavefronts interact and merge, and that accounts for the non-periodic boundary conditions. In the next chapter, we strive to develop a MOR framework that resolves the former open question.

Chapter 9

Model order reduction for problems with moving discontinuities

In this chapter, the focus is to develop efficient reduced-order model representations for (non-linear) transport-dominated problems and hyperbolic partial differential equations. In particular, we aim to develop a novel model order reduction (MOR) framework for tackling problems with multiple moving, interacting and possibly merging discontinuous features. The main ingredient of the proposed approach is a novel decomposition of the solution into a function that tracks evolving discontinuities and a residual part that is devoid of shock features. This decomposition ansatz is then combined with Proper Orthogonal Decomposition applied only to the residual part to develop an efficient reduced-order model representation for the problems of interest. Numerical case-studies show the potential of the approach in terms of computational accuracy compared with standard MOR techniques.

9.1 Introduction

We recall from Chapters 6 - 8 that moving discontinuities, such as shock-fronts, and interactions between the propagating wavefronts are representative features of models governed by scalar or systems of non-linear hyperbolic partial differential equations (PDEs). As mentioned/observed in Chapters 7 - 8, such

moving, interacting and merging discontinuous wavefronts pose a major hindrance to obtain effective reduced-order model representations; also see [135]. Although we made progress in the scope of developing reduced-order model representations for problems with multiple moving wavefronts in Chapter 8, we do not yet have a model order reduction (MOR) framework that is suitable and generic enough to particularly deal with the scenarios where the (discontinuous) wavefronts interact and merge.

We recall from Chapter 8 that the complete MOR framework, built upon the "Method of Freezing", depends on the choice of the partition functions and the number of co-moving frames (see Sections 8.2 and 8.3.1), or relies on some kind of a post-processing (see Section 8.3.2) to obtain features that are easy to reduce. On the one hand, the choice of ingredients, such as the partition functions, the number of co-moving frames, etc., do not have a sound theoretical basis and need to be tuned for any problem of interest. On the other hand, the post-processing based framework lacks the (online-efficient) automated identification of the switching point from a multiple wavefront setting to a single wavefront setting upon the merging of the wavefronts, and is quite specifically tailored to reduce the Burgers' equation. We draw insights from the "Method of Freezing" and propose an approach that is a stepping stone towards resolving the aforementioned issues.

The main contribution of this chapter is to propose a new decomposition ansatz that separates the solution into a basis function that tracks the evolving discontinuities and a residual part that is expected to be devoid of shock features. This decomposition renders the residual part to be amenable for reduced-order approximation. We then use these generated bases to apply Proper Orthogonal Decomposition (POD) on the residual part and later reconstruct the solution by lifting it to the high-dimensional problem space. We finally assess the combined performance of decomposition, reduction and reconstruction approach (as opposed to conventional reduction and reconstruction approach) in the scope of transport-dominated problems with moving, interacting and merging discontinuities.

The outline of this chapter is as follows. In Section 9.2, we provide the mathematical formulation of the proposed approach. Section 9.3 is devoted to numerically assess the performance of the proposed approach on several challenging test cases and compare it with the existing standard projection-based MOR approach. Finally, Section 9.4 ends the chapter with conclusions.

9.2   Mathematical Formulation

We consider a scalar one-dimensional (1D) conservation equation of the form:

$$\partial_t u(x,t) + \partial_x f(u(x,t)) = 0, \quad x \in \Omega = [0,L], \quad t \in [0,t_f], \tag{9.1}$$

with $u(x,0) = u_0(x)$ and suitable boundary conditions. Here Ω stands for the spatial domain of interest, L stands for the length of the spatial domain, and t_f indicates the final simulation time.

We assume that $u(x,0) = u_0(x)$ already has S number of discontinuities at locations $x_1(0),...,x_S(0)$ with values $u^-(x_s(0),0)$, $s = 1,...,S$ from the left and values $u^+(x_s(0),0)$, $s = 1,...,S$ from the right. We associate a single basis function $\sigma_s(x - x_s(t))$ to each discontinuity at their respective locations. This basis function has a jump of height 1, i.e., $\sigma_s^+(0) - \sigma_s^-(0) = 1$, at the location of the discontinuity and can have any (preferably continuous and smooth) shape away from the discontinuity.

We now decompose the solution of (9.1) in the following way:

$$u(x,t) = \sum_{s=1}^{S} j_s(t)\sigma_s(x - x_s(t)) + u_r(x,t),$$

where

$$j_s(t) = u^-(x_s(t),t) - u^+(x_s(t),t). \tag{9.2}$$

If $x_s(t)$ exactly matches the shock locations and (9.2) is exactly fulfilled, then $u_r(x,t)$ does not contain any discontinuities and is amenable to a low-rank approximation.

The time-stepping scheme is defined in the following way. In each time step, we:

- Compute the updated shock locations $x_s(t^{n+1})$ at $n+1$-th time instant t^{n+1} for all $s = 1,...,S$ using the Rankine-Hugoniot condition [177].

- Compute $u^\pm(x_s(t^{n+1}), t^{n+1})$ in a neighborhood of $x_s(t^{n+1})$ and define the jumps, $j_s(t^{n+1})$, via (9.2) for all $s = 1,...,S$.

- Compute the residual part $u_r(x, t^{n+1})$ from

$$u_r(x, t^{n+1}) - u_r(x, t^n) = \sum_{s=1}^{S} j_s(t^n)\sigma_s(x - x_s(t^n)) -$$

$$\Delta t \partial_x f(u(x,t^n)) - \sum_{s=1}^{S} j_s(t^{n+1})\sigma_s(x - x_s(t^{n+1})). \tag{9.3}$$

The standard way to construct a reduced-order model (ROM) is to reduce (9.1) by applying Galerkin projection on u; also see Chapter 7. Instead, we reduce (9.3) via Galerkin projection onto $V_N \subseteq V_h$, where V_N is a N-dimensional reduced space spanned by the functions obtained from a truncated singular value decomposition of the snapshot matrix composed of the high-fidelity computations of the residual part u_r, and V_h is a h-dimensional high-fidelity space. Upon considering the projection operator $P_N : V_h \rightarrow V_N$, the reduced scheme takes the following form:

$$u_{r,N}^{k+1} = u_{r,N}^k + P_N\Big(\sum_{s=1}^S j_{s,N}(t^k)\sigma_s(x - x_{s,N}(t^k)) - \Delta t \partial_x f(P_N^T u_N^k) -$$

$$\sum_{s=1}^S j_{s,N}(t^{k+1})\sigma_s(x - x_{s,N}(t^{k+1}))\Big), \quad (9.4)$$

where $u_{r,N}^k \in V_N$, the superscript k (resp. $k+1$) indicates the k-th (resp. $k+1$-th) time instant, u_N^k is defined in the following form:

$$P_N^T u_N^k = \sum_{s=1}^S j_{s,N}(t^k)\sigma_s(x - x_{s,N}(t^k)) + P_N^T u_{r,N}^k, \quad (9.5)$$

and, $j_{s,N}$ and $x_{s,N}$ are, respectively, the jumps and shock locations computed during the ROM time-stepping. The jumps $j_{s,N}$ and shock locations $x_{s,N}$ can be obtained in a manner similar to the steps carried out during the full-order model (FOM) time-stepping.

Remark 9.2.1 *It is well known that the projection alone is not sufficient to reduce the costs of computing the solution of a reduced-order model if the Finite Volume operators are non-linear in nature. The Empirical Operator Interpolation [52] method can be used here as a recipe for hyper-reduction. We do not delve into the full and efficient offline and online decomposition as its discussion is not within the scope of this work. However, we mention that we need to know $j_{s,N}(t^k)$ and $u_{r,N}(x_{s,N}(t^k), t^k)$ for computing $x_{s,N}(t^{k+1})$. In a reduced scheme this means that we need to keep the entire reduced basis in memory. However, the basis vectors are only evaluated at the shock locations at each time step. The same consideration holds for the computation of $j_{s,N}(t^{k+1})$.*

The proposed MOR approach is summarized below in an algorithmic man-

ner:

    Full-order model computation;

    initialization;

    while $t \leq t_f$ **do**

        Compute updated shock locations $x_s(t^{n+1})$ $\forall s = \{1, 2, \ldots, S\}$;

        Compute jumps $j_s(t^{n+1})$ $\forall s = \{1, 2, \ldots, S\}$;

        Compute the residual part $u_r(x, t^{n+1})$ by solving (9.3);

    end

    Perform POD on the residual part u_r **to obtain** P_N;

    Reduced-order model computation (Galerkin Projection);

    initialization using the built projector P_N, i.e., $u_{r,N}^0 = P_N(u_r^0)$;

    while $t \leq t_f$ **do**

        if *case==1* **then**

            Compute updated shock locations $x_{s,N}(t^{n+1})$ $\forall s = \{1, 2, \ldots, S\}$;

            Compute jumps $j_{s,N}(t^{n+1})$ $\forall s = \{1, 2, \ldots, S\}$;

            Compute the residual part $u_{r,N}(x, t^{n+1})$ by solving (9.4);

        else

            Use $x_s(t^{n+1})$ and $j_s(t^{n+1})$ computed during FOM time-stepping;

            Compute the residual part $u_{r,N}(x, t^{n+1})$ by solving (9.4);

        end

    end

Algorithm 2: Proposed MOR Approach.

9.3 Test case scenarios

We numerically test the new approach and show its potential as a MOR technique. We reduce Burgers' equation, which is given by:

$$\partial_t u + \partial_x(\frac{u^2}{2}) = 0, x \in [0, L]. \tag{9.6}$$

The case studies consider that the shock is already present in the initial data, which for single and multiple wavefront scenarios, is respectively given by:

$$u(x,0) = u_0(x) = \begin{cases} x, & 0 \leq x \leq 1, \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad u(x,0) = u_0(x) = \begin{cases} x - 2, & 2 \leq x \leq 4, \\ \frac{(x-5)}{2}, & 5 \leq x \leq 7, \\ 0, & \text{otherwise}. \end{cases}$$

Furthermore, we consider only periodic boundary conditions. We use an up-wind Finite Volume (FV) scheme for the spatial discretization and a first-order Forward Euler for the time-stepping.

We quantify the performance of the standard and the proposed approach by computing the reduced-order modeling (ROM) error. As described in Chapter 7, the ROM error can be defined in several ways. We refer the reader to (7.21), (7.22) and (7.23) for different definitions of the ROM error incurred due to the standard approach. In a similar manner, the ROM error incurred due to the proposed approach is defined as described below.

- $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time absolute ROM error is given by:

$$e_{rom} = \sqrt{\Delta t \sum_{k=1}^{N_T+1} \Delta x \sum_{i=1}^{N_x} | u_{i,k} - (P_N^T u_N^k)_{i,k} |^2}, \tag{9.7}$$

where Δt is the temporal step size, Δx is the spatial step size, N_T is the number of time-steps and N_x is the number of FV elements. And, $u_{i,k}$ stands for u at $x = x_i$ and $t = t_k$ (similarly for $(P_N^T u_N^k)_{i,k}$). Herewith, (9.7) expresses the error which is the difference between the high-fidelity numerical solution u and the reconstruction given by (9.5).

- $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time relative ROM error is given by:

$$e_{rom}^{rel(\mathscr{L}^2,\mathscr{L}^2)} = \frac{\sqrt{\Delta t \sum_{k=1}^{N_T+1} \Delta x \sum_{i=1}^{N_x} | u_{i,k} - (P_N^T u_N^k)_{i,k} |^2}}{\sqrt{\Delta t \sum_{k=1}^{N_T+1} \Delta x \sum_{i=1}^{N_x} | u_{i,k} |^2}}, \tag{9.8}$$

where the meaning of the notations is as explained earlier.

- $\mathscr{L}^2$ in space and $\mathscr{L}^\infty$ in time relative ROM error is given by:

$$e_{rom}^{rel(\mathscr{L}^2,\mathscr{L}^\infty)} = \frac{\max\limits_{k=\{1,2,...,N_T+1\}} \left(\Delta x \sum_{i=1}^{N_x} | u_{i,k} - (P_N^T u_N^k)_{i,k} |^2 \right)}{\max\limits_{k=\{1,2,...,N_T+1\}} \left(\Delta x \sum_{i=1}^{N_x} | u_{i,k} |^2 \right)}. \tag{9.9}$$

Single wavefront scenario

We first consider the scenario where only a single discontinuous front evolves

across the spatial domain. Here, we use the following shape function for $\sigma_s(x - x_s)$:

$$\sigma_s(x - x_s) = \begin{cases} 1 + x - x_s, & x_s - 1 \leq x \leq x_s, \\ 0, & \text{otherwise.} \end{cases} \quad , \quad s = 1, .., S, \qquad (9.10)$$

with $S = 1$, $x_s(t = 0) = 1$.

In the scope of this scenario, we consider the spatial domain to be of length $L = 10$. We take 8000 steps in time, i.e., $t \in [0, 4]$ with a temporal step size of 0.0005. We consider three different spatial mesh resolutions, i.e., spatial step size of 0.005, 0.002 and 0.001.

Multiple wavefront scenario

Here, we consider the setting where multiple (discontinuous) wavefronts evolve across the spatial domain and also interact non-linearly with each other. We particularly study the scenario where two wavefronts are present in the spatial domain and the left front propagates faster than the right one. We study this scenario under two different settings. We first restrict the study to only assess the performance of the proposed approach in dealing with the interaction of the head of one wavefront with the tail of the other one, and stop the simulation before the propagating fronts completely merge and subsequently evolve as a single front. We then consider a larger simulation run-time and check the ability of the proposed approach in automatically dealing with the merging of the wavefronts. The former setting is referred to as Case A, while the latter setting is referred to as Case B in the sequel.

We use the following shape function for $\sigma_s(x - x_s)$ to study this scenario for both Case A and Case B.

$$\sigma_s(x - x_s) = \begin{cases} 1 + \frac{1}{2}(x - x_s), & x_s - 2 \leq x \leq x_s, \\ 0, & \text{otherwise} \end{cases} \quad , \quad s = 1, .., S, \qquad (9.11)$$

with $S = 2$, $x_{s=1}(t = 0) = 4$ and $x_{s=2}(t = 0) = 7$.

In the scope of Case A, we consider the spatial domain to be of length $L = 10$. We take 8000 steps in time, i.e., $t \in [0, 4]$ with a temporal step size of 0.0005. We also consider three different spatial mesh resolutions, i.e., spatial step size of 0.005, 0.002 and 0.001.

In the scope of Case B, we consider the spatial domain to be of length $L = 12$. We take 20000 steps in time, i.e., $t \in [0, 10]$ with a temporal step size of 0.0005, and we consider a spatial step size of 0.002.

9.3.1   Numerical Results

Interpolation of $\sigma_s(x - x_s(t))$ onto the FV mesh results in numerical approximation error. As a result, we observe residual jumps in the residual part, $u_r(x, t)$, computed during FOM simulation. We recall from Algorithm 2 that the aim is to build a reduced space by applying POD on the residual part. One option could be to build the bases (or reduced space) from the computed residual part (with residual jumps). An other alternative could be to post-process the residual part (computed during FOM) in order to get rid of the residual jumps. Yet another alternative could be to compute the exact residual part analytically, which is feasible in the scope of scenarios of interest. The post-processed residual part or the exact residual part, which are even more low-rank approximable than the residual part with residual jumps, can be then used to build the (effective) reduced space. We invoke one of these ways to generate the bases and build a reduced-order model.

We first consider the single wavefront scenario and Case A in the scope of multiple wavefront scenario. We further consider the setting where the shock locations and jumps computed during FOM simulation are used during the ROM time-stepping, i.e., we assume that $j_{s,N} = j_s$ and $x_{s,N} = x_s$. Moreover, we use the computed residual part (with residual jumps) to generate the bases by performing POD with respect to the natural inner product and, subsequently, build a ROM. We can clearly see the benefits of the proposed approach in Figure 9.1, which shows the behaviour of the ROM error for increasing basis sizes N across different mesh resolutions. Firstly, the initial error incurred via the proposed approach is clearly lower than that of the standard approach. This is attributed to the fact that our decomposition approach associates a basis function corresponding to the travelling discontinuity. Secondly, the rate of decay of the error is better for the proposed approach compared to the standard approach. We also see that the ROM error for the standard approach is larger for finer mesh-sizes. This occurs as the effect of the shock becomes more pronounced for finer meshes. Also, the finer mesh implies less numerical viscosity. We also observe that the ROM error for the standard approach could even increase with an increment in the basis size. It can be argued that this could occur as a result of insufficiently many basis functions. However, the ROM error for the proposed approach decreases with an increment in basis size. Moreover, the ROM error is lower and stagnates later for finer mesh-sizes. This can be argued from the fact that the proposed approach is able to resolve the shock more accurately at finer meshes. This error behaviour is clearly in contrast to that of the standard approach which fails to efficiently capture the shock. As a result, the difference
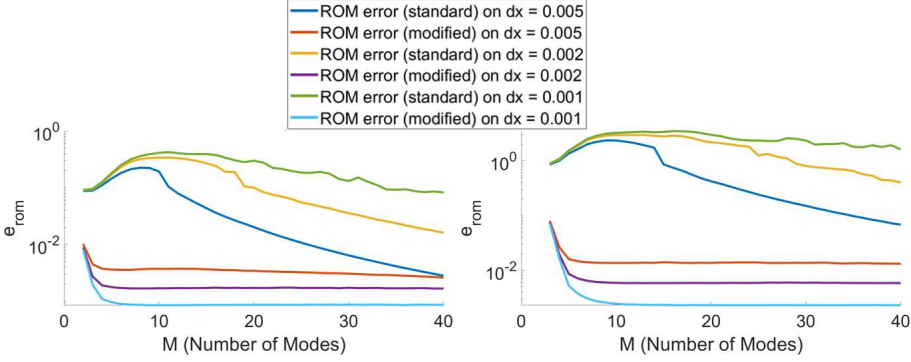
Figure 9.1: $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time absolute ROM error upon using shock locations and jumps computed during FOM simulation: (left) single wavefront scenario and (right) multiple wavefront scenario (Case A).

between the ROM error (at a certain number of basis function) computed via standard and proposed approach becomes more pronounced for finer meshes.

Figure 9.2 demonstrates the performance for fully ROM computations, i.e., shocks locations, $x_{s,N}$ and jumps, $j_{s,N}$ are computed during the ROM time-stepping. We perform post-processing on the residual part computed during FOM. The residual part u_r is post-processed by a linear interpolation between the locations x_s^+, x_s^- where the local minimum u^+ and maximum u^- in a neighborhood of x_s is attained. We then generate the bases from this post-processed residual part. The post-processing was not needed in an earlier setting (discussed in the paragraph above) as we used accurate shock locations and jumps computed during the FOM time-stepping. However, it becomes essential here in order to approximate $x_{s,N}$ and $j_{s,N}$ within the ROM time-stepping with good accuracy. We observe that the proposed approach still performs better than the standard approach. However, the proposed approach seems to incur larger ROM error for larger POD mode numbers. Such issues do not exist when we use the shock locations and jumps from FOM during the ROM time-stepping. Hence, the issues could be caused from a poor approximation of the shock. A possible explanation could be that we have more oscillations (around the shock position in the residual part) as the number of POD modes increases. The oscillations, which appear due to the reduced regularity of the residual part, lead to wrong computation of $x_{s,N}$ and $j_{s,N}$. It is clear that $x_{s,N}$ (and $j_{s,N}$) need to
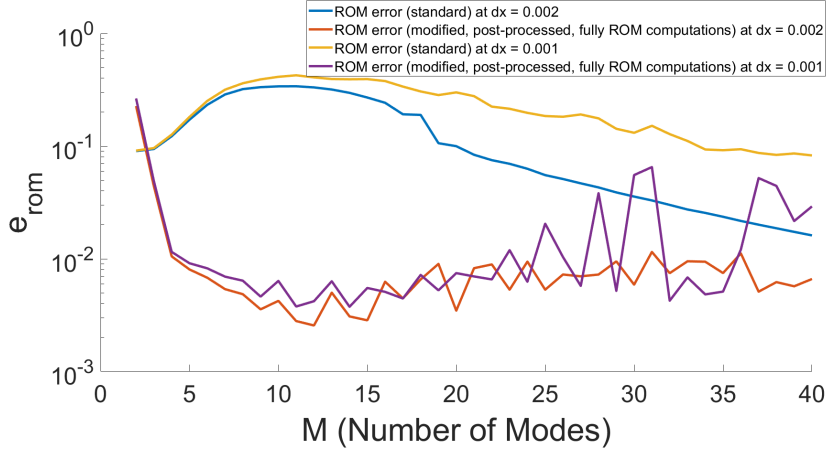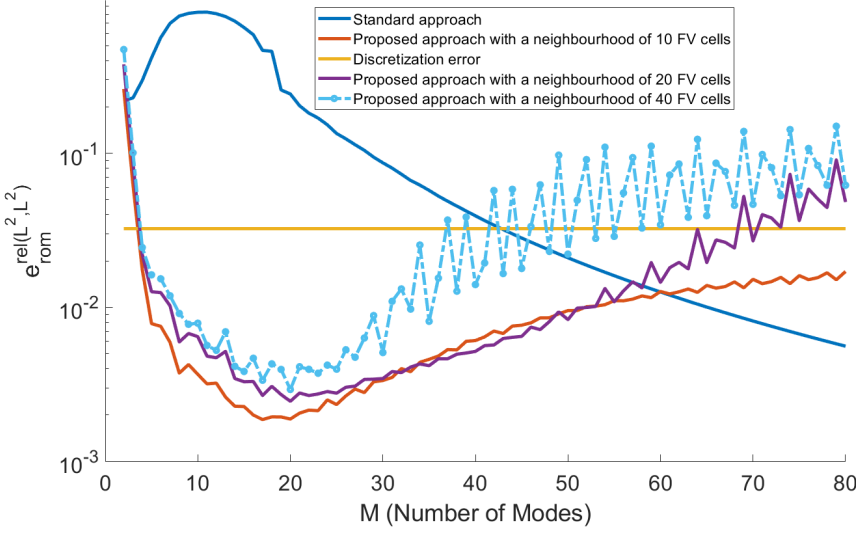
Figure 9.2: $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time absolute ROM error under fully ROM computations for the single wavefront scenario.

be approximated with good accuracy. The error in $x_{s,N}$, which would increase over time, should be in the order of the discretization error to achieve an overall ROM error in the order of the discretization error. The high-frequency modes could also be a source of the problem.

Next, we seek to identify the reasons behind the performance observed in Figure 9.2. To this end, we pursue the following thought process. Depending upon the spatial mesh resolution and the numerical viscosity of an implemented scheme for the (spatial) discretization, the shock features are known to be distributed over several FV cells. In order to compute the jumps j_s or $j_{s,N}$ for all $s = \{1, 2, \ldots, S\}$, a neighbourhood around the shock locations x_s or $x_{s,N}$ needs to be defined. It is intuitively clear that not any neighbourhood will yield correct computation of the jumps in the context of both the full-order model and the reduced-order model. Since there is no sound underlying theory for defining such a neighbourhood, we resort to numerically tuning the width of the neighbourhood around the shock locations and assess its impact on the computation of the jumps and, ultimately, on the quality of the ROM solution. Furthermore, among various ways to generate the bases, we choose to do so by applying POD (with respect to the $\mathscr{L}^2$ inner product) on the exact residual part. We do this in order to remove the (possible) sources of error due to post-processing of the computed residual part (with residual jumps). Another reason to use the
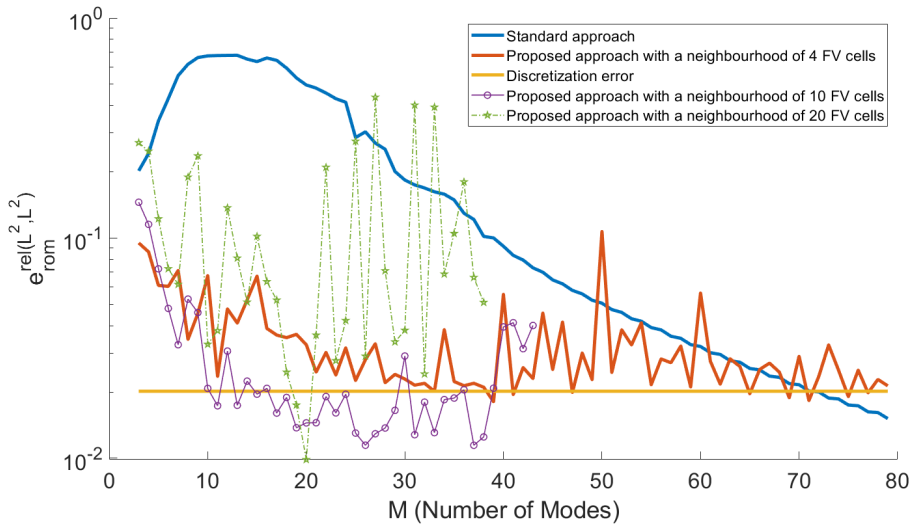
Figure 9.3: $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time relative ROM error under fully ROM computations for the single wavefront scenario.

exact residual part for generating the bases is that the computed residual part with residual jumps is less favorable for an efficient low-rank approximation. We also compute the discretization error and use it to argue the possibility of further improvements in the reduced-order models. Furthermore, as in the previous paragraph, we seek to develop a reduced-order model without resorting to any FOM computations, i.e., the counter "case" in Algorithm 2 is equal to one. Moreover, we consider the spatial step size $\Delta x = 0.002$ and the temporal step size $\Delta t = 0.0005$, and use the relative ROM error as defined in (7.22) and (9.8) to assess the performance of the standard and the proposed approach. The numerical observations in the scope of the single wavefront scenario and the multiple wavefront scenario (Case A and Case B) are discussed in the following paragraphs.

Figures 9.3 and 9.4, respectively, show the behaviour of the $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time relative ROM error versus the number of the POD modes for the single wavefront scenario and for Case A in the scope of the multiple wavefront scenario. In Figure 9.3, the ROM error behaves (qualitatively) similarly for the

first few number of the POD modes irrespective of the width of the neighbourhood around the shock location. The width of the neighbourhood around the shock location clearly has as impact on the error behaviour, particularly for a large number of the POD modes. Although the ROM error increases for higher POD mode numbers, it stays (to a large extent) lower than the discretization error, particularly for the case where 10 FV cells are employed in the neighbourhood of the shock location. Any behaviour below the discretization error is not of interest since it is known that if the ROM error is of the order of the discretization error, then the corresponding ROM is acceptable. In the scope of the single wavefront scenario, the choice of 10 FV cells in the neighbourhood can be considered to be (nearly) optimal as the corresponding ROM error, after showing an increasing behaviour, almost stagnates around a level which is well below the discretization error. For Case A in the scope of the multiple wavefront scenario, we consider 4, 10 and 20 FV cells in the neighbourhood around the shock locations. From Figure 9.4, we can observe that the ROM error obtained from the proposed approach with the choice of 4 FV cells in the neighbourhood is lower than the one computed from the standard methodology. Furthermore,
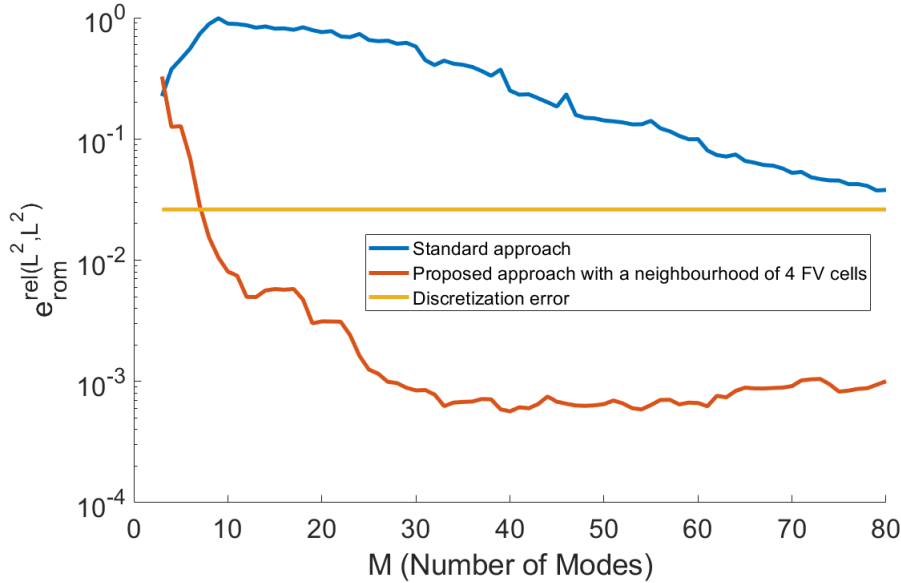

Figure 9.4: $\mathcal{L}^2$ in space and $\mathcal{L}^2$ in time relative ROM error under fully ROM computations for Case A in the scope of the multiple wavefront scenario.

this ROM error (in the proposed approach) decays to the level of the discretization error and then nearly stagnates around it. The choice of 10 FV cells in the neighbourhood yields a quite similar performance to that with 4 FV cells. Also, it gives a relatively better performance than that observed with 4 FV cells for POD mode number $M \leq 40$. However, the choice of 10 FV cells in the neighbourhood renders much higher (unacceptable) errors in the jump computations for a larger number of the POD modes, and destroys the quality of the ROM. Unlike the choice of 4 and 10 FV cells in the neighbourhood, invoking 20 FV cells in the neighbourhood results in a highly oscillatory behaviour of the ROM error, which also becomes (unacceptably) large for higher POD mode numbers. Clearly, the ROM constructed with 4 FV cells in the neighbourhood outperforms the other selections. Based on these observations, we will employ 4 FV cells in the neighbourhood for further discussions in the scope of the cases falling in the realm of the multiple wavefront scenario.

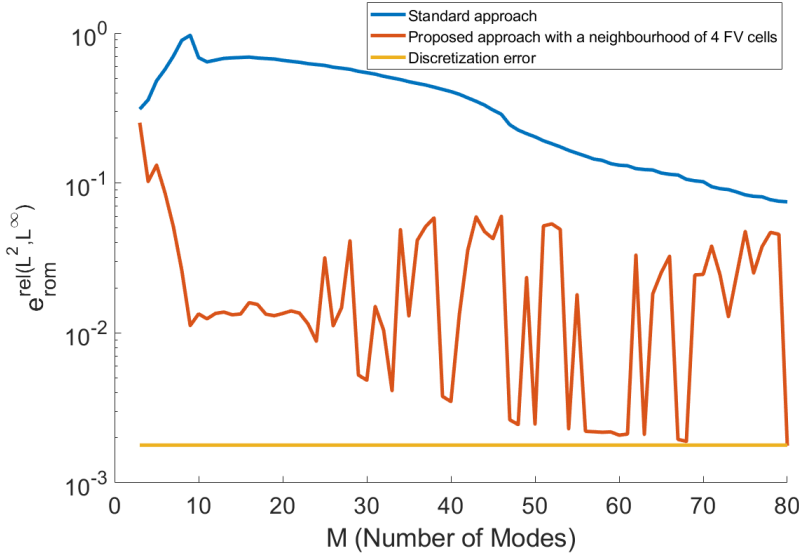We now consider Case B in the scope of the multiple wavefront scenario.

Figure 9.5: $\mathcal{L}^2$ in space and $\mathcal{L}^2$ in time relative ROM error under fully ROM computations for Case B in the scope of the multiple wavefront scenario.
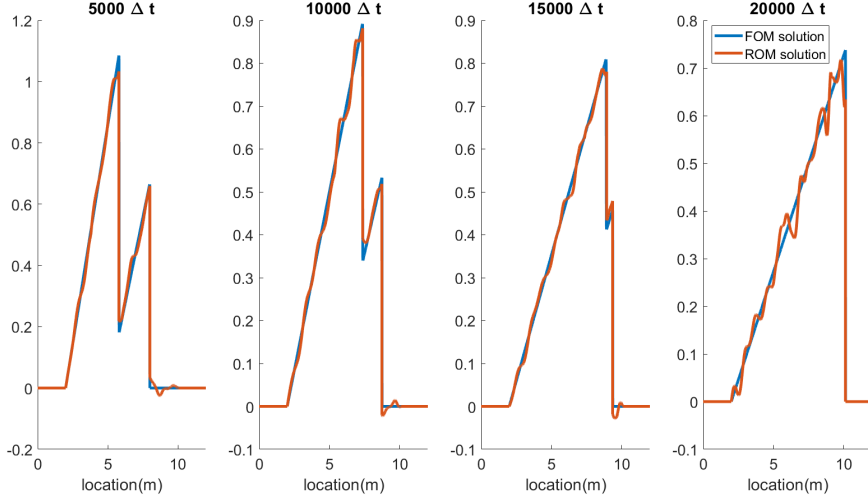
Figure 9.6: Evolution of the FOM and the ROM (obtained using the proposed approach) over equidistant time samples for Case B in the scope of the multiple wavefront scenario. Here, blue colour indicates the high-fidelity solution, and the red colour indicates the (reconstructed) solution of the reduced-order model with reduced dimension equal to 15.

Figure 9.5 demonstrates $\mathcal{L}^2$ in space and $\mathcal{L}^2$ in time relative ROM error versus the number of the POD modes. We can observe that the ROM error incurred in the proposed approach is lower than the one in the standard methodology. Also, the rate of the decay of the ROM error is better for the proposed approach compared to the standard one. Furthermore, the ROM error due to the proposed approach is observed to be below the discretization error just by using 7 or more POD modes. The ROM error due to the standard approach, on the other hand, does not reach to the level of the discretization error even with 80 POD modes. Consider that an orthogonal projector is built using 15 dominant modes, and subsequently a reduced-order model is constructed. Figure 9.6 shows the evolution of the solution obtained from the high-fidelity and the reduced-order model computation at 4 equidistant time instants, i.e., $5000\Delta t, 10000\Delta t, 15000\Delta t$ and $20000\Delta t$. The dimension of the reduced-order model has been considered to be equal to 15, which is a value around the small stagnation region after the sharp decay in the ROM error. Unlike the standard approach which gives highly

Figure 9.7: $\mathscr{L}^2$ in space and $\mathscr{L}^\infty$ in time relative ROM error under fully ROM computations for Case B in the scope of the multiple wavefront scenario.

oscillatory modes and is unable to resolve the shock, the proposed approach is able to (almost) resolve the shock even after the discontinuous wavefronts completely merge and evolve as a single entity. However, it can be observed that the jumps are not computed accurately. Furthermore, there are time-instants during the evolution where the error in the jumps are even larger than the ones observed in the evolution shown in Figure 9.6. If one is interested in the accurate computation of the jumps (and the shock locations), the evaluation of the $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time relative ROM error is not indicative enough as it overlooks this aspect. A better measure to quantify the performance in such a situation is to compute the $\mathscr{L}^2$ in space and $\mathscr{L}^\infty$ in time relative ROM error. Figure 9.7 depicts the behaviour of the aforementioned ROM error versus the number of the POD modes. The ROM error via the proposed approach is observed to again decay sharply for first few POD modes. It then stagnates for next few POD mode numbers, and finally oscillates around some stagnation level. It is clear that more than 60 POD modes are required to attain the level of the discretization error. Although the proposed approach fares better than the standard one, there is a potential to obtain much more effective reduction.

9.4   Conclusions

We have proposed a decomposition ansatz and used it in conjunction with the POD-Galerkin methodology. We have show-cased the performance of the standard and the proposed approach on Burgers' equation. Unlike, the standard approach, the proposed one is able to resolve the discontinuities and also offers (significant) reduction in the ROM error. We have also shown that a care should be taken in the selection of the "norm" that is used to compute the reduced-order modelling error. Overall, the results are promising in the scope of developing a model order reduction framework for problems with moving, interacting and merging discontinuities. However, issues exist in the proposed approach, particularly for larger POD mode numbers. These issues are to some extent associated to the selection of the neighbourhood around the shock location for the computation of the jumps. While the behaviour of the ROM error at higher POD mode numbers is attributed to the possible poor approximation of the moving shock (both location and the height/jump), the reason is still not fully clear, and needs to be further investigated in the future. A mitigating measure could be to improve shock approximation similar to [172] or to filter out the high-frequency modes when advancing the shock. Furthermore, in view of the numerical observations, a sound theoretical paradigm should be laid to identify an optimal neighbourhood around the shock locations. An inspiration could be drawn from the results on the width of the boundary layer for singularly perturbed problems.

In view of the promising results, particularly where the shock locations and jumps computed during the full-order model time-stepping are used, the proposed approach needs to also be assessed on other (parametrized) scalar and system of conservation laws. Furthermore, although we considered periodic boundary conditions, the mathematical formulation can easily be adapted to deal with the realistic non-periodic boundary conditions. Such aspects should also be considered in future works.

One of the limitations of the proposed approach is that the initial data is required to be non-smooth. Such a requirement is quite restrictive. It should also be mentioned that although the "Method of Freezing" with reduced-basis approximations suffers under wavefront interactions, it is able to deal with the smooth initial data (see Chapter 8). A further study, using the principles of the "Method of Freezing" and the ones proposed in this chapter, is required to develop a model order reduction framework that can handle both the smooth/non-smooth initial data and the merging/interaction of the wavefronts.

Chapter 10

Conclusions and Recommendations

10.1 Conclusions

To conclude, in this dissertation, we developed a much needed framework for modelling, (structure-preserving) discretization and model order reduction in the scope of drilling automation. In Chapter 2, we introduced Managed Pressure Drilling (MPD) systems and provided physical insights, particularly highlighting the properties that need to be captured by the high-fidelity and reduced-order models. The rest of this dissertation has been divided into two main parts. In Part I (Chapters 3 - 6), we discussed about modelling and (structure-preserving) discretization of hydraulic systems in the scope of MPD applications. These hydraulic systems, governed by non-linear hyperbolic partial differential equations (PDEs), pose several theoretical and computational challenges. With these open scientific challenges in mind, we provided answers to research questions Q1 - Q5 formulated in Chapter 1. To this end, we met research objectives O1 - O7, which have been outlined in the introduction of this dissertation. In Part II (Chapters 7 - 9), we discussed about model-complexity reduction of (high-fidelity) hydraulic systems. In particular, we delved into model order reduction of hyperbolic partial differential equations characterized by dominant transport effects, and strove to find answers to several important open-ended questions. We addressed research questions Q6 - Q7 by making promising steps in the direction of achieving research objectives O8 - O10 that have been described in

Chapter 1. The main results of the two parts of this dissertation are summarized next in Sections 10.1.1 and 10.1.2.

10.1.1 Part I: Modelling and (structure-preserving) discretization of hydraulic systems

In Chapter 3, we developed port-Hamiltonian formulations of two different types of two-phase flow models: the Two-Fluid Model (TFM) and the Drift Flux Model (DFM) (without slip) across a *constant* geometrical cross-section of the drilling well. These port-Hamiltonian formulations are developed by, firstly, deriving dissipative Hamiltonian representations of the TFM and the DFM (without slip). Secondly, the obtained dissipative Hamiltonian representations are extended with boundary port variables, and the *state-dependent* Stokes-Dirac structures associated with the skew-adjoint Hamiltonian operators are derived. In this chapter, we also demonstrated that the DFM with the Zuber-Findlay slip conditions is not an energy-consistent model for two-phase flow. The port-Hamiltonian framework, developed in this chapter, is then extended to account for *spatially and temporally varying* geometrical cross-sections in Chapter 4. In particular, in Chapter 4, we developed port-Hamiltonian formulations of single-phase flow model and the DFM (without slip) across *spatially and temporally varying* geometrical cross-sections. It is observed that a special care needs to be taken while formulating port-Hamiltonian representations, particularly in terms of *non-conservative* state variables. Using the aforementioned results, a compositional port-Hamiltonian model is then constructed for MPD systems in Chapter 4. The single-phase flow model describing the dynamics inside the drill string, the drill bit model, and the two-phase flow model describing the dynamics inside the annulus are interconnected in a power-preserving manner. Furthermore, we accounted for the interaction between the reservoir and the wellbore. Composing the MPD system so as to obtain an aggregated port-Hamiltonian structure led to finally deriving the condition for the passivity of the composed MPD system. This derived condition is then numerically assessed on a specific drilling well to identify the set of admissible operating conditions in the scope of the drilling operations.

In Chapter 5, we developed a framework for the structure-preserving discretization of non-linear distributed-parameter port-Hamiltonian systems associated with (extended) *state-dependent* Stokes-Dirac structures obtained in Chapters 3 and 4. We also accounted for the presence and absence of dissipative/resistive effects while developing the discretization framework. We par-

ticularly employed the mixed finite element method for the spatial discretization of the TFM and the DFM (without slip), and obtained corresponding (extended) continuous-time, finite-dimensional Dirac structures. Subsequently, the (symplectic) Gauss-Legendre collocation methods are invoked to perform the temporal discretization while accounting for *state-dependent* system matrices and *feedthrough* terms in state-space port-Hamiltonian representations. This helps to obtain discrete-time, finite-dimensional Dirac structures. In the pursuit of developing the discretization framework, we derived the conditions that must hold for the existence of the continuous-time and discrete-time, finite-dimensional Dirac structures. Furthermore, the properties of finite-dimensional Dirac structures have been exploited to interconnect the (discretized) mesh cells in a power-preserving manner and, ultimately, obtain a complete finite-dimensional port-Hamiltonian structure. Finally, we assessed the discretization framework on a simple, yet representative test case.

We then forayed into the numerical analysis of the two-phase flow model, in particular the DFM, in Chapter 6. It is proposed to account for the compressiblity effects of the phases under consideration in the quest to obtain the model for the speed of sound in the two-phase mixture. We numerically studied the behaviour of the wave propagation (for, e.g., the location and speed of the wavefronts) under the usage of the surrogate sound speed model and the proposed one in the discretization framework based on the Finite Volume method. We tested the model and numerical methods with exact reference solutions, and also employed the non-linear error transport approach to estimate the discretization error. This, to the best of our knowledge, has been done for the first time in the scope of the DFM. Furthermore, several key numerical properties, for, e.g., whether the numerical schemes follow Abgrall's principle, are assessed for different challenging test cases. Finally, the discretization error estimates along with the performance metrics (e.g., the observed order of accuracy, preservation of key numerical properties, etc.) paved the road for establishing the relative order of merit of the numerical schemes.

10.1.2 Part II: Model order reduction for transport-dominated problems

In Chapter 7, we developed reduced-order models of non-linear hyperbolic PDEs by invoking standard projection-based model order reduction (MOR) technique, namely, the Proper Orthogonal Decomposition in conjunction with Galerkin projection. In particular, we performed MOR on Burgers' equation and a two-

phase flow model, namely, the DFM. It is shown that the standard projection-based MOR techniques offer limited success in dealing with hyperbolic PDEs due to their inherent *transport-dominated* nature. The challenges in significantly reducing the dimension of a high-fidelity counterpart and achieving acceptable computational accuracy are compounded in the presence of non-smooth features, such as shocks, which move and interact across the spatial domain. The numerical results point out the need of developing a novel MOR framework in order to significantly reduce the dimension of hyperbolic problems (with moving, interacting and merging discontinuities), while still maintaining an acceptable computational accuracy.

Chapters 8 and 9 dealt with the development of the much needed MOR framework in the scope of *transport-dominated* problems, in particular for the scenarios in which the (multiple) moving wavefronts hamper the efficiency of the standard projection-based MOR techniques. In Chapter 8, the "Method of Freezing" in conjunction with reduced-basis approximations is proposed as a novel MOR framework. We invoked the "Method of Freezing", a symmetry reduction framework, and employed it for dealing with hyperbolic PDEs. We also proposed flux re-distribution principles to resolve numerical issues, such as additional travelling structures and numerical instabilities upon the merging of wavefronts, observed in the high-fidelity computations of the transformed model problem obtained after employing the "Method of Freezing". The transformed model problem is then reduced by applying the standard projection-based MOR techniques. Although the proposed MOR framework yields promising results in the scope of reducing problems with multiple evolving wavefronts, it cannot deal with the scenario of the (full) merging of wavefronts. Furthermore, the aforementioned framework lacks generic applicability to a large class of hyperbolic PDEs/*transport-dominated* problems. This further instigated the need to develop an approach that can mitigate the open issues. To this end, we proposed a new MOR approach in Chapter 9. We decomposed the solution into a function that tracks the evolving discontinuities and a residual part that is devoid of shock features. We then applied the standard projection-based MOR techniques only on the residual part. This approach helps to develop suitable reduced-order model representations for problems with multiple moving and possibly merging discontinuous features. It is shown that the proposed approach is better than the existing standard approaches in terms of the attainable reduction in the dimension of the high-fidelity model and the computational accuracy. The role of the *norms* used to compute the reduced-order modelling error is also made evident by the reported numerical observations. For instance, it is shown that $\mathscr{L}^2$ in space and $\mathscr{L}^2$ in time (relative) reduced-

order modelling error can mask the possible errors in the computation of the *jump*, i.e, the height of the sharp gradient features, which eventually reflects in the form of mismatch with the "true" time response behaviour of the model of interest. $\mathscr{L}^2$ in space and $\mathscr{L}^\infty$ in time (relative) reduced-order modelling error is shown to be a more representative measure for error computations in the scope of the problems with discontinuities. Although the proposed approach is rather generic and works better than the existing approaches, we clearly see a potential in further improvements, which is deferred to future works.

Although the mathematical construction of the approaches developed in Chapters 8 and 9 is different, both of these exploit the underlying travelling characteristic of the problem. Qualitatively speaking, the approaches proposed in each of the aforementioned chapters have their own pros and cons as briefly outlined next. On the one hand, the "Method of Freezing" with reduced-basis approximations, proposed in Chapter 8, is able to handle the smooth initial data, but suffers to (generically and efficiently) deal with the scenarios where the wavefronts interact and merge. On the other hand, the approach proposed in Chapter 9 is quite effective at dealing with the merging and the interaction of the wavefronts, but is limited to deal with the non-smooth initial data. A dedicated study is required in the future to develop a generic MOR framework that can effectively handle any type of the initial data and also deal with the merging and the interaction of the propagating wavefronts.

10.2   Outlook and recommendations for future work

In this section, recommendations are presented for further research in the scope of modelling and structure-preserving discretization of hydraulic systems, and in the scope of model order reduction of transport-dominated problems. We provide recommendations in a broader sense both from the perspectives of theory and application. The recommendations are summarized below.

R1 In Chapters 3 and 4, we presented port-Hamiltonian formulations for single-phase and two-phase flow models accounting for flows across *constant* and *smoothly varying* geometrical cross-sections. We recall that the proposed/derived Stokes-Dirac structures are defined under the smoothness and differentiability constraints, for instance on the state variables, see Theorem 3.4.9, Theorem 3.4.11, Proposition 4.2.9 and Proposition 4.2.10. Such constraints are an outcome of the classical theory that hinges on the fundamental lemma of calculus of variations, i.e., Lemma 3.4.8,

and it is mathematically hard to define the Stokes-Dirac structures under non-classical evolution of states. However, it is well-known that single-phase and two-phase flow models can lead to shock fronts, rarefaction waves and contact discontinuities. Moreover, non-smooth variations in cross-sections are also admissible in real life applications, including MPD. All these aforementioned situations eventually result in an instantaneous jump in the state variables. Such a discontinuous evolution renders the classical definitions of Stokes-Dirac structures inapplicable. In other words, the structures defined in this dissertation are restrictive and fail to account for discontinuous behaviour (or an instantaneous jump). A parallel can, however, be drawn between switched dynamical systems and the problems with evolving discontinuities. Few major works in the direction of port-Hamiltonian representations for switched dynamical systems include [180, 183]. These kind of dynamical systems encompass the applications that require modelling fast transitions in physical systems, which include power converters, robotics, hydraulic networks, etc. There exist notions of switched Dirac structures. However, most of the work in the realm of switched Dirac structures has been performed for finite-dimensional dynamical systems. In the scope of the aforementioned challenge, it would be interesting to *extend the definitions of finite-dimensional switched Dirac structures to an infinite-dimensional setting and, ultimately, obtain infinite-dimensional Stokes-Dirac structures in the scope of a large class of system of hyperbolic PDEs*.

R2 An alternative thread could be pursued in order to meet the objective set out in **R1**. For instance, we know that the DFM can admit three discontinuous wavefronts, and that the solutions are smooth away from the location of discontinuities. Also, the relation between the states at the left and the right of the discontinuity can be obtained via Rankine-Hugoniot conditions, for instance, see [177] in the scope of Euler equations. The Rankine-Hugoniot principles can be used to find the relations between the left and right states of a discontinuity in the scope of the DFM (and also the TFM). Furthermore, we already know that Stokes-Dirac structures, derived in Chapters 3 and 4, hold in the regions away from the discontinuities. In the spirit of compositional port-Hamiltonian modelling pursued in Chapter 4, and with the knowledge in hand, the objective laid in **R1** would be achieved if a finite-dimensional Dirac structure can be shown to exist across the interface of a discontinuity. Similar ideas, though in the scope of extrusion process modelling, have been pursued in the past [50]. Ba-

sically, port-Hamiltonian systems have previously been coupled through fixed and moving interface(s) [49], and an inspiration can be drawn to prove the existence of the Dirac structure at the moving interface in various contexts. To summarize, the objective (also related to **R1**) would be to *study the dynamics at the interface and cast it into a finite-dimensional Dirac structure representation*.

R3 Furthermore, looking from another viewpoint but related to **R2** to some extent, it is known that a drilling scenario, called tripping, includes axial pipe movements, i.e., movements in the drill string. These axial movements, which lead to so called water hammer effects, are essential to be incorporated in the mathematical model governing the hydraulic dynamics. While we composed the sub-systems in the scope of MPD applications in Chapter 4, we did not account for the moving dynamics of the drill string. Encoding such moving dynamics into a port-Hamiltonian structure will take us a step closer to obtain a holistic modelling and control framework for MPD systems.

R4 In Chapters 3 and 4, we have derived the dissipative Hamiltonian representations of the DFM and the associated Stokes-Dirac structures under the assumption of zero slip conditions. It is also worth mentioning that some of our attempts in the scope of developing a dissipative Hamiltonian representation for the DFM with the Zuber-Findlay slip conditions have been futile. There is an impression that the DFM with the Zuber-Findlay slip conditions admits a port-Hamiltonian structure that holds conditionally. Such a conditional existence could potentially be linked to the conditional hyperbolicity of the DFM under the Zuber-Findlay slip conditions. In the first place, the applicability of the DFM under the loss of hyperbolicity is a serious (open) question. It could be argued that the model is no longer physical in the regime dictating the loss of hyperbolicity and, eventually, a corresponding port-Hamiltonian structure, which has physical premise, cannot be envisioned. It would be valuable to anyways *assess the link between the conditional hyperbolicity of the model and the conditional existence of a corresponding port-Hamiltonian structure in a rather rigorous sense*. Furthermore, it would be worth to *investigate whether the dissipative Hamiltonian representations and the associated Stokes-Dirac structures can be developed for the Zuber-Findlay slip conditions and general slip laws*. Alternatively, a potential research direction could be to *find the slip laws that are energy-consistent, and also render port-Hamiltonian representations*.

R5 In Chapters 3 and 4, the boundary port variables are described in a non-unique manner. In principle, it would be ideal to derive an admissible set of boundary conditions in a parametrized way similar to [101], where a parametrization was derived for a canonical skew-symmetric differential operator. A similar parametrization is, however, non-trivial for *state-dependent* Stokes-Dirac structures obtained for the TFM and the DFM without slip. As a consequence, *fundamental modifications/extensions to the theory in [101, 194] are needed to obtain parametrization of boundary port variables in the scope of non-linear distributed parameter port-Hamiltonian systems*.

R6 The theoretical development of the framework for the structure-preserving discretization constituted a major emphasis in Chapter 5. In the scope of the numerical analysis, we conducted only one case study and assessed the performance of the structure-preserving discretization framework, in particular of the structure-preserving spatial discretization methodology. Firstly, a *rigourous numerical assessment should be performed for supporting the laid theoretical foundations*. Secondly, the *topic of the numerical stability deserves a dedicated attention*. Recall that, under certain spatial and temporal mesh resolutions, we observed numerical instabilities after a finite number of time steps for the test case with uniform conditions and without any resistive/dissipative effects. We care to stress that the numerical case study considered the structure-preserving spatial discretization and employed non structure-preserving explicit time-integrators for the reasons mentioned in Chapter 5. One could associate the observed instabilities to non structure-preserving time-discretization techniques. However, there are definitely some phenomena occurring at a numerical level, which are potentially not due to the choice of time-discretization techniques, since the numerical instabilities are not observed at relatively coarser spatial mesh resolutions even for a larger temporal step size. Generically speaking, the finer the mesh, the better are the results, which eventually tend to converge to the "truth" (mesh-independent) solutions. However, this does not seem to be the case here, and the source of the aforementioned numerical phenomena needs to be better understood. It would be valuable to *derive an analogue of the CFL condition in order to quantify the regime of numerical stability in the scope of the structure-preserving port-Hamiltonian discretization framework*. While such a measure would largely be problem specific, the *development of a general theory for the numerical (in)stability, particularly for the (lossless) conservative setting,*

would aid significant advances in the port-Hamiltonian community. It is also worth mentioning that the numerical stability is not offered as 'a free lunch' if the structure-preserving spatial and temporal discretization techniques are employed. The quantity characteristic to port-Hamiltonian systems is the dissipation inequality or passivity, which is required to be preserved while developing finite-dimensional port-Hamiltonian representations. While the passivity and the stability can be closely related for the linear systems, it is hard to make any statements on the stability for a passive non-linear system. A theoretical advancement is needed to *fill the gap by coining a methodology to study the (non-linear) stability with an a-priori knowledge of the passivity of the system.*

R7 In Chapter 5, we essentially discretized (extended) Stokes-Dirac structures to obtain (extended) continuous-time, finite-dimensional and discrete-time, finite-dimensional Dirac structures. Since the Stokes-Dirac structures are associated to classical solutions, the initial conditions and the final simulation time is constrained to be chosen such that the solution still belongs to the admissible function class. However, such a numerical setting defers us from numerical simulations of interesting and relevant features (such as moving shocks) of two-phase flow models. Closely tied to the modelling aspect mentioned in **R1**, *extending the developed theoretical framework of structure-preserving discretization to be able to deal with moving sharp gradient features* would be of added value. Alternatively, it might be meaningful to *assess the developed theoretical structure-preserving discretization framework even on situations admitting non-smooth solution behaviour.* The results can be subsequently studied/analyzed numerically. In particular, it would be of significant interest to quantify the deviation of the obtained results from the exact or accepted "truth" solutions.

R8 In Chapters 3 and 4, we resorted to model reformulations by eliminating auxilliary variables to obtain a system of PDEs with as many unknowns as equations. In other words, we converted an implicit system of conservation and balance equations together with closure equations into an explicit system of conservation and balance equations. Subsequently, we developed port-Hamiltonian representations. However, such a methodology is not modular, for instance, model reformulations are a must and all the steps have to be repeated even for a small change in the type of equation of state or the slip law. Furthermore, the algebraic constraints (or closure laws) may not be exactly preserved in the discretized

(or reduced) models if the starting point of port-Hamiltonian representations is the aforementioned model reformulation. In the pursuit to develop a modular framework and preserve algebraic constraints, port-Hamiltonian descriptor realizations, i.e., port-Hamiltonian partial differential algebraic equations (pH-PDAEs), are deemed suitable. General representations of port-Hamiltonian differential algebraic equations (pH-DAEs) have recently been introduced in [123]. The authors in [123] also claim that this framework can be used to describe pH-PDAEs. However, *the existing framework in the scope of pH-PDAEs still has a room of generalization* since the two-phase flow models, such as the DFM, cannot be formulated in terms of pH-PDAE representations as in [123]. Furthermore, *structure-preserving discretization (and reduction) of pH-PDAEs is largely an open question*, and has not yet been taken up in the community. In the scope of this open challenge, *structure-perserving spatial discretization of pH-PDAEs needs significant attention* since other existing concepts can then be utilized to proceed with the discretization and, ultimately, finite-dimensional realizations.

R9 The framework developed in Chapters 3,4 and 5 should be exploited to *perform model order reduction, and controller and observer design in the scope of MPD systems, and other applications* where the studied single- and two-phase flow models are also employed.

R10 In scope of the DFM, we have either assumed zero slip between the phases or considered the Zuber-Findlay slip conditions; see Chapters 3 - 7. In other words, we have considered a specific flow regime throughout the wellbore. However, in principle, various flow regimes exist across different sections of the wellbore. This indicates that the flux function in the underlying hyperbolic PDE will be spatially dependent. Hence, overall, the flux function will depend on both state and its spatial derivative. Although some work has been done to develop numerical methods for spatially dependent flux functions for hyperbolic PDEs, such an aspect has not been considered so far in the scope of the TFM and the DFM. It would be worth to *incorporate the effect of multiple flow regimes across the wellbore (mathematically speaking, spatially dependent flux functions) to simulate a scenario closer to reality and, subsequently, to validate the results against field experiments*.

R11 In the scope of the MOR framework proposed in Chapter 8, it would be of significant interest to *pursue theoretical analysis pertaining to the design*

of partition functions, the choice of number of co-moving frames, and flux re-distribution principle. The vision of such a theoretical analysis would be to yield the decomposed solution features that are appealing from the model order reduction point of view. Such a study would also help to increase the level of automation in the proposed framework, i.e., fewer "tunable" parameters would have to be manually tuned, and the choice of "tunable" parameters would, in principle, have a sound theoretical basis. This study will also potentially serve as a stepping stone to the possibility of generalization of the framework.

R12 The MOR frameworks, proposed in Chapters 8 and 9, specifically deal with one-dimensional scalar conservation laws, and in particular Burgers' equation. Out of the two similar, yet different frameworks, the one proposed in Chapter 9 seems promising in dealing with multiple moving, interacting and merging scenarios. Hence, this framework deserves further investigation. Although the proposed framework fares better than the standard MOR framework, it is worth to *study the possibility to further reduce the dimension of the obtained reduced-order models under interaction/merger of wavefronts*. Also, the *role of the number of neighbouring mesh cells in the computation of the "jump" and, eventually, in the behaviour of the reduced-order model needs to be rigorously investigated theoretically*. Furthermore, it is of interest to *extend the framework in Chapter 9 to be able to deal with system of non-linear conservation laws, such as the DFM and the TFM, and also deal with multi-dimensional transport-dominated problems*. Moreover, in the quest of achieving significant cost savings, the *performance of the proposed framework should be assessed in conjunction with the Empirical Operator Interpolation*.

R13 The reduced-order models, developed in Chapters 7 - 9, possess an issue of stability apart from the ones related to the resulting computational accuracy and the (reduced) dimension. This occurs since the stability is not inherently taken into account in the discussed frameworks. The topic of stability is critical, particularly for problems with (moving) discontinuities and non-differentiable features. It would be worth to *develop a framework or to extend the one proposed in Chapter 9 such that the stability of the reduced-order model can also be guaranteed*. In the scope of this aim, it would be significantly valuable to *delve into the concepts of Discontinuous Petrov-Galerkin methods, which possess a nice property of stability, and use the idea to embed the stability in the reduced-order models*.

Bibliography

[1] *IADC Drilling Manual Vol 1 of 2*. Houston, Texas, USA., 12th edition, edition. (Cited on pages ix and 20.)

[2] U.J.F. Aarsnes. *Modeling of two-phase flow for estimation and control of drilling operations*. PhD thesis, NTNU, 2016. (Cited on pages 6, 37, 72, 160, and 161.)

[3] Ulf Jakob F. Aarsnes, Adrian Ambrus, Florent Di Meglio, Ali Karimi Vajargah, Ole Morten Aamo, and Eric van Oort. A simplified two-phase flow model using a quasi-equilibrium momentum balance. *International Journal of Multiphase Flow*, 83:77 – 85, 2016. (Cited on page 3.)

[4] Ulf Jakob F. Aarsnes, Adrian Ambrus, Ali Karimi Vajargah, Ole Morten Aamo, and Eric van Oort. A simplified gas-liquid flow model for kick mitigation and control during drilling operations. In *ASME 2015 Dynamic Systems and Control Conference*, pages V002T20A002–V002T20A002. American Society of Mechanical Engineers, 2015. (Cited on page 158.)

[5] Ulf Jakob F. Aarsnes, Florent Di Meglio, Steinar Evje, and Ole Morten Aamo. Control-Oriented Drift-Flux Modeling of Single and Two-Phase Flow for Drilling. page V003T37A003, San Antonio, Texas, USA, 2014. American Society of Mechanical Engineers. (Cited on pages 26, 37, 42, and 80.)

[6] M.H. Abbasi, H. Bansal, H. Zwart, L. Iapichino, W. Schilders, and N. van de Wouw. Power-preserving interconnection of single- and two-

phase flow models for managed pressure drilling. *Accepted for American Control Conference (ACC 2020)*, 2020. (Cited on page 56.)

[7] M.H. Abbasi, S. Naderi Lordejani, N. Velmurugan, C. Berg, L. Iapichino, W.H.A. Schilders, and N. van de Wouw. A Godunov-type Scheme for the Drift Flux Model with Variable Cross Section. *Journal of Petroleum Science and Engineering*, 179, 2019. (Cited on pages 5, 156, 160, and 163.)

[8] Mohammad Hossein Abbasi, Laura Iapichino, Sajad Naderi Lordejani, Wil Schilders, and Nathan van de Wouw. Reduced basis method for managed pressure drilling based on a model with local nonlinearities. *International Journal for Numerical Methods in Engineering*, 2020. (Cited on page 211.)

[9] Remi Abgrall. How to prevent pressure oscillations in multicomponent flow calculations: A quasi conservative approach. *Journal of Computational Physics*, 125(1):150 – 160, 1996. (Cited on page 4.)

[10] Remi Abgrall and Smadar Karni. Computations of Compressible Multifluids. *Journal of Computational Physics*, 169, 2001. (Cited on page 4.)

[11] Remi Abgrall and Richard Saurel. Discrete equations for physical and numerical compressible multiphase mixtures. *Journal of Computational Physics*, 186, 2003. (Cited on page 4.)

[12] Mark Ainsworth and J.Tinsley Oden. A posteriori error estimation in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 142, 1997. (Cited on page 157.)

[13] R. Altmann and P. Schulze. A port-Hamiltonian formulation of the Navier-Stokes equations for reactive flows. *Systems & Control Letters*, pages 51–55, February 2017. (Cited on pages 38, 71, and 110.)

[14] A. Ambroso, C. Chalons, and P.-A. Raviart. A Godunov-type method for the seven-equation model of compressible two-phase flow. *Computers & Fluids*, 54, 2012. (Cited on pages 3, 119, and 156.)

[15] Athanasios C. Antoulas. *7. Balancing and Balanced Approximations*, pages 207–247. 2005. (Cited on page 211.)

[16] A. Moses Badlyan, B. Maschke, C. Beattie, and V. Mehrmann. Open physical systems: from GENERIC to port-Hamiltonian systems. *arXiv:1804.04064 [math]*, April 2018. (Cited on page 71.)

[17] Nicodemus Banagaaya, Peter Benner, and Sara Grundel. Index-Preserving Model Order Reduction for Differential-Algebraic Systems Arising in Gas Transport Networks. In Istvan Farago, Ferenc Izsak, and Peter L. Simon, editors, *Progress in Industrial Mathematics at ECMI 2018*, volume 30, pages 291–297. Springer International Publishing, Cham, 2019. Series Title: Mathematics in Industry. (Cited on page 110.)

[18] J.W. Banks, J.A.F. Hittinger, J.M. Connors, and C.S. Woodward. Numerical error estimation for nonlinear hyperbolic PDEs via nonlinear error transport. *Computer Methods in Applied Mechanics and Engineering*, 213-216, 2012. (Cited on pages 157, 184, 185, 186, and 187.)

[19] H. Bansal, P. Schulze, M.H. Abbasi, H. Zwart, L. Iapichino, W.H.A. Schilders, and N. van de Wouw. Port-Hamiltonian formulation of two-phase flow models. *submitted to Systems & Control Letters*, 2020. (Cited on pages 70, 79, 82, 95, 110, 111, and 113.)

[20] H. Bansal, S. Weiland, L. Iapichino, W. Schilders, and N. van de Wouw. Structure-preserving Spatial Discretization of a Two-Fluid Model. *Accepted for 59th Conference on Decision and Control (CDC 2020)*, 2020. (Cited on pages 110 and 120.)

[21] H. Bansal, H. Zwart, L. Iapichino, W. Schilders, and N. van de Wouw. Port-Hamiltonian modelling of fluid dynamics models with variable cross-section. *Accepted for 24th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2020)*, 2020. (Cited on pages 111 and 112.)

[22] Maxime Barrault, Yvon Maday, Ngoc Cuong Nguyen, and Anthony T. Patera. An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339(9):667–672, 2004. (Cited on pages 9 and 211.)

[23] Luca Bassi, Alessandro Macchelli, and Claudio Melchiorri. An Algorithm to Discretize One-Dimensional Distributed Port Hamiltonian Systems. volume 366. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. (Cited on page 113.)

[24] Kjell H. Bendiksen. An experimental investigation of the motion of long bubbles in inclined tubes. *International Journal of Multiphase Flow*, 10, 1984. (Cited on page 27.)

[25] G Berkooz, PJ Holmes, and John Lumley. *The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows*, volume 25. 2003. (Cited on pages 9, 211, and 216.)

[26] W. J. Beyn and V. Thummler. Freezing Solutions of Equivariant Evolution Equations. *SIAM Journal on Applied Dynamical Systems*, 3(2):85–116, January 2004. (Cited on pages 239, 240, 242, 244, and 247.)

[27] Wolf-Jurgen Beyn, Sabrina Selle, and Vera Thummler. Freezing multipulses and multifronts. *SIAM Journal on Applied Dynamical Systems*, 7(2):577–608, 2008. (Cited on pages 240, 242, 244, and 247.)

[28] Felix Black, Philipp Schulze, and Benjamin Unger. Projection-Based Model Reduction with Dynamically Transformed Modes. *arXiv:1912.11138 [cs, math]*, June 2020. arXiv: 1912.11138. (Cited on pages 238, 239, and 241.)

[29] Julia Brunken, Kathrin Smetana, and Karsten Urban. (Parametrized) First Order Transport Equations: Realization of Optimally Stable Petrov-Galerkin Methods. *arXiv:1803.06925 [math]*, March 2018. arXiv: 1803.06925. (Cited on page 238.)

[30] Annalisa Buffa, Yvon Maday, Anthony T. Patera, Christophe Pruddhomme, and Gabriel Turinici. *A priori* convergence of the Greedy algorithm for the parametrized reduced basis method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46, 2012. (Cited on page 214.)

[31] Nicolas Cagniart, Yvon Maday, and Benjamin Stamm. Model Order Reduction for Problems with Large Convection Effects. In B. N. Chetverushkin, W. Fitzgibbon, Y.A. Kuznetsov, P. Neittaanmäki, J. Periaux, and O. Pironneau, editors, *Contributions to Partial Differential Equations and Applications*, volume 47, pages 131–150. Springer International Publishing, Cham, 2019. Series Title: Computational Methods in Applied Sciences. (Cited on pages 238 and 241.)

[32] Jiaxiang Cai, Yushun Wang, and Chaolong Jiang. Local structure-preserving algorithms for general multi-symplectic Hamiltonian PDEs. *Computer Physics Communications*, 235:210–220, February 2019. (Cited on page 110.)

[33] F. L. Cardoso-Ribeiro. *Port-Hamiltonian modeling and control of a fluid-structure system*. PhD thesis, University of Toulouse, 2016. (Cited on pages 112, 126, and 128.)

[34] F.L. Cardoso-Ribeiro, D. Matignon, and V. Pommier-Buddinger. Modeling of a fluid-structure coupled system using port-Hamiltonian formulation. *IFAC-PapersOnLine*, 48(13):217–222, 2015. (Cited on pages 38, 71, and 110.)

[35] Flavio Luiz Cardoso-Ribeiro, Denis Matignon, and Laurent Lefevre. A Partitioned Finite Element Method for power-preserving discretization of open systems of conservation laws. *arXiv:1906.05965 [cs, math]*, June 2019. arXiv: 1906.05965. (Cited on page 112.)

[36] Kevin Carlberg. Adaptive h-refinement for reduced-order models. *International Journal for Numerical Methods in Engineering*, 102(5):1192–1210, 2015. (Cited on page 238.)

[37] C. E. Castro and E. F. Toro. Godunov schemes for compressible multiphase flows. 2006. (Cited on page 4.)

[38] Ismail Celik and Gusheng Hu. Single Grid Error Estimation Using Error Transport Equation. *Journal of Fluids Engineering*, 126, 2004. (Cited on pages 157 and 184.)

[39] Elena Celledoni, Sølve Eidnes, Brynjulf Owren, and Torbjørn Ringholm. Energy-preserving methods on Riemannian manifolds. *Mathematics of Computation*, 89(322):699–716, September 2019. (Cited on page 110.)

[40] Elena Celledoni, Marta Farre Puiggali, Eirik Hoel Høiseth, and David Martin de Diego. Energy-Preserving Integrators Applied to Nonholonomic Systems. *Journal of Nonlinear Science*, 29(4):1523–1562, August 2019. (Cited on page 110.)

[41] Y.A. Cengel and M. A. Boles. *Thermodynamics: An Engineering Approach*. New York City, NY, USA, 2006. (Cited on page 46.)

[42] J. Cervera, A.J. van der Schaft, and A. Banos. Interconnection of port-Hamiltonian systems and composition of dirac structures. *Automatica*, 43(2):212–225, 2007. (Cited on page 71.)

[43] S. Chaturantabut, C. Beattie, and S. Gugercin. Structure-Preserving Model Reduction for Nonlinear Port-Hamiltonian Systems. *SIAM J. Sci. Comput*, 38(5):B837–B865, January 2016. (Cited on pages 38, 71, and 110.)

[44] Saifon Chaturantabut and Danny C. Sorensen. Nonlinear Model Reduction via Discrete Empirical Interpolation. *SIAM Journal on Scientific Computing*, 32(5):2737–2764, 2010. (Cited on pages 9 and 211.)

[45] Francisco Chinesta, Roland Keunings, and Adrien Leygue. *The Proper Generalized Decomposition for Advanced Numerical Simulations*. Springer-Briefs in Applied Sciences and Technology. Springer International Publishing, Cham, 2014. (Cited on page 211.)

[46] Jinho Choi, Eduardo Pereyra, Cem Sarica, Changhyup Park, and Joe M. Kang. An Efficient Drift-Flux Closure Relationship to Estimate Liquid Holdups of Gas-Liquid Two-Phase Flow in Pipes. *Energies*, 5, 2012. (Cited on page 27.)

[47] Frederic Coquel, Jean-Marc Herard, and Khaled Saleh. A splitting method for the isentropic Baer-Nunziato two-phase flow model. *ESAIM: Proceedings*, 38, 2012. (Cited on page 4.)

[48] H.W.P. de Wilde. *Port-Hamiltonian discretization of gas pipeline networks*. Masters thesis, University of Groningen, 2015. (Cited on pages 38, 70, 74, 110, and 111.)

[49] Mamadou Diagne and Bernhard Maschke. Boundary Port Hamiltonian systems of conservation laws coupled by a moving interface. *4th IFAC Workshop on Lagrangian and Hamiltonian Methods for Non Linear Control*, 45(19):265–270, January 2012. (Cited on page 295.)

[50] Mamadou Lamine Diagne. *Modelling and control of systems of conservation laws with a moving interface: an application to an extrusion process*. PhD thesis, Universite Claude Bernard - Lyon I, 2013. (Cited on page 294.)

[51] Donald A. Drew and Richard T. Lahey. Application of general constitutive principles to the derivation of multidimensional two-phase flow equations. *International Journal of Multiphase Flow*, 5, 1979. (Cited on pages 7 and 164.)

[52] Martin Drohmann, Bernard Haasdonk, and Mario Ohlberger. Reduced Basis Approximation for Nonlinear Parametrized Evolution Equations based on Empirical Operator Interpolation. *SIAM Journal on Scientific Computing*, 34, 2012. (Cited on pages 211, 212, 221, and 276.)

[53] Vincent Duindam, Alessandro Macchelli, Stefano Stramigioli, and Herman Bruyninckx, editors. *Modeling and control of complex physical systems: the port-Hamiltonian approach*. Springer, Berlin, 2009. (Cited on pages 38, 46, 55, 56, 58, 61, 71, 82, 88, 111, and 112.)

[54] H. Egger. Structure preserving approximation of dissipative evolution problems. *Numerische Mathematik*, 143(1):85–106, 2019. (Cited on page 110.)

[55] Virginie Ehrlacher, Damiano Lombardi, Olga Mula, and François-Xavier Vialard. Nonlinear model reduction on metric spaces. Application to one-dimensional conservative PDEs in Wasserstein spaces. *ESAIM: Mathematical Modelling and Numerical Analysis*, March 2020. (Cited on page 238.)

[56] S. Evje and T. Flatten. On the wave structure of two-phase flow models. *SIAM J. Appl. Math*, 67(2):487–511, 2007. (Cited on pages 5, 26, 27, 29, 30, 38, 43, 44, 65, 94, and 163.)

[57] Steinar Evje and Kjell K. Fjelde. Hybrid Flux-Splitting Schemes for a Two-Phase Flow Model. *Journal of Computational Physics*, 175(2):674–701, January 2002. (Cited on pages 3, 5, 6, 7, 26, 43, 47, 119, 155, 156, 158, 160, 161, 164, 166, 168, 169, 171, 172, 173, 174, 176, 181, 182, 195, 201, 204, 226, 227, and 233.)

[58] Steinar Evje and Tore Flatten. Hybrid flux-splitting schemes for a common two-fluid model. *Journal of Computational Physics*, 192, 2003. (Cited on pages 3, 5, 6, 7, 119, 156, 157, 168, and 174.)

[59] Steinar Evje and Tore Flatten. Hybrid central-upwind schemes for numerical resolution of two-phase flows. *ESAIM: Mathematical Modelling and Numerical Analysis*, 39(2), 2005. (Cited on pages 3, 5, 7, 119, 156, 168, 171, and 172.)

[60] Steinar Evje and Tore Flatten. Weakly Implicit Numerical Schemes for a Two-Fluid Model. *SIAM Journal on Scientific Computing*, 26(5), 2005. (Cited on pages 3, 4, 5, 7, 119, 156, and 168.)

[61] Steinar Evje and Tore Flatten. CFL-Violating Numerical Schemes for a Two-Fluid Model. *Journal of Scientific Computing*, 29(1), 2006. (Cited on pages 3, 5, 119, and 156.)

[62] O. Farle, D. Klis, M. Jochum, O. Floch, and R. Dyczij-Edlinger. A port-hamiltonian finite-element formulation for the maxwell equations. In *2013 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pages 324–327, 2013. (Cited on page 71.)

[63] Kang Feng and Mengzhao Qin. Symplectic Difference Schemes for Hamiltonian Systems. In *Symplectic Geometric Algorithms for Hamiltonian Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. (Cited on page 110.)

[64] Joel H. Ferziger and Milovan Peric. *Computational Methods for Fluid Dynamics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002. (Cited on page 168.)

[65] A.B. Figueiredo, R.M. Baptista, F.B. de Freitas Rachid, and G.C..R. Bodstein. Numerical simulation of stratified-pattern two-phase flow in gas pipelines using a two-fluid model. *International Journal of Multiphase Flow*, 88:30–49, 2017. (Cited on page 98.)

[66] Tore Flatten and Alexandre Morin. On interface transfer terms in two-fluid models. *International Journal of Multiphase Flow*, 45(Supplement C):24 – 29, 2012. (Cited on page 26.)

[67] Tore Flatten and Svend Tollak Munkejord. The approximate Riemann solver of Roe applied to a drift-flux two-phase flow model. *ESAIM: Mathematical Modelling and Numerical Analysis*, 40(4), 2006. (Cited on pages 3, 5, 7, 119, 156, 168, and 172.)

[68] Flatten, Tore and Munkejord, Svend Tollak. The approximate riemann solver of roe applied to a drift-flux two-phase flow model. *ESAIM: M2AN*, 40(4):735–764, 2006. (Cited on page 4.)

[69] F. Franca and R.T. Lahey. The use of drift-flux techniques for the analysis of horizontal two-phase flows. *International Journal of Multiphase Flow*, 18, 1992. (Cited on page 27.)

[70] William D. Fullmer, Martin A. Lopez de Bertodano, and Xiaoying Zhang. Verification of a Higher-Order Finite Difference Scheme for the One-Dimensional Two-Fluid Model. *The Journal of Computational Multiphase Flows*, 5, 2013. (Cited on pages 119 and 156.)

[71] S.L. Gavrilyuk and J. Fabre. Lagrangian coordinates for a drift-flux model of a gas-liquid mixture. *International Journal of Multiphase Flow*, 22(3), June 1996. (Cited on page 161.)

[72] R.J Gelinas, S.K Doss, and Keith Miller. The moving finite element method: Applications to general partial differential equations with multiple large gradients. *Journal of Computational Physics*, 40(1):202–249, March 1981. (Cited on page 169.)

[73] Jean-Frederic Gerbeau and Damiano Lombardi. Approximated Lax pairs for the reduced order integration of nonlinear evolution equations. *Journal of Computational Physics*, 265:246–269, May 2014. (Cited on pages 238 and 239.)

[74] Jean-Frederic Gerbeau, Damiano Lombardi, and Elisa Schenone. Approximated lax pairs and empirical interpolation for nonlinear parabolic partial differential equations. 2016. (Cited on page 239.)

[75] Goran Golo, Viswanath Talasila, Arjan van der Schaft, and Bernhard Maschke. Hamiltonian discretization of boundary control systems. *Automatica*, 40(5):757 – 771, 2004. (Cited on pages 110, 111, 112, 113, 119, 120, 123, 126, and 127.)

[76] Jan Einar Gravdal, Hardy Siahaan, and Knut S. Bjorkevoll. Back-Pressure MPD in Extended-Reach Wells - Limiting Factors for the Ability to Achieve Accurate Pressure Control. Society of Petroleum Engineers, 2014. (Cited on pages ix, 24, 30, and 31.)

[77] Constantin Greif and Karsten Urban. Decay of the Kolmogorov N-width for wave problems. *Applied Mathematics Letters*, 96, 2019. (Cited on pages 9 and 215.)

[78] Martin A. Grepl, Yvon Maday, Ngoc C. Nguyen, and Anthony T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41, 2007. (Cited on page 211.)

[79] Christian Grossmann, Hans-Gorg Roos, and Martin Stynes. *Numerical Treatment of Partial Differential Equations*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. (Cited on page 169.)

[80] Ernst Hairer and Gerhard Wanner. *Solving Ordinary Differential Equations II*, volume 14 of *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1996. (Cited on pages 110, 141, and 144.)

[81] Jan S Hesthaven, Gianluigi Rozza, and Benjamin Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. 2016. (Cited on page 9.)

[82] J.S. Hesthaven and C. Pagliantini. Structure-preserving reduced basis methods for hamiltonian systems with a nonlinear poisson structure. *under review*, 2018. (Cited on page 110.)

[83] J.S. Hesthaven and S. Ubbiali. Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363:55–78, June 2018. (Cited on page 238.)

[84] D. D. Holm and B. A. Kupershmidt. Hydrodynamics and electrohydrodynamics of adiabatic multiphase fluids and plasmas. *Int. J. Multiphas. Flow*, 12(4):667–680, 1986. (Cited on page 38.)

[85] Willem Hundsdorfer and Jan Verwer. *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, volume 33 of *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003. (Cited on page 5.)

[86] Angelo Iollo and Damiano Lombardi. Advection modes by optimal mass transfer. *Phys. Rev. E*, 89:022923, Feb 2014. (Cited on page 238.)

[87] M. Ishii and T. Hibiki. *Thermo-Fluid Dynamics of Two-Phase Flow*. Springer, New York, NY, USA, 2006. (Cited on page 65.)

[88] Mamoru Ishii and Takashi Hibiki. One-Dimensional Drift-Flux Model. In *Thermo-Fluid Dynamics of Two-Phase Flow*. Springer New York, New York, NY, 2011. (Cited on page 27.)

[89] Birgit Jacob and Hans Zwart. *Linear port-Hamiltonian systems on infinite-dimensional spaces*. Springer, New York, 2012. (Cited on pages 38 and 45.)

[90] Kjell Kå. *Computers & Fluids*. (Cited on pages 5, 156, 158, 169, 171, 179, 181, 182, 227, and 232.)

[91] Glenn-Ole Kaasa, Øyvind N. Stamnes, Ole Morten Aamo, and Lars S. Imsland. Simplified Hydraulics Model Used for Intelligent Estimation of Downhole Pressure for a Managed-Pressure-Drilling Control System. *SPE Drilling & Completion*, 27(01):127–138, March 2012. (Cited on page 3.)

[92] Smadar Karni, Eduard Kirr, Alexander Kurganov, and Guergana Petrova. Compressible two-phase flows by central and upwind schemes. *ESAIM: Mathematical Modelling and Numerical Analysis*, 38(3), 2004. (Cited on page 4.)

[93] Kouhei Kawanishi, Yasuhiko Hirao, and Ayao Tsuge. An experimental study on drift flux parameters for two-phase flow in vertical round tubes. *Nuclear Engineering and Design*, 120, 1990. (Cited on page 27.)

[94] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, NJ, USA, third edition, 2002. (Cited on page 8.)

[95] Paul Kotyczka. Discretized models for networks of distributed parameter port-Hamiltonian systems. In *nDS '13; Proceedings of the 8th International Workshop on Multidimensional Systems*, September 2013. (Cited on pages 71, 111, and 113.)

[96] Paul Kotyczka and Laurent Lefevre. Discrete-time port-Hamiltonian systems: A definition based on symplectic integration. *Systems & Control Letters*, 133:104530, 2019. (Cited on pages 110, 114, 141, 142, 143, 146, and 149.)

[97] Paul Kotyczka, Bernhard Maschke, and Laurent Lefevre. Weak form of Stokes-Dirac structures and geometric discretization of port-Hamiltonian systems. *Journal of Computational Physics*, 361, May 2018. (Cited on pages 71, 112, 119, 120, 126, 129, and 153.)

[98] H.-O. KREISS and G. SCHERER. Finite Element and Finite Difference Methods for Hyperbolic Partial Differential Equations. In Carl de Boor, editor, *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 195–212. Academic Press, January 1974. (Cited on page 156.)

[99] M. Kurula, A. van der Schaft, and H.J. Zwart. Composition of infinite-dimensional dirac structures. *17th International Symposium on Mathematical Theory of Networks and Systems, MTNS*, pages 27–32, 2006. (Cited on pages 56 and 71.)

[100] Jens Lang. Adaptive FEM for reaction-diffusion equations. *Applied Numerical Mathematics*, 26(1-2):105–116, January 1998. (Cited on page 169.)

[101] Y. Le Gorrec, H. Zwart, and B Maschke. Dirac structures and boundary control systems associated with skew-symmetric differential operators. *SIAM J. Control. Optim*, 44(5):1864–1892, 2005. (Cited on pages 38, 55, 56, 57, 58, 82, 88, 111, 116, and 296.)

[102] Kookjin Lee and Kevin Carlberg. Deep Conservation: A latent dynamics model for exact satisfaction of physical conservation laws. *arXiv:1909.09754 [physics]*, September 2019. arXiv: 1909.09754. (Cited on page 238.)

[103] Barbara Lee Keyfitz, Richard Sanders, Michael Sever, ,Department of Mathematics, University of Houston, Houston, Texas 77204-3008, and ,Department of Mathematics, The Hebrew University, Jerusalem. Lack of hyperbolicity in the two-fluid model for two-phase incompressible flow. *Discrete & Continuous Dynamical Systems - B*, 3, 2003. (Cited on page 30.)

[104] Christoph Lehrenfeld. The Nitsche XFEM-DG Space-Time Method and its Implementation in Three Space Dimensions. *SIAM Journal on Scientific Computing*, 37(1):A245–A270, January 2015. (Cited on page 169.)

[105] Randall J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002. (Cited on pages 25, 160, 171, 248, and 249.)

[106] Zhiyuan Li, Naira Hovakimyan, and Glenn-Ole Kaasa. Bottomhole Pressure Estimation and $\mathscr{L}_1$ Adaptive Control in Managed Pressure Drilling System. *IFAC Proceedings Volumes*, 45, 2012. (Cited on pages ix and 22.)

[107] Shan Liang, Wei Liu, and Li Yuan. Solving seven-equation model for compressible two-phase flow using multiple gpus. *Computers Fluids*, 99(Supplement C):156 – 171, 2014. (Cited on page 26.)

[108] Meng-Sing Liou. A Sequel to AUSM: AUSM+. *Journal of Computational Physics*, 129(2):364–382, December 1996. (Cited on pages 172 and 174.)

[109] Meng-Sing Liou. A sequel to AUSM, Part II: AUSM+-up for all speeds. *Journal of Computational Physics*, 214(1):137–170, May 2006. (Cited on page 172.)

[110] Meng-Sing Liou, Loc Nguyen, Chih-Hao Chang, Sveta Sushchikh, Robert Nourgaliev, and Theo Theofanous. Hyperbolicity, Discontinuities, and Numerics of Two-Fluid Models. In Herman Deconinck and E. Dick, editors, *Computational Fluid Dynamics 2006*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. (Cited on page 30.)

[111] Meng-Sing Liou and Christopher J. Steffen. A New Flux Splitting Scheme. *Journal of Computational Physics*, 107(1):23–39, July 1993. (Cited on page 172.)

[112] Sajad Naderi Lordejani, Bart Besselink, Mohammad Hossein Abbasi, Glenn-Ole Kaasa, Wil H. A. Schilders, and Nathan van de Wouw. Control-Oriented Modeling for Managed Pressure Drilling Automation Using Model Order Reduction. *IEEE Transactions on Control Systems Technology*, 2020. (Cited on page 211.)

[113] A. Macchelli, Y. Le Gorrec, and H. Ramirez. Boundary energy-shaping control of an ideal compressible isentropic fluid in 1-d. *IFAC-PapersOnLine*, 50(1):5598–5603, 2017. (Cited on page 38.)

[114] A. Macchelli and C. Melchiorri. Modeling and control of the timoshenko beam. the distributed port Hamiltonian approach. *SIAM J. Control. Optim*, 43(2):743–767, 2004. (Cited on page 38.)

[115] A. Macchelli and C. Melchiorri. Modeling and control of the timoshenko beam. the distributed port hamiltonian approach. *SIAM J. Control. Optim*, 43(2):743–767, 2004. (Cited on pages 38 and 110.)

[116] A. Macchelli, A.J. van der Schaft, and C. Melchiorri. Port Hamiltonian formulation of infinite dimensional systems I. Modeling. In *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, pages 3762–3767, Vol.4, Nassau, Bahamas, 2004. (Cited on pages 52 and 71.)

[117] V.D.S. Martins, B. Maschke, and Y. Le Gorrec. Hamiltonian approach to the stabilization of systems of two conservation laws. *IFAC Proceedings Volumes*, 43(14):581–586, 2010. (Cited on pages 38 and 71.)

[118] B. Maschke and A. van der Schaft. Port-Controlled Hamiltonian Systems: modelling origins and system-theoretic properties. *IFAC Proceedings Volumes*, 25:359 – 365, 1992. (Cited on pages 38 and 71.)

[119] B.M. Maschke and A. J. van der Schaft. From conservation laws to port-hamiltonian representations of distributed-parameter systems. *IFAC Proceedings Volumes*, 38(1):483–488, 2005. (Cited on page 71.)

[120] B.M. Maschke and Arjan van der Schaft. *Interconnected mechanical systems, part I: geometry of interconnection and implicit Hamiltonian systems*, pages 1–16. Imperial College Press, United Kingdom, 1997. (Cited on page 71.)

[121] B.M. Maschke and Arjan van der Schaft. *Interconnected mechanical systems, part II: the dynamics of spatial mechanical networks*, pages 17–30. Imperial College Press, United Kingdom, 1997. (Cited on page 71.)

[122] Denis Matignon and Thomas Helie. A class of damping models preserving eigenspaces for linear conservative port-Hamiltonian systems. *European Journal of Control*, 19, 2013. (Cited on page 77.)

[123] Volker Mehrmann and Riccardo Morandin. Structure-preserving discretization for port-Hamiltonian descriptor systems. *arXiv:1903.10451 [math]*, 2019. (Cited on pages 41, 75, 90, 110, 154, and 298.)

[124] Morin, Alexandre and Flatten, Tore. A two-fluid four-equation model with instantaneous thermodynamical equilibrium. *ESAIM: M2AN*, 50(4):1167–1192, 2016. (Cited on page 26.)

[125] P. J. Morrison. Structure and structure-preserving algorithms for plasma physics. *Physics of Plasmas*, 24(5):055502, 2017. (Cited on page 110.)

[126] R. Moulla, L. Lefevre, and B. Maschke. Pseudo-spectral methods for the spatial symplectic reduction of open systems of conservation laws. *Journal of Computational Physics*, 231(4):1272 – 1292, 2012. (Cited on pages 111, 112, and 126.)

[127] Svend Tollak Munkejord. *Analysis of the two-fluid model and the drift-flux model for numerical calculation of two-phase flow*. PhD thesis, NTNU, 2006. (Cited on page 30.)

[128] Svend Tollak Munkejord, Steinar Evje, and Tore Flatten. The multi-stage centred-scheme approach applied to a drift-flux two-phase flow model. *International Journal for Numerical Methods in Fluids*, 52(6), 2006. (Cited on pages 3, 5, 7, 119, 156, and 168.)

[129] Angelo Murrone and Herve Guillard. A five equation reduced model for compressible two phase flow problems. *Journal of Computational Physics*, 202(2):664 – 698, 2005. (Cited on page 26.)

[130] S. Naderi Lordejani, M. H. Abbasi, N. Velmurgan, C. Berg, J. Å. Stakvik, B. Besselink, L. Iapichino, F. Di Meglio, W.H.A. Schilders, and N. van de Wouw. Modeling and numerical implementation of managed pressure drilling systems for evaluating pressure control systems. *SPE Drilling & Completion*, 2020. (Cited on pages 70, 98, 111, 155, and 156.)

[131] Nirmal J. Nair and Maciej Balajewicz. Transported snapshot model order reduction approach for parametric, steady-state fluid flows containing parameter-dependent shocks. *International Journal for Numerical Methods in Engineering*, 117(12):1234–1262, March 2019. (Cited on page 238.)

[132] Yang-Yao Niu, Yu-Chin Lin, and Chih-Hao Chang. A further work on multi-phase two-fluid approach for compressible multi-phase flows. *International Journal for Numerical Methods in Fluids*, 58(8), 2008. (Cited on page 4.)

[133] Gerhard Nygaard and Geir Nævdal. Nonlinear model predictive control scheme for stabilizing annulus pressure during oil well drilling. *Journal of Process Control*, 16(7):719–732, August 2006. (Cited on page 3.)

[134] Mario Ohlberger and Stephan Rave. Nonlinear reduced basis approximation of parameterized evolution equations via the method of freezing. *Comptes Rendus Mathematique*, 351(23-24):901–906, 2013. (Cited on pages 238, 239, and 240.)

[135] Mario Ohlberger and Stephan Rave. Reduced basis methods: Success, limitations and future challenges. *Proceedings of the Conference Algoritmy*, pages 1–12, 2016. (Cited on pages 9, 214, 215, and 274.)

[136] R. Ortega, A. van der Schaft, B. Maschke, and G. Escobar. Interconnection and damping assignment passivity-based control of port-controlled hamilonian systems. *Automatica*, 38:585–596, 2002. (Cited on page 38.)

[137] Petter Osmundsen, Kristin Helen Roll, and Ragnar Tveterays. Exploration drilling productivity at the Norwegian shelf. *Journal of Petroleum Science and Engineering*, 73, 2010. (Cited on page 1.)

[138] N. Panicker, A. Passalacqua, and R.O. Fox. On the hyperbolicity of the two-fluid model for gas-liquid bubbly flows. *Applied Mathematical Modelling*, 57, 2018. (Cited on page 30.)

[139] R. Pasumarthy and A. van der Schaft. Achievable Casimirs and its implications on control of port-Hamiltonian systems. *International Journal of Control*, 80, September 2007. (Cited on page 71.)

[140] Ramkrishna Pasumarthy and Arjan van der Schaft. A Finite Dimensional Approximation of the shallow water Equations: The port-Hamiltonian Approach. pages 3984–3989, San Diego, CA, USA, 2006. Proceedings of the 45th IEEE Conference on Decision and Control. (Cited on pages 38, 56, and 113.)

[141] Benjamin Peherstorfer. Model reduction for transport-dominated problems via online adaptive bases and adaptive sampling. *arXiv:1812.02094 [cs, math]*, December 2018. arXiv: 1812.02094. (Cited on page 238.)

[142] Benjamin Peherstorfer, Daniel Butnaru, Karen Willcox, and Hans-Joachim Bungartz. Localized discrete empirical interpolation method. *SIAM Journal on Scientific Computing*, 36, 2014. (Cited on page 211.)

[143] Tyrone S. Phillips, Joseph M. Derlaga, Christopher J. Roy, and Jeff Borggaard. Error transport equation boundary conditions for the Euler and Navier-Stokes equations. *Journal of Computational Physics*, 330, 2017. (Cited on pages 157 and 184.)

[144] Allan Pinkus. *n-Widths in Approximation Theory*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1985. (Cited on page 214.)

[145] A. Poullikkas. Effects of two-phase liquid-gas flow on the performance of the nuclear reactor cooling pumps. *Progress in Nuclear Energy*, 42(1):3–10, 2003. (Cited on page 37.)

[146] Alfio Quarteroni, Andrea Manzoni, and Federico Negri. *Reduced Basis Methods for Partial Differential Equations*, volume 92 of *UNITEXT*. Springer International Publishing, Cham, 2016. (Cited on page 213.)

[147] Alfio Quarteroni and Gianluigi Rozza, editors. *Reduced order methods for modeling and computational reduction: selected peer-reviewed contributions submitted by the plenary speakers to the Workshop "Reduced Basis, POD and Reduced Order Methods for Model and Computational Reduction:*

Towards Real-Time Computing and Visualization" funded by CECAM (European Center for Atomistic and Molecular Computing), hosted at Ecole Polytechnique Federale de Lausanne, Switzerland on 14-16 May 2012. Number 9 in MS&A - Modeling, simulation and applications. Springer, Cham, 2014. (Cited on page 210.)

[148] V R Ambati R Pasumarthy and A. van der Schaft. Port-hamiltonian formulation of shallow water equations with coriolis force and topography. *Proceedings of the 18th International symposium on Mathematical Theory of Networks and Systems, Virginia Tech*, 2008. (Cited on page 113.)

[149] V.H Ransom and D.L Hicks. Hyperbolic two-pressure models for two-phase flow. *Journal of Computational Physics*, 53, 1984. (Cited on pages 7 and 164.)

[150] J. Reiss, P. Schulze, J. Sesterhenn, and V. Mehrmann. The Shifted Proper Orthogonal Decomposition: A Mode Decomposition for Multiple Transport Phenomena. *SIAM Journal on Scientific Computing*, 40(3):A1322–A1344, January 2018. (Cited on pages 238, 239, 241, and 247.)

[151] Donsub Rim, Scott Moe, and Randall J. LeVeque. Transport reversal for model reduction of hyperbolic partial differential equations. *SIAM/ASA Journal on Uncertainty Quantification*, 6(1):118–150, 2018. (Cited on pages 238 and 241.)

[152] Jens Rottmann-Matthes. An IMEX-RK scheme for capturing similarity solutions in the multidimensional Burgers' equation. *arXiv:1612.04127 [math]*, 2016. arXiv: 1612.04127. (Cited on pages 239 and 264.)

[153] C. W. ROWLEY. Model reduction for fluids, using balanced proper orthogonal decomposition. *International Journal of Bifurcation and Chaos*, 15(03):997–1013, 2005. (Cited on page 216.)

[154] Clarence W. Rowley, Ioannis G. Kevrekidis, Jerrold E. Marsden, and Kurt Lust. Reduction and reconstruction for self-similar dynamical systems. *Nonlinearity*, 16(4):1257, 2003. (Cited on pages 238, 242, and 264.)

[155] Clarence W. Rowley and Jerrold E. Marsden. Reconstruction equations and the Karhunen Love expansion for systems with symmetry. *Physica D: Nonlinear Phenomena*, 142(1):1 – 19, 2000. (Cited on pages 238 and 242.)

[156] V.V. Rusanov. Calculation of Interaction of Non-steady Shock Waves with Obstacles. 1961. (Cited on page 65.)

[157] Richard Saurel and Remi Abgrall. A Multiphase Godunov Method for Compressible Multifluid and Multiphase Flows. *Journal of Computational Physics*, 150, 1999. (Cited on pages 7 and 164.)

[158] Martin Schemann and Folkmar A. Bornemann. An adaptive Rothe method for the wave equation. *Computing and Visualization in Science*, 1(3):137–144, November 1998. (Cited on page 169.)

[159] W. H. A. Schilders, H. A. van der Vorst, and Joost Rommes, editors. *Model order reduction: theory, research aspects and applications*. Number 13 in Mathematics in industry ; The European Consortium for Mathematics in Industry. Springer, 2008. (Cited on pages 210 and 211.)

[160] Anass Serhani, Ghislain Haine, and Denis Matignon. Anisotropic heterogeneous n-D heat equation with boundary control and observation: I. Modeling as port-Hamiltonian. *3rd IFAC Workshop on Thermodynamic Foundations for a Mathematical Systems Theory TFMST 2019*, 52, 2019. (Cited on page 112.)

[161] Anass Serhani, Ghislain Haine, and Denis Matignon. Anisotropic heterogeneous n-D heat equation with boundary control and observation: II. Structure-preserving discretization. *3rd IFAC Workshop on Thermodynamic Foundations for a Mathematical Systems Theory TFMST 2019*, 52, 2019. (Cited on page 112.)

[162] Anass Serhani, Denis Matignon, and Ghislain Haine. A Partitioned Finite Element Method for the Structure-Preserving Discretization of Damped Infinite-Dimensional Port-Hamiltonian Systems with Boundary Control. In Frank Nielsen and Frederic Barbaresco, editors, *Geometric Science of Information*, volume 11712, pages 549–558. Springer International Publishing, Cham, 2019. (Cited on pages 112 and 113.)

[163] Marko Seslija, Jacquelien M. A. Scherpen, and Arjan van der Schaft. Explicit simplicial discretization of distributed-parameter port-Hamiltonian systems. *Automatica*, pages 369 – 377, 2014. (Cited on pages 110 and 112.)

[164] Robert D. Skeel. Thirteen ways to estimate global error. *Numerische Mathematik*, 48:1–20, 1986. (Cited on page 157.)

[165] Joel Smoller. *Shock Waves and Reaction-Diffusion Equations*, volume 258 of *Grundlehren der mathematischen Wissenschaften*. Springer New York, New York, NY, 1994. (Cited on page 161.)

[166] Jin Ho Song and M. Ishii. The well-posedness of incompressible one-dimensional two-fluid model. *International Journal of Heat and Mass Transfer*, 43, 2000. (Cited on pages 7 and 164.)

[167] T. Flatten S.T. Munkejord, S. Evje. A musta scheme for a nonconservative two-fluid model. *SIAM Journal of Scientific Computing*, 31(4):2587–2622, 2009. (Cited on pages 3, 5, 7, 119, 156, and 168.)

[168] Joseph L Steger and R.F Warming. Flux vector splitting of the inviscid gasdynamic equations with application to finite-difference methods. *Journal of Computational Physics*, 40(2):263–293, April 1981. (Cited on page 172.)

[169] Hans J. Stetter. The defect correction principle and discretization methods. *Numerische Mathematik*, 29, 1978. (Cited on page 157.)

[170] J.H. Stuhmiller. The influence of interfacial pressure forces on the character of two-phase flow model equations. *International Journal of Multiphase Flow*, 3, 1977. (Cited on pages 7 and 164.)

[171] P. K. Sweby. High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws. *SIAM Journal on Numerical Analysis*, 21(5):995–1011, October 1984. (Cited on page 177.)

[172] T. Taddei, S. Perotto, and A. Quarteroni. Reduced basis techniques for nonlinear conservation laws. *ESAIM: M2AN*, 49(3):787–814, 2015. (Cited on pages 238 and 288.)

[173] Emanuele Tassi. Hamiltonian closures in fluid models for plasmas. *Eur. Phys. J. D*, 71(269), 2017. (Cited on page 51.)

[174] Mai Duc Thanh. A well-balanced numerical scheme for a model of two-phase flows with treatment of nonconservative terms. *Advances in Computational Mathematics*, 45, 2019. (Cited on page 5.)

[175] J. W. Thomas. *Numerical Partial Differential Equations*, volume 33 of *Texts in Applied Mathematics*. Springer New York, New York, NY, 1999. (Cited on page 177.)

[176] I. Tiselj and S. Petelin. Second order numerical method for two-fluid model of air-water flow. *Nuclear Society of Slovenia*, 1995. (Cited on page 94.)

[177] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer, Dordrecht ; New York, 3rd ed edition, 2009. (Cited on pages 160, 161, 167, 168, 172, 176, 177, 275, and 294.)

[178] Ngoc Minh Trang VU, Laurent Lefevre, and Bernhard Maschke. Port-Hamiltonian formulation for systems of conservation laws: application to plasma dynamics in Tokamak reactors. *IFAC Proceedings*, 45:108–113, 2012. (Cited on pages 38 and 71.)

[179] Vincent Trenchant, Hector Ramirez, Yann Le Gorrec, and Paul Kotyczka. Finite differences on staggered grids preserving the port-Hamiltonian structure with application to an acoustic duct. *Journal of Computational Physics*, 373:673–697, 2018. (Cited on page 112.)

[180] Claire Valentin, Miguel Magos, and Bernhard Maschke. A port-Hamiltonian formulation of physical switching systems with varying constraints. *Automatica*, 43(7):1125–1133, July 2007. (Cited on page 294.)

[181] E. H. van Brummelen and B. Koren. A pressure-invariant conservative godunov-type method for barotropic two-fluid flows. *J. Comput. Phys.*, 185(1):289–308, February 2003. (Cited on pages 4, 172, 189, and 192.)

[182] A. van der Schaft and D. Jeltsema. Port-Hamiltonian systems theory: An introductory overview. *Foundations and Trends in Systems and Control*, 1(2):173–378, 2014. (Cited on page 38.)

[183] A.J. van der Schaft and M.K. Camlibel. A state transfer principle for switching port-Hamiltonian systems. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 45–50, Shanghai, China, December 2009. IEEE. (Cited on page 294.)

[184] A.J. van der Schaft and B.M. Maschke. Hamiltonian formulation of distributed-parameter systems with boundary energy flow. *J. Geom. Phys*, 42:166–194, 2002. (Cited on pages 38, 58, and 71.)

[185] Arjan van der Schaft. Port-Hamiltonian Modeling for Control. *Annual Review of Control, Robotics, and Autonomous Systems (2020)*. (Cited on pages 8 and 71.)

[186] Arjan van der Schaft. *Interconnection and geometry*, pages 203–218. Eigen beheer, 1999. (Cited on page 71.)

[187] Arjan van der Schaft. Port-Hamiltonian Systems. In *L2-Gain and Passivity Techniques in Nonlinear Control*. Springer International Publishing, Cham, 2017. (Cited on page 121.)

[188] Arjan van der Schaft and Bernhard Maschke. Generalized Port-Hamiltonian DAE Systems. *arXiv:1808.01845 [math]*, 2018. (Cited on page 71.)

[189] Arjan van der Schaft and Bernhard Maschke. Dirac and Lagrange algebraic constraints in nonlinear port-Hamiltonian systems. *arXiv:1909.07025 [math]*, 2019. (Cited on page 71.)

[190] J.J.W. van der Vegt and J.J. Sudirham. A space-time discontinuous Galerkin method for the time-dependent Oseen equations. *Special Issue in Honor of Piet Hemker*, 58(12), December 2008. (Cited on page 169.)

[191] Bram van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *Journal of Computational Physics*, 32(1):101–136, July 1979. (Cited on page 177.)

[192] J.S.B. van Zwieten, B. Sanderse, M.H.W. Hendrix, C. Vuik, and R.A.W.M. Henkes. Efficient simulation of one-dimensional two-phase flow with a high-order h-adaptive space-time Discontinuous Galerkin method. *Ninth International Conference on Computational Fluid Dynamics (ICCFD9)*, 156, 2017. (Cited on pages 119 and 156.)

[193] E. I. Verriest. Time Variant Balancing and Nonlinear Balanced Realizations. In Hans-Georg Bock, Frank de Hoog, Avner Friedman, Arvind Gupta, Helmut Neunzert, William R. Pulleyblank, Torgeir Rusten, Fadil Santosa, Anna-Karin Tornberg, Luis L. Bonilla, Robert Mattheij, Otmar Scherzer, Wilhelmus H. A. Schilders, Henk A. van der Vorst, and Joost Rommes, editors, *Model Order Reduction: Theory, Research Aspects and Applications*, volume 13. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. (Cited on page 211.)

[194] J.A. Villegas. *A Port-Hamiltonian approach to distributed parameter systems*. PhD thesis, 2007. (Cited on pages 56, 57, 58, and 296.)

[195] Stefan Volkwein. Model reduction using proper orthogonal decomposition. *Lecture Notes, Institute of Mathematics and Scientific Computing, University of Graz.*, 2011. (Cited on pages 9, 211, and 216.)

[196] Stefan Volkwein. Proper orthogonal decomposition:theory and reduced-order modelling. *Lecture Notes, Department of Mathematics and Statistics, University of Konstanz.*, 2013. (Cited on pages 216, 217, and 218.)

[197] Jeroen Wackers and Barry Koren. A fully conservative model for compressible two-fluid flow. *International Journal for Numerical Methods in Fluids*, 47(10-11):1337–1343, 2005. (Cited on page 4.)

[198] Yasuhiro Wada and Meng Liou. A flux splitting scheme with high-resolution and robustness for discontinuities. *32nd Aerospace Sciences Meeting and Exhibit*, 07 1994. (Cited on page 172.)

[199] Yasuhiro Wada and Meng-Sing Liou. An Accurate and Robust Flux Splitting Scheme for Shock and Contact Discontinuities. *SIAM Journal on Scientific Computing*, 18(3):633–657, May 1997. (Cited on page 172.)

[200] Xia Wang and Xiaodong Sun. Hyperbolicity of One-Dimensional Two-Fluid Model With Interfacial Area Transport Equations. In *FEDSM2009*, Volume 1: Symposia, Parts A, B and C, 2009. (Cited on page 30.)

[201] Na Wei, Chaoyang Xu, Yingfeng Meng, Gao Li, Xiao Ma, and Anqi Liu. Numerical simulation of gas-liquid two-phase flow in wellbore based on drift flux model. *Applied Mathematics and Computation*, 338, 2018. (Cited on pages 156 and 160.)

[202] G. Welper. Interpolation of Functions with Parameter Dependent Jumps by Transformed Snapshots. *SIAM Journal on Scientific Computing*, 39(4):A1225–A1250, January 2017. (Cited on page 238.)

[203] G. Welper. Transformed Snapshot Interpolation with High Resolution Transforms. *arXiv:1901.01322 [math]*, January 2019. arXiv: 1901.01322. (Cited on page 238.)

[204] Anders Willersrud. *Model-Based Diagnosis of Drilling Incidents*. PhD thesis, NTNU, 2015. (Cited on pages ix and 18.)

[205] X.D. Zhang, J.-Y. Trepanier, and R. Camarero. A posteriori error estimation for finite-volume solutions of hyperbolic conservation laws. *Computer Methods in Applied Mechanics and Engineering*, 185(1), April 2000. (Cited on pages 157 and 184.)

[206] V. D. Zhibaedov, N. A. Lebedeva, A. A. Osiptsov, and K. F. Sinkov. On the hyperbolicity of one-dimensional models for transient two-phase flow in a pipeline. *Fluid Dynamics*, 51, 2016. (Cited on page 30.)

[207] W. Zhou, B. Hamroun, Y. Le Gorrec, and F. Couenne. Infinite Dimensional Port Hamiltonian Representation of reaction diffusion processes. *IFAC-PapersOnLine*, 48:476–481, 2015. (Cited on pages 38, 71, and 110.)

Summary

Structure-preserving model order reduction for drilling automation

The aim of the thesis is to develop a framework for multi-phase hydraulic modelling and model-complexity reduction for conservation laws governed by non-linear hyperbolic partial differential equations (PDEs) in the scope of drilling automation. This work is a first step towards obtaining accurate and simple enough (multi-phase) hydraulic models. The central goal is to develop model reduction techniques that enable effective handling of distributed nonlinearities and delays (due to pressure wave propagation), guarantee the preservation of key system properties (such as stability and input-output behavior), efficiently capture the relevant phenomena over the various spatio-temporal scales and solve at a good level of accuracy within a reasonable computing time. Such reduced-order models, developed by uniquely combining high predictive capacity and low complexity, will enable the usage of hydraulic models in virtual drilling scenario testing, estimator and controller design and drilling automation strategies for down-hole pressure management in the case of gas-influx into the well.

Hydraulics in a Managed Pressure Drilling (MPD) model can be characterized by interconnection of subsystems governed by single- and two-phase flow models, and, mathematical models governed by non-linear ordinary differential equations (for instance, a hydraulic model across a bit). We adopt a port-Hamiltonian (pH) framework for modelling and (structure-preserving) discretization of hydraulic systems since such formulations have several key properties that include passivity, shifted passivity, existence of Casimirs and compositionality. In particular, we develop pH representations (with respect to Stokes-Dirac structures) for single-phase flow models with spatially (and

temporally) varying cross-sections and for two-phase flow models, namely the Two-Fluid Model (TFM) and the Drift Flux Model (DFM), with constant and spatially (and temporally) varying cross-sections. We also exploit the property of compositionality and compose a network of pH systems for an MPD set-up via a power-preserving interconnection. Furthermore, we develop a framework for structure-preserving discretization of state-dependent, non-linear infinite-dimensional pH representations of two-phase flow models by employing (and extending) the concept of mixed Finite Element Methods and Gauss-Legendre collocation methods. Moreover, we perform the numerical analysis of the two-phase flow model, in particular the DFM, by using the principles of the Finite Volume Method, and numerically estimate the discretization error.

The response of single- and multi-phase hydraulic models is known to be dominated by wave propagation characteristics. Moving discontinuities (shock-fronts) are representative features of such class of models. These moving discontinuous features pose a major hindrance to obtain effective reduced-order model representations. Hence, we investigate and propose efficient, advanced and automated approaches to obtain reduced models, while still guaranteeing the accurate approximation of wave-front propagation (and wave interaction) phenomena (induced by slow or fast transients). We propose new model order reduction (MOR) approach(es) to obtain effective reduction for such transport-dominated problems or hyperbolic PDEs. We test our framework on challenging scenarios, such as multiple moving and possibly merging discontinuities, that embody the features of critical interest in the scope of multi-phase hydraulic models. Numerical case-studies show the potential of the approach in terms of computational accuracy compared with standard existing MOR techniques.

The two contributions of this thesis include: i) (compositional) modelling and structure-preserving discretization of pH systems, and, ii) MOR for transport-dominated problems as a drive towards obtaining an online-effective reduced-order representation. On the one hand, the structure-preserving reduced-order counterpart developed using a pH framework may be inherently stable/passive, but not of lowest possible dimension. On the other hand, the novel model reduction approach, built around the principle of the "method of freezing"/symmetry reduction, yields (nearly) online-effective reduced-order models, but is not necessarily stable. The two fields, treated separately in this work, have been pursued with a bigger goal in mind of future unification of these two aspects of minimality and stability in order to obtain a minimal structure-preserving reduced-order model for transport-dominated problems or hyperbolic PDEs, particularly characterized by moving discontinuous features for scalar and system of conservation laws.

Curriculum Vitae

Harshit Bansal was born on September 6, 1993 in Agra, Uttar Pradesh, India. He pursued the dual-degree program in Mechanical Engineering at the Indian Institute of Technology, Kharagpur, India. He obtained Bachelor of Technology (Honors) in Mechanical Engineering and Master of Technology in Mechanical Engineering with specialization in Mechanical Systems Design in 2016. After finishing his masters, he started his PhD in September 2016 (under the guidance of Prof. dr. Wil Schilders, Prof. dr. ir. Nathan van de Wouw and Dr. Laura Iapichino) in the Centre for Analysis, Scientific Computing and Applications (CASA) at the Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands. His PhD project falls under the aegis of Shell-NWO/FOM Programme in Computational Sciences for Energy Research. The main focus of the project is to develop structure-preserving model order reduction techniques in the scope of drilling automation. The results are presented in this dissertation.

Harshit's research interests include port-Hamiltonian modelling, structure-preserving discretization and model order reduction of distributed parameter systems and transport-dominated problems.

List of Publications

Harshit Bansal has the following publications:

Journals

Under Review/to-be submitted:

- H. Bansal, P. Schulze, M. H. Abbasi, H. Zwart, L. Iapichino, W.H.A. Schilders, N. van de Wouw, Port-Hamiltonian Formulation of Two-phase Flow Models, submitted to Systems and Control Letters (under review)

- H. Bansal, L. Iapichino, W.H.A. Schilders, N. van de Wouw, Numerical Analysis of the Drift Flux Model, to be submitted.

In preparation:

- H. Bansal, S. Weiland, L. Iapichino, W. Schilders, N. van de Wouw, Structure-preserving Discretization of Two-Phase Flow Models

- H. Bansal, S. Rave, L. Iapichino, W. Schilders, N. van de Wouw, Model order reduction for transport-dominated problems

Peer Reviewed Conference Proceedings (Full papers)

- H. Bansal, S. Weiland, L. Iapichino, W. Schilders, N. van de Wouw, Structure-preserving Spatial Discretization of a Two-Fluid Model, Accepted for 59th Conference on Decision and Control (CDC 2020)

- E.J.I Hoeijmakers, H. Bansal, T.M. van Opstal, P.A. Bobbert, Reduced order modelling for wafer heating with the Method of Freezing, Accepted for Proceedings of 13th International Conference on Scientific Computing in Electrical Engineering (SCEE 2020), to be published in Springer series "Scientific Computing in Electrical Engineering"

- H. Bansal, H. Zwart, L. Iapichino, W. Schilders, N. van de Wouw, Port-Hamiltonian modelling of fluid dynamics models with variable cross-section, Accepted for 24th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2020)

- M. H. Abbasi, H. Bansal, L. Iapichino, W. Schilders, N. van de Wouw, Power-Preserving Interconnection of Single- and Two-Phase Flow Models for Managed Pressure Drilling, 2020 American Control Conference (ACC 2020)

- H. Bansal, S. Rave, L. Iapichino, W. Schilders, N. van de Wouw, Model order reduction framework for problems with moving discontinuities, Accepted for European Numerical Mathematics and Advanced Applications Conference 2019 Proceedings (ENUMATH 2019)

Technical Reports (Non-Refereed)

- SciSports: Learning football kinematics through two dimensional tracking data, Proceedings of the 127th European Study Group Mathematics with Industry, 2018
  (with Anatoliy Babic, Gianluca Finocchio, Julian Golak, Mark Peletier, Jim Portegies, Clara Stegehuis, Anuj Tyagi, Roland Vincze and William Weimin Yoo)

- Synopsys: Latency prediction for on-chip communication, Proceedings of the 148th Study Group Mathematics with Industry, 2019
  (with Xingang Cao, Kristof Cools, Nhung Dang, Yingjun Deng, Tieme Goedendorp, Anastasiia Hraivoronska, Amir Parsa Sadr, Mikola Schlottke, Oliver Sheridan-Methven and Niels van der Wekken)

Acknowledgments

Writing this part already makes me feel nostalgic as it marks the end of a beautiful journey that had its share of ups and downs. The last four years have been extremely memorable for me, and this would not have been possible without the support of many people, both in a professional and personal capacity.

I would not have applied for the Shell NWO/FOM Ph.D. programme had it not been an indirect peer influence. *Shashank Maurya*, my friend from IIT Kharagpur, you made me to apply for the position at the final hour of the deadline. Thanks for that push!! I would also like to thank *Koondi* for his suggestions during the selection process, and for helping me get an understanding of the environment at CASA, which was useful in making the right career choice. I would also like to thank *Pushkar Mondal*, hostel mate from my undergraduate days. I believe that the time I spent with you brought out the researcher within me.

I would like to express my gratitude to my promotors, *Prof. Wil Schilders* and *Prof. Nathan van de Wouw*, and co-promotor *Dr. Laura Iapichino*. I could not have asked for a better combination of supervisors.

Wil, you played an instrumental role in shaping me to an individual I am today. Apart from interesting and valuable discussions on the topics pertaining to this dissertation, you gave me an opportunity to organize several workshops, mini-symposia in international conferences, summer school, etc., and this helped me both professionally and personally. You helped me to extend the network and get the scientific support from the experts in their respective fields. We also often talked about other mathematical problems, including numerical linear algebra, exascale computing, etc., and I learnt a lot from you.

Nathan, it was great to have a promotor and a supervisor like you. Your involvement, profound knowledge, and detailed critical feedback on reports/-papers/thesis helped me a lot to improve the quality of the scientific outputs. The coaching meetings with you were always very fruitful. It challenged the

thinking and enhanced to develop and formulate the theory in a better and rigorous manner. Meeting the thesis submission deadline would not have been possible without your support and a gentle push. Thanks for everything!!

Laura, you were more a friend than a supervisor to me. You know the equation we had between us, both at a professional and a personal level, and I am thankful to you for creating the atmosphere so that I could open-up with you. You helped me to extend the scientific network, and supported me whenever I needed you. You were the one who suggested me to attend *MoRePaS 2018* in Nantes, which, in hindsight, was one of the best advice. Your support was valuable during the writing phase of the thesis. There were times when a lot of things were going on in your personal life, but you did your best so as not to affect my Ph.D. timeline in any way. Thank you for being there!

I also had the pleasure of collaborating with experts from their respective fields. This dissertation would not have been the same without their guidance.

Prof. Hans Zwart, I want to thank you for insights into operator theory and for helping to make the theorems on Stokes-Dirac structures mathematically rigorous. The discussions with you helped me to get over my fear of functional analysis, and now I feel a bit more comfortable while reading related text.

Dr. Stephan Rave, we started collaborating in January 2018, and it has been a nice experience to work under your guidance. I learned a lot from your previous experience on a similar problem and acquired a better understanding of some of the core concepts of model order reduction.

Prof. Siep Weiland, I am very thankful to you for giving me an opportunity to discuss about port-Hamiltonian systems with you during the late phase of the Ph.D., and for providing me a better and thorough understanding of the finite-dimensional Dirac structures that were the basis of structure-preserving discretization methodology I developed.

Philipp Schulze, I have always enjoyed discussing model order reduction, port-Hamiltonian systems, and other topics with you. It was a good experience to collaborate with you on the modelling of port-Hamiltonian systems, and I still remember how it began from a bar at Glasgow during *ECCM-ECFD 2018*.

Prof. Barry Koren, I had a few meetings with you to discuss the numerical methods in the scope of the two-phase flow models, and I really appreciate your guidance since it helped me a lot to perform the required numerical analysis.

I also had the pleasure of visiting the *University of Münster, Germany*, and *Kelda Drilling Controls, Norway*, for a few weeks. Firstly, I would like to thank *Prof. Mario Ohlberger* for hosting me at the *University of Münster*, during two research visits, and for financially supporting my stay. The experience of staying at Münster was enhanced by Andreas Buhr, Felix Schindler, and Stephan Rave,

the *CASA Day* held in April 2019. I also enjoyed working on the Synopsys problem along with you, *Xingang*, and *Anastasiia*, during the *SWI 2019*. I still often recall and laugh on our inside jokes about Brownian Motion, Martingales, AM method, etc. *Rene*, also, thank you for keeping me informed about the happenings in the fluid dynamics community, particularly when some events used to be organized by *JMBC*. *Arthur*, the moving Wikipedia (actually more than that), during the first days at the group, I thought that you are the professor since you had knowledge across disciplines, but soon I realized that you are one among the PhDs, which made me respect you even more. *Anastasiia*, thank you for making me participate in several social activities, thus bringing a better balance in my life, and for sharing your opinion on a couple of versions of the design of the cover. *Lotte*, among a few joint activities, I will especially remember the time we spent at the Bowling alley during the *ENUMATH 2019*, and you know the reason. *Michiel*, thank you for sharing memes and cracking jokes to bring a smile on my face, particularly during the final phase of my Ph.D. Also, thank you very much for the discussions on some aspects related to numerical linear algebra. *Andrea Fuster*, I still remember getting a surprise gift from you in return for helping you with an instruction duty. That moment remains memorable for me even today since that was the first-ever gift that I got from a non-Indian. *Karen-Veroy*, I want to thank you for financially supporting my participation in an upcoming virtual conference. *Oliver*, thank you for sharing some insights on the topic of optimal transport. *Mark* and *Jim*, it was a good learning experience while working along with you during the *SWI 2018*.

I had the pleasure of sharing the office with several colleagues. In particular, I would like to thank *Amir*, *Faeze*, *Giulia*, *Perfect* for making the office my *casa*. *Amir*, you made me inculcate a habit of working until late in the office. You (and *Sajad*) tried a lot to make me join the Sports Centre while you were at the office, but could not succeed :P *Faeze*, you changed the dynamics of the office. Remember our neighboring officemates saying that we only laugh and do not work? :P. Apart from the work-related discussions, late evenings, quite often, used to be the fun part for various reasons, and I will really miss those moments. Also, thanks for your inputs on the design of the cover of my thesis. *Giulia*, the COVID-19 pandemic forced us to work from home, but it turned out to be a boon for our friendship, which strengthened over the period. Thank you for de-stressing me after my thesis submission and for taking care of the plants that *Nhung* had given me. *Perfect*, it was nice to share the office with you, and I want to thank you for telling me about the NA Digest Archive.

CASA lunches and gatherings used to be a perfect platform to come together and get to know each other on a personal level. These would not have been fun