Automatica 49 (2013) 2087–2097

Contents lists available at SciVerse ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Steady-state performance optimization for nonlinear control systems of Lur'e type ☆

A. Pavlov [a], B.G.B. Hunnekens [b,1], N.v.d. Wouw [b], H. Nijmeijer [b]

[a] Statoil ASA, R&D center, P.B. 1004, NO-3905, Porsgrunn, Norway

[b] Eindhoven University of Technology, Department of Mechanical Engineering, P.O. Box 513, NL 5600 MB Eindhoven, The Netherlands

ARTICLE INFO

Article history:
Received 6 February 2012
Received in revised form
6 February 2013
Accepted 7 April 2013
Available online 14 May 2013

Keywords:
Steady-state behavior
Nonlinear systems
Performance optimization
Variable gain control

ABSTRACT

In this paper, we develop a methodology for the steady-state performance optimization, in terms of the sensitivity to disturbances, for Lur'e type nonlinear control systems. For linear systems, steady-state performance is well defined and related to frequency-domain characteristics. The definition and analysis of steady-state performance of nonlinear systems are, however, far from trivial. For a practically relevant class of nonlinear systems and disturbances, this paper provides a computationally efficient method for the computation of the steady-state responses and, therewith, for the efficient performance assessment of the nonlinear system. Based on these analysis tools, a strategy for performance optimization is proposed, which can be employed for the optimized tuning of system and controller parameters. The results are illustrated by application to a variable gain controlled short-stroke wafer stage of a wafer scanner.

1. Introduction

Steady-state performance of a control system relates to the sensitivity of its steady-state response to perturbations. For linear systems, powerful tools for performance evaluation exist, which are well-known among control engineers and have been crucial to the success of linear control in industrial practice. Linear control systems' performance is usually assessed by investigating frequency-domain characteristics, such as the sensitivity, process sensitivity, and complementary sensitivity functions. The power of these frequency-domain techniques hinges on the fact that the steady-state response is unique, and it can be easily computed in the frequency domain from the input and the corresponding frequency response functions. The latter fully characterizes the mapping from the input to the steady-state output due to the superposition principle. However, for nonlinear systems such a frequency response function is not defined, and, moreover, the superposition principle does not hold. As a consequence, steady-state performance analysis for nonlinear control systems is a challenging task and the majority of the works on nonlinear control focus on stability rather than performance.

Still, there are several tools available to evaluate the steady-state performance of nonlinear systems in the face of disturbances. For example it can be assessed through an L_2-gain between input and output (van der Schaft, 2000) or *ISS*-gain between input and state (Sontag, 2007) for the whole class of (bounded in the corresponding norm) inputs. These approaches provide an upper bound on the norm of the steady-state response given an upper bound on the norm of the input. The benefit of such an approach is that it is valid for a generic class of bounded disturbances. On the other hand, this generality results in conservative estimates when considering particular classes of disturbances. Moreover, such a bound tends to be rather conservative because estimated L_2- and *ISS*-gains for nonlinear systems are generally conservative.

Quite often we do have more knowledge about the disturbances than a mere bound on their magnitude. For example, in many practically relevant cases, disturbances can be modeled as being periodic. This is the case, for example, if perturbations are induced by a mass-unbalance in rotor dynamics systems (Huang, Chao, Kang, & Sung, 2002), due to narrow-band filtering of resonances in the system dynamics in e.g. mechanical systems, or due to periodicity of reference trajectories to be tracked. In addition to periodicity, from the physical properties of the system we may even know the shape of disturbances possibly parameterized in

☆ This research is financially supported by the Dutch Technology Foundation STW. The material in this paper was partially presented at the 47th IEEE Conference on Decision and Control (CDC), December 9–11, 2008, Cancun, Mexico. This paper was recommended for publication in revised form by Associate Editor Murat Arcak under the direction of Editor Andrew R. Teel.

E-mail addresses: alepav@statoil.com (A. Pavlov), B.G.B.Hunnekens@tue.nl (B.G.B. Hunnekens), n.v.d.wouw@tue.nl (N.v.d. Wouw), H.Nijmeijer@tue.nl (H. Nijmeijer).

[1] Tel.: +31 40 247 4828; fax: +31 40 246 1418.

0005-1098/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.automatica.2013.04.017

some way. In this case the estimates on the steady-state responses obtained through the L_2-gain or *ISS*-gain approach, which do not take into account this information, will be too conservative.

To overcome this problem and evaluate the steady-state performance in a more accurate way, an alternative approach is proposed in this paper. It is based on *numerical computation* of the steady-state solutions as opposed to *estimating* their quantitative characteristics as in the previously mentioned approaches. For example, performance of a controlled robot doing a finite number of repetitive tasks can be evaluated by computing its steady-state solutions for the corresponding periodic reference trajectories. In addition to this, with an efficient algorithm for numerical computation of the steady-state solution's sensitivity with respect to control system parameters, we can employ a gradient-like method to *optimize* system performance. Of course, the efficiency and practical feasibility of such an approach is based on the efficiency of the underlying numerical algorithms.

Even though this approach may sound straightforward, it is far from trivial to make it feasible from both theoretical and computational points of view, since we are dealing with *nonlinear* systems. To cope with this problem, we limit our analysis to the case of periodic excitations (disturbances or reference signals) and to the practically relevant class of Lur'e nonlinear control systems. As it has been mentioned above, the choice of periodic disturbances is practically relevant for a number of applications. Modeling disturbances as being periodic may usually involve some form of approximation. However, it is well worth adopting such an approximation if more accurate characterizations of the steady-state response can be obtained, when compared to employing an L_2-gain or *ISS*-gain approach. The choice of Lur'e systems is explained by simple and easily verifiable conditions under which such a system exhibits a unique periodic steady-state response to a periodic excitation (Pavlov, van de Wouw, & Nijmeijer, 2005; Yakubovich, 1964). It is exactly this property that will allow us to uniquely characterize steady-state performance. For nonlinear systems that are not necessarily of Lur'e type, some conditions for the existence of the unique periodic steady-state response to a periodic excitation can be found, for example, in Angeli (2002), Demidovich (1961) and Russo, di Bernardo, and Sontag (2010). For general nonlinear systems, in contrast, the steady-state response to periodic disturbances may not be well-defined: it may be non-unique (i.e. dependent on the initial conditions) and/or not periodic. Examples of Lur'e type systems include variable-gain controlled motion systems (Fromion & Scorletti, 2002; Heertjes, Schuurbiers, & Nijmeijer, 2009; Heertjes & van de Wouw, 2006; Jiang & Gao, 2002; van de Wouw, Pastink, Heertjes, Pavlov, & Nijmeijer, 2008; Zheng, Guo, & Wang, 2005), and mechanical systems with local nonlinearities such as friction or one-sided supports (Bonsel, Fey, & Nijmeijer, 2004).

In general, periodic steady-state responses of a nonlinear system can be computed using several methods. Well-known methods for calculation of periodic solutions include, for example, period solvers Ascher, Mattheij, and Russell (1995) and Parker and Chua (1988) (e.g. the shooting method and finite difference method), or routines using simple forward integration in time. However, these methods are in general computationally rather expensive. A computationally less expensive method is to use describing function methods (Khalil, 2002), but the disadvantage of these methods is that only a single-harmonic approximation of the response is obtained. In another approach, which is especially applicable to Lur'e systems, the periodic solution is computed iteratively through finding the response of the linear part of the system in the frequency domain, and computing the response of the nonlinearity in the time domain, see, e.g. Cardona, Lerusse, and Géradin (1998), Semlyen and Medina (1995) and Telang and Hunt (2001). This method is very efficient especially in combination with Fast Fourier Transforms, which are used for transitions between the time- and frequency domains.

In contrast to the results in the literature mentioned above, in this paper we, firstly, prove that under the same conditions that guarantee a unique periodic steady-state response of a Lur'e system to a periodic excitation, this iterative method will converge from any initial guess for the steady-state solution. Secondly, we provide estimates on the accuracy of the algorithm if higher harmonics are truncated in each algorithm step. Both of these contributions are essential for practical application of the algorithm. Thirdly, we prove that the sensitivity of the steady-state response with respect to control system parameters is a unique periodic steady-state solution of another Lur'e system satisfying the same conditions as the original system. Thus it can be computed using the same iterative mixed time–frequency domain algorithm. Efficient computation of both the steady-state solution and its sensitivity with respect to control parameters opens the possibility for a gradient-based strategy for the steady-state performance *optimization*.

The developed optimization method allows us to solve the challenging problem of performance-based tuning of a variable gain controller for the linear motion stage of a wafer scanner (Heertjes et al., 2009; Heertjes & van de Wouw, 2006). Linear motion systems, of which such a motion stage is an example, are nowadays still often controlled by linear proportional–integral–differential (PID) controllers. However, it is well known that linear controllers suffer from inherent performance limitations such as the waterbed-effect (Freudenberg, Middleton, & Stefanopoulou, 2000; Seron, Braslavsky, & Goodwin, 1997): an inherent trade-off between low-frequency tracking and sensitivity to high-frequency disturbances and measurement noise. To overcome such linear performance limitations, nonlinear PID control, also called variable gain control, has been employed (Fromion & Scorletti, 2002; Heertjes et al., 2009; Heertjes & van de Wouw, 2006; Jiang & Gao, 2002; van de Wouw et al., 2008; Zheng et al., 2005). In these references, it has been shown that variable gain control can outperform linear control strategies. However, the performance-based tuning of the variable gain controllers is far from trivial and typically done in an heuristic fashion. The last contribution of this paper is therefore the performance-based tuning of variable-gain controllers for a motion stage of an industrial wafer scanner. This method is based on the developed performance optimization method.

As can be seen, in this paper we pursue a *model-based* approach to performance optimization. Alternatively, also *data-based* approaches, such as e.g. extremum seeking (Krstić & Wang, 2000; Tan, Nešić, & Mareels, 2006), could be employed. The benefit of extremum seeking based approaches is that no accurate system and disturbance models need to be available. We note that in the scope of the application domain considered in this paper, i.e. control of high-precision motion stages in wafer scanners, accurate model information is typically available, which motivates the pursuit of a model-based approach. Moreover, a model-based approach is beneficial, firstly, in a system design-phase where no machine is available yet, secondly, in situations where performing many experiments becomes prohibitive and, thirdly, when performing parameter studies of the closed-loop system.

The paper is organized as follows. In Section 2, we will introduce the class of convergent Lur'e systems, and propose a method for steady-state performance analysis and optimization for such systems. Section 3 will present an efficient iterative numerical procedure for the computation of the steady-state responses and sensitivity of these steady-state responses to control system parameters. The theory will be applied in Section 4 to a variable gain controlled motion stage of a wafer scanner, to show the effectiveness of the proposed performance optimization strategy. Conclusions are presented in Section 5.

1.1. Notation and mathematical preliminaries

Throughout this paper, the following notation will be used. By $\mathbb{Z}$ we denote the set of integer numbers. By $L_2(T)$ we denote the space of piecewise-continuous real-valued T-periodic scalar functions $y(t)$ satisfying $\|y\|_{L_2} < +\infty$, where $\|y\|_{L_2}^2 := \frac{1}{T}\int_0^T |y(t)|^2 dt$. By l_2 we denote the space of complex-valued sequences $W = \{W[m]\}_{m\in\mathbb{Z}}$ satisfying $\|W\|_{l_2} < +\infty$, where $\|W\|_{l_2}^2 = \sum_{m\in\mathbb{Z}} |W[m]|^2$. Both $L_2(T)$ and l_2 are Banach spaces.

The sequence of Fourier coefficients of $y \in L_2(T)$ is denoted by Y. The elements of this sequence are given by

$$Y[m] = \frac{1}{T}\int_0^T y(t)e^{-im\omega t}dt, \quad m \in \mathbb{Z}, \tag{1}$$

where $\omega := 2\pi/T$. The inverse Fourier transform is given by

$$y(t) = \sum_{m\in\mathbb{Z}} Y[m]e^{im\omega t}. \tag{2}$$

For any $y \in L_2(T)$ and its Fourier coefficients Y, Parseval's equality holds:

$$\|y\|_{L_2} = \|Y\|_{l_2}. \tag{3}$$

For a linear single-input–single-output system

$$\dot{x} = Ax + Bu \tag{4}$$
$$y = Cx,$$

excited by a T-periodic input $u(t)$, $u \in L_2(T)$, if the matrix A is Hurwitz, there exists a unique globally exponentially stable T-periodic steady-state solution $\bar{x}_u(t)$ with the corresponding steady-state output $\bar{y}_u(t)$ ($\bar{y}_u \in L_2(T)$). Hence, system (4) defines a linear operator $\mathcal{G}_{yu} : L_2(T) \to L_2(T)$ according to $\mathcal{G}_{yu}u(t) = \bar{y}_u(t)$. In the frequency domain, we define the linear operator $\hat{\mathcal{G}}_{yu} : l_2 \to l_2$ that maps the Fourier coefficients U of the function $u(t)$ to the Fourier coefficients $\bar{Y}_U$ of the function $\bar{y}_u(t)$, i.e. $\hat{\mathcal{G}}_{yu}U := \bar{Y}_U$. It is known that

$$\left(\hat{\mathcal{G}}_{yu}U\right)[m] = G_{yu}(im\omega)U[m], \quad m \in \mathbb{Z}, \tag{5}$$

where $G_{yu}(s) := C(sI - A)^{-1}B$, $s \in \mathbb{C}$, is the transfer function of system (4) from input u to output y. Due to (5) it is straightforward to verify that

$$\|\hat{\mathcal{G}}_{yu}U\|_{l_2} \le \sup_{m\in\mathbb{Z}} |G_{yu}(im\omega)|\, \|U\|_{l_2} \tag{6}$$

and, by the Parseval's equality (3), we also conclude that

$$\|\mathcal{G}_{yu}u\|_{L_2} \le \sup_{m\in\mathbb{Z}} |G_{yu}(im\omega)|\, \|u\|_{L_2}. \tag{7}$$

2. Performance analysis and optimization

2.1. Convergent Lur'e systems

Let us consider Lur'e systems of the form

$$\dot{x} = Ax + Bu + Hw(t) \tag{8}$$
$$y = Cx + Dw(t) \tag{9}$$
$$u = -\varphi(y, w(t), \theta) \tag{10}$$
$$e = C_e x + D_e w(t), \tag{11}$$

where $x \in \mathbb{R}^n$ is the state, $y \in \mathbb{R}$ is the output, $w(t) \in \mathbb{R}^m$ is a piecewise-continuous input, and $e \in \mathbb{R}$ is a performance output. We assume that the nonlinearity $\varphi : \mathbb{R} \times \mathbb{R}^m \times \Theta \to \mathbb{R}$ is memoryless and may depend on n_θ parameters collected in the vector $\theta = [\theta_1, \ldots, \theta_{n_\theta}]^T \in \Theta \subset \mathbb{R}^{n_\theta}$. We also assume that $\varphi(0, w, \theta) = 0\ \forall w \in \mathbb{R}^m$ and $\theta \in \Theta$. For simplicity, we only consider the case in which the parameters θ appear in the nonlinearity φ and none of the system matrices. An extension to the situation where this is the case is relatively straightforward. The functions

$$G_{yu}(s) = C(sI - A)^{-1}B \tag{12}$$
$$G_{yw}(s) = C(sI - A)^{-1}H + D \tag{13}$$
$$G_{eu}(s) = C_e(sI - A)^{-1}B \tag{14}$$
$$G_{ew}(s) = C_e(sI - A)^{-1}H + D_e, \tag{15}$$

are transfer functions from inputs u and w to outputs y and e. In this paper, we consider the case of periodic disturbances $w(t)$. The following theorem provides conditions under which system (8)–(11) excited by a periodic input has a uniquely defined steady-state solution.

Theorem 1 (Yakubovich, 1964). *Consider system (8)–(11). Suppose*

A1 *The matrix A is Hurwitz;*

A2 *There exists a $K > 0$ such that the nonlinearity $\varphi(y, w, \theta)$ satisfies*

$$|\varphi(y_2, w, \theta) - \varphi(y_1, w, \theta)| \le K\,|y_2 - y_1|, \tag{16}$$

for all $y_1, y_2, w \in \mathbb{R}^m, \theta \in \Theta$;

A3 *The transfer function $G_{yu}(s)$ given by (12) satisfies*

$$\sup_{\omega\in\mathbb{R}} |G_{yu}(i\omega)| =: \gamma_{yu} < \frac{1}{K}. \tag{17}$$

Then for any $\theta \in \Theta$ and any T-periodic piecewise continuous input $w(t)$, system (8)–(11) has a unique T-periodic solution $\bar{x}_w(t, \theta)$, which is globally exponentially stable.

We will call $\bar{x}_w(t, \theta)$ the steady-state solution. Systems with a uniquely defined bounded globally asymptotically stable solution (for arbitrary bounded inputs $w(t)$) are called convergent (Demidovich, 1961; Pavlov et al., 2005). It can be shown that systems satisfying the conditions of Theorem 1 are also incrementally stable (Angeli, 2002) and contracting (Forni & Sepulchre, 2012; Lohmiller & Slotine, 1998). See Rüffer, van de Wouw, and Mueller (2013) for a detailed comparison between convergence and incremental stability. However, we will not need either of the above stability properties for the analysis in the paper; we will use only technical conditions A1–A3.

2.2. Steady-state performance analysis

Once the steady-state solution is uniquely defined, we can define a performance objective to quantify the steady-state performance of the system for a particular T-periodic input $w(t)$ and particular parameter θ. For example, it can be defined as

$$J(\theta) = \frac{1}{T}\int_0^T \bar{e}_w(t, \theta)^2 dt, \tag{18}$$

where $\bar{e}_w(t, \theta)$ is the value of the performance output corresponding to the steady-state solution. If we are interested in quantifying simultaneously the steady-state performance corresponding to several disturbances, $w_1(t), w_2(t), \ldots, w_N(t)$, with periods $T_1, \ldots, T_N$, we can choose a functional depending on all the corresponding steady-state performance outputs. For example, it can be defined as

$$J(\theta) = \sum_{k=1}^{N} \frac{1}{T_k}\int_0^{T_k} \bar{e}_{w_k}(t, \theta)^2 dt. \tag{19}$$

The choice of the performance objective strongly depends on needs of the particular application.

This or any other steady-state performance objective can be evaluated by computing the corresponding steady-state solutions. A numerical algorithm presented in Section 3 allows one to do it in a computationally very efficient way. This gives us a non-conservative method to evaluate the steady-state performance for convergent Lur'e systems. Moreover, it is a key enabler for performance optimization, as described in the next subsection.

2.3. Steady-state performance optimization

System (8)–(11) may represent a closed-loop nonlinear control system with θ being a vector of controller parameters. Ultimately we aim to optimize the steady-state performance of this system by tuning $\theta \in \Theta$. With the efficient numerical method for evaluation of steady-state performance, as presented later in Section 3, this can be done by computing the value of the performance objective $J(\theta)$ for a sufficiently dense grid of parameters θ inside the set Θ and then choosing an optimal θ. This approach, however, becomes computationally prohibitive if the set Θ is large or multidimensional. Also it becomes prohibitive if the performance objective depends on multiple steady-state solutions (as in (19)), which have to be computed for each value of θ.

To cope with this problem, we propose to use gradient-like optimization algorithms, which provide a direction for decrease of $J(\theta)$ based on the gradient of $\partial J/\partial\theta(\theta)$. This approach requires the computation of the gradient of $J(\theta)$. For the performance objective as in (18), the gradient equals

$$\frac{\partial J}{\partial \theta}(\theta) = \frac{2}{T}\int_0^T \bar{e}_w(t,\theta)\frac{\partial \bar{e}_w}{\partial \theta}(t,\theta)dt, \tag{20}$$

under the condition that $\bar{e}_w(t,\theta)$ is C^1 with respect to θ. Here, we see that to compute the gradient of $J(\theta)$ we need to know both $\bar{e}_w(t,\theta)$ and $\partial\bar{e}_w/\partial\theta(t,\theta)$. The following theorem provides, firstly, conditions under which $\bar{x}_w(t,\theta)$ (and therefore $\bar{e}_w(t,\theta)$) is C^1 with respect to θ, and, secondly, gives us an equation for the computation of $\partial\bar{e}_w/\partial\theta(t,\theta)$.

Theorem 2. *Under the conditions of Theorem 1, if the nonlinearity $\varphi(y,w,\theta)$ is C^1 for all $y \in \mathbb{R}$, $w \in \mathbb{R}^m$ and θ in the interior of Θ, then the steady-state solution $\bar{x}_w(t,\theta)$ is C^1 in θ. The corresponding partial derivatives $\partial\bar{x}_w/\partial\theta_i(t,\theta)$ and $\partial\bar{e}_w/\partial\theta_i(t,\theta)$ are, respectively, the unique T-periodic solution $\bar{\Psi}(t)$ and the corresponding periodic output $\bar{\mu}(t)$ of the system*

$$\dot{\Psi} = A\Psi + BU + BW_i(t) \tag{21}$$

$$\lambda = C\Psi \tag{22}$$

$$U = -\frac{\partial\varphi}{\partial y}(\bar{y}(t,\theta), w(t), \theta)\lambda \tag{23}$$

$$\mu = C_e\Psi, \tag{24}$$

where $W_i(t) = -\partial\varphi/\partial\theta_i(\bar{y}_w(t,\theta), w(t), \theta)$.

Proof. The proof can be found in the Appendix. □

One may be tempted to recognize in Theorem 2 a classical result on the sensitivity of solutions of differential equations with respect to its parameters, see e.g. Khalil (2002). However, such classical sensitivity result deals with the sensitivity of a solution corresponding to a particular initial condition with respect to parameters. Note that Theorem 2 is a result on the sensitivity of the *steady-state solution* $\bar{x}(t,\theta)$ and that if the parameter θ changes, then $\bar{x}(t,\theta)$ changes typically also at the initial time $t = 0$. Hence, the statement and proof of this theorem is essentially different from the classical sensitivity result and, consequently, needs an explicit statement and proof.

As follows from the proof of Theorem 2, system (21)–(24) with the T-periodic input $(W_i(t), \bar{y}(t,\theta), w(t))$ has the same form as system (8)–(11) and satisfies the same conditions of Theorem 1. Thus, after computing $\bar{y}_w(t,\theta)$ from system (8)–(11), we will know the input to the sensitivity system (21)–(24), and, using the same numerical method as for system (8)–(11), we will be able to compute the periodic steady-state solution $\partial\bar{x}_w/\partial\theta_i(t,\theta)$ of the sensitivity system and the corresponding output $\partial\bar{e}_w/\partial\theta_i(t,\theta)$.

Having computed $\bar{e}_w(t,\theta)$ and $\partial\bar{e}_w/\partial\theta(t,\theta)$, we can compute $J(\theta)$ (e.g. from (18) or (19)) and its Jacobian $\partial J/\partial\theta(\theta)$. With this, we can apply some gradient-based optimization method (such as steepest-descent or Quasi-Newton methods) to find an optimum of J over the set Θ.

3. Computation of periodic responses

One way of computing a steady-state solution of system (8)–(11) (or (21)–(24)) is simply to simulate it for an arbitrary initial condition for sufficiently long time. Since the steady-state solution is globally exponentially stable, the simulated solution will eventually converge to the steady-state solution with any desired accuracy. Still, this method is computationally expensive and alternative methods are needed to make the computation efficient, especially in the context of optimization. Such a method is presented in this section. A preliminary version of the results in Section 3.1 have been presented in Pavlov and van de Wouw (2008). To simplify notations, we will denote the steady-state solutions and the corresponding outputs by $\bar{x}(t)$, $\bar{y}(t)$ and $\bar{e}(t)$, omitting their dependency on $w(t)$ and θ, which are considered fixed.

3.1. Iterative computation of periodic responses with convergence and accuracy guarantees

The idea for the numerical method comes from the Banach fixed point theorem (Khalil, 2002; Kreyszig, 1978). As follows from the system equations (8)–(11), the steady-state output $\bar{y}(t)$ is a solution of the following equation

$$\bar{y} = \mathcal{G}_{yu} \circ \mathcal{F}\bar{y} + \mathcal{G}_{yw}w, \tag{25}$$

where $\mathcal{G}_{yu}$ and $\mathcal{G}_{yw}$ are the operators mapping $L_2(T) \to L_2(T)$ as defined in Section 1.1, and $\mathcal{F} : L_2(T) \to L_2(T)$ is the operator defined by $\mathcal{F}y(t) := -\varphi(y(t), w(t), \theta)$. We will show that the operator $\mathcal{G}_{yu} \circ \mathcal{F}$ is a contraction mapping. Thus, as follows from the Banach fixed point theorem, $\bar{y}$ – being the solution of (25) – can be found as the limit of the iterative process

$$y_{k+1} = \mathcal{G}_{yu} \circ \mathcal{F}y_k + \mathcal{G}_{yw}w \tag{26}$$

with *an arbitrary* initial value $y_0 \in L_2(T)$. This iterative process forms the core of the numerical method.

Let us show that $\mathcal{G}_{yu} \circ \mathcal{F}$ is a contraction mapping. Since $\mathcal{G}_{yu}$ is a linear operator, $\mathcal{G}_{yu}u_1 - \mathcal{G}_{yu}u_2 = \mathcal{G}_{yu}(u_1 - u_2)$. Applying (7) and (17) to the last equality, we conclude that

$$\|\mathcal{G}_{yu}u_1 - \mathcal{G}_{yu}u_2\|_{L_2} \le \gamma_{yu}\|u_1 - u_2\|_{L_2}, \tag{27}$$

for any $u_1, u_2 \in L_2(T)$. Consider the nonlinear operator $\mathcal{F}$. Due to condition A2 of Theorem 1, this operator satisfies

$$\|\mathcal{F}y_1 - \mathcal{F}y_2\|_{L_2} \le K\|y_1 - y_2\|_{L_2}. \tag{28}$$

From (27) and (28), we conclude that

$$\begin{aligned}\|\mathcal{G}_{yu} \circ \mathcal{F}y_1 - \mathcal{G}_{yu} \circ \mathcal{F}y_2\|_{L_2} &\le \gamma_{yu}\|\mathcal{F}y_1 - \mathcal{F}y_2\|_{L_2} \\ &\le \gamma_{yu}K\|y_1 - y_2\|_{L_2}.\end{aligned} \tag{29}$$

Finally, condition A3 in Theorem 1 implies $\gamma_{yu}K < 1$, which, together with (29), yields that $\mathcal{G}_{yu} \circ \mathcal{F}$ is a contraction mapping.

To implement the iterative process (26), we decompose it into the following equivalent one:

$$u_{k+1} = \mathcal{F}y_k \tag{30}$$

$$y_{k+1} = \mathcal{G}_{yu}u_{k+1} + \mathcal{G}_{yw}w. \tag{31}$$

Computationally it is cheaper to implement this algorithm in frequency domain by representing the T-periodic functions $u_k(t)$, $y_k(t)$ and $w(t)$ by their respective Fourier coefficients U_k, Y_k and W, and substituting the operators $\mathcal{G}_{yu}$, $\mathcal{G}_{yw}$ and $\mathcal{F}$ by their frequency-domain counterparts $\hat{\mathcal{G}}_{yu}, \hat{\mathcal{G}}_{yw}$ and $\hat{\mathcal{F}}$, respectively. Then the algorithm (30), (31) takes the form

$$U_{k+1} = \hat{\mathcal{F}} Y_k \tag{32}$$

$$Y_{k+1} = \hat{\mathcal{G}}_{yu} U_{k+1} + \hat{\mathcal{G}}_{yw} W. \tag{33}$$

Using inequalities (27), (28) and (29) and taking into account Parseval's equality (3), one can show that the operator $\hat{\mathcal{G}}_{yu} \circ \hat{\mathcal{F}}$ is a contraction on l_2 and by the Banach fixed point theorem, the iterative process (32), (33) will exponentially converge to the unique solution $\bar{Y}$ of the equation

$$\bar{Y} = \hat{\mathcal{G}}_{yu} \circ \hat{\mathcal{F}} \bar{Y} + \hat{\mathcal{G}}_{yw} W. \tag{34}$$

As a final step, we can calculate the steady-state performance output $\bar{E}$ as

$$\bar{E} = \hat{\mathcal{G}}_{eu} \circ \hat{\mathcal{F}} \bar{Y} + \hat{\mathcal{G}}_{ew} W. \tag{35}$$

The main advantage of such a frequency-domain implementation is the fact that the computation in (31), which is the most computationally demanding in the time-domain implementation, is now replaced by the computationally cheap step (33). Indeed, in (33) only the product of the Fourier coefficients U (W) and the frequency response function G_{yu} (G_{yw}) need to be calculated, see (5). However, frequency-domain implementation of the nonlinear operator $\hat{\mathcal{F}}$ becomes prohibitive. For a general nonlinearity $\varphi(y, w(t), \theta)$, it is impossible to find an analytic expression for the implementation of $\hat{\mathcal{F}}$. Therefore, it is suggested to firstly transform the Fourier coefficients Y_k to the periodic function $y_k(t)$ in the time domain using the inverse Fourier transform, compute $u_{k+1}(t) = -\varphi_{k+1}(y_k(t), w(t), \theta)$ in the time domain, and then apply the Fourier transform to transform $u_{k+1}(t)$ into U_{k+1}. Since the calculations are carried out iteratively in the time- and frequency domain we will refer to this algorithm as the Mixed-Time–Frequency (MTF) algorithm. Numerical implementation of this algorithm can be done very efficiently using Fast Fourier Transform algorithms (Brigham, 1976).

Practical implementation of the MTF algorithm will require truncations of U_{k+1} and W: due to the nonlinear operator $\hat{\mathcal{F}}$, the number of nonzero entries in U_{k+1} will, in general, be infinite. Moreover, the spectrum W of the periodic input $w(t)$ can have an infinite number of nonzero entries. So, we need to truncate U_{k+1} and W at each step. Another argument for truncation stems from using Fast Fourier Transforms. In practice, applying a Fast Fourier Transform operation, will always imply a truncation of U_{k+1}. Thus the MTF algorithm becomes

$$U_{k+1} = (\hat{\mathcal{F}} Y_k)_N \tag{36}$$

$$Y_{k+1} = \hat{\mathcal{G}}_{yu} U_{k+1} + \hat{\mathcal{G}}_{yw} (W)_N, \tag{37}$$

where $(\cdot)_N$ denotes a truncation operation:

$$(U)_N[m] = \begin{cases} U[m], & \text{for } |m| \le N \\ 0, & \text{for } |m| > N, \end{cases} \tag{38}$$

and $N > 0$ is a truncation parameter. In general, introduction of truncation in such an iterative algorithm can cause large errors in the limit solution and even prevent the convergence of the algorithm. However, in the next theorem we prove that, in fact, under the conditions of Theorem 1, the iterative sequence (36), (37) will converge for any value of the truncation parameter N. Moreover, we obtain an estimate on the accuracy of the algorithm with truncation.

Theorem 3. *Under the conditions of Theorem 1, for any $N > 0$ there is a unique limit $\bar{Y}^N$ for the sequence Y_k, $k = 1, 2, \ldots$, resulting from the iterative process with truncation (36), (37). Moreover,*

$$\|\bar{Y} - \bar{Y}^N\|_{l_2} \le \left\{ \sup_{|m|>N} |G_{yu}(im\omega)| \gamma_{yw} \frac{K\|W\|_{l_2}}{1 - \gamma_{yu} K} + \gamma_{yw} \|(W)_N^{res}\|_{l_2} \right\} \frac{1}{1 - \gamma_{yu} K}, \tag{39}$$

where $\gamma_{yw} := \sup_{m\in\mathbb{Z}} |G_{yw}(im\omega)|$ and $(W)_N^{res} := W - (W)_N$.

Proof. Notice that, as follows from (5), for any $U \in l_2$ it holds that $\hat{\mathcal{G}}_{yu}(U)_N = (\hat{\mathcal{G}}_{yu})_N U$, where $(\hat{\mathcal{G}}_{yu})_N : l_2 \to l_2$ is a linear operator defined as

$$(\hat{\mathcal{G}}_{yu})_N U[m] = \begin{cases} G_{yu}(im\omega) U[m], & \text{for } |m| \le N \\ 0, & \text{for } |m| > N. \end{cases} \tag{40}$$

Hence, instead of (36), (37) one can consider the equivalent iterative process

$$U_{k+1} = \hat{\mathcal{F}} Y_k \tag{41}$$

$$Y_{k+1} = (\hat{\mathcal{G}}_{yu})_N U_{k+1} + \hat{\mathcal{G}}_{yw} (W)_N, \tag{42}$$

which is of a similar form as (32), (33). So, in order to prove its convergence we only need to show that $(\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}}$ is a contraction mapping from l_2 to l_2.

It is straightforward to verify that

$$\|(\hat{\mathcal{G}}_{yu})_N U\|_{l_2} \le \sup_{|m|\le N} |G_{yu}(im\omega)| \|U\|_{l_2}. \tag{43}$$

Taking into account (17), we obtain $\|(\hat{\mathcal{G}}_{yu})_N U\|_{l_2} \le \gamma_{yu} \|U\|_{l_2}$. From this fact and from the linearity of $(\hat{\mathcal{G}}_{yu})_N$ we conclude that for any $U_1, U_2 \in l_2$ it holds that

$$\|(\hat{\mathcal{G}}_{yu})_N U_1 - (\hat{\mathcal{G}}_{yu})_N U_2\|_{l_2} \le \gamma_{yu} \|U_1 - U_2\|_{l_2}. \tag{44}$$

Using Parseval's equality (3) and (28) we conclude that

$$\|\hat{\mathcal{F}} Y_1 - \hat{\mathcal{F}} Y_2\|_{l_2} \le K \|Y_1 - Y_2\|_{l_2}, \tag{45}$$

for any $Y_1, Y_2 \in l_2$. In the same way as in (29), inequalities (44) and (45) imply

$$\|(\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}} Y_1 - (\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}} Y_2\|_{l_2} \le \gamma_{yu} K \|Y_1 - Y_2\|_{l_2}. \tag{46}$$

Since $\gamma_{yu} K < 1$ (see condition A3 in Theorem 1), the operator $(\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}}$ is a contraction. By the Banach fixed point theorem, there exists a unique $\bar{Y}^N \in l_2$ satisfying

$$\bar{Y}^N = (\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}} \bar{Y}^N + \hat{\mathcal{G}}_{yw} (W)_N, \tag{47}$$

and this solution $\bar{Y}^N$ can be found as a limit of the iterative sequence (41), (42) or, equivalently, of the sequence (36), (37).

It remains to show that the error bound in (39) holds. From (34) and (47), we conclude that

$$\begin{aligned} \|\bar{Y} - \bar{Y}^N\|_{l_2} &= \|\hat{\mathcal{G}}_{yu} \circ \hat{\mathcal{F}} \bar{Y} + \hat{\mathcal{G}}_{yw} W \\ &\quad - ((\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}} \bar{Y}^N + \hat{\mathcal{G}}_{yw} (W)_N)\|_{l_2} \\ &\le \|(\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}} \bar{Y} - (\hat{\mathcal{G}}_{yu})_N \circ \hat{\mathcal{F}} \bar{Y}^N\|_{l_2} \\ &\quad + \|(\hat{\mathcal{G}}_{yu})_N^{res} \circ \hat{\mathcal{F}} \bar{Y}\|_{l_2} + \|\hat{\mathcal{G}}_{yw} W_N^{res}\|_{l_2}, \end{aligned} \tag{48}$$

where $(\hat{\mathcal{G}}_{yu})_N^{res} := \hat{\mathcal{G}}_{yu} - (\hat{\mathcal{G}}_{yu})_N$ and $(W)_N^{res} = W - (W)_N$. Taking into account (46), we obtain

$$\|\bar{Y} - \bar{Y}^N\|_{l_2} \le \gamma_{yu} K \|\bar{Y} - \bar{Y}^N\|_{l_2} + \|(\hat{\mathcal{G}}_{yu})_N^{res} \circ \hat{\mathcal{F}} \bar{Y}\|_{l_2} + \|\hat{\mathcal{G}}_{yw} (W)_N^{res}\|_{l_2}. \tag{49}$$

Since $\gamma_{yu}K < 1$, it follows that

$$\|\bar{Y} - \bar{Y}^N\|_{l_2} \leq \frac{1}{1-\gamma_{yu}K}\left(\|(\hat{\mathcal{G}}_{yu})_N^{res} \circ \hat{\mathcal{F}}\bar{Y}\|_{l_2} + \|\hat{\mathcal{G}}_{yw}(W)_N^{res}\|_{l_2}\right). \tag{50}$$

Notice that $(\hat{\mathcal{G}}_{yu})_N^{res}$ is defined as

$$(\hat{\mathcal{G}}_{yu})_N^{res}U[m] = \begin{cases} 0, & \text{for } |m| \leq N \\ G_{yu}(im\omega)U[m], & \text{for } |m| > N. \end{cases} \tag{51}$$

Hence it can be easily verified that

$$\|(\hat{\mathcal{G}}_{yu})_N^{res}U\|_{l_2} \leq \sup_{|m|>N}|G_{yu}(im\omega)|\|U\|_{l_2}. \tag{52}$$

Since $\hat{\mathcal{F}}$ is Lipschitz with the Lipschitz constant K and $\hat{\mathcal{F}}0 = 0$ (this follows from the condition that $\varphi(0, w, \theta) = 0\ \forall\theta \in \Theta,\ w \in \mathbb{R}^m$, see Section 2.1), we obtain

$$\|\hat{\mathcal{F}}\bar{Y}\|_{l_2} \leq K\|\bar{Y}\|_{l_2}. \tag{53}$$

Next, let us determine a bound on $\|\bar{Y}\|_{l_2}$. Due to Parseval's equality $\|\bar{Y}\|_{l_2} = \|\bar{y}\|_{L_2}$, so we can estimate $\|\bar{y}\|_{L_2}$ instead. Since $\bar{y}$ satisfies (25), using the triangular inequality, it holds that

$$\|\bar{y}\|_{L_2} \leq \|\mathcal{G}_{yu} \circ \mathcal{F}\bar{y}\|_{L_2} + \|\mathcal{G}_{yw}w\|_{L_2}. \tag{54}$$

Applying inequality (29) for $y_1 = \bar{y}$ and $y_2 = 0$ and expressing $\|\bar{y}\|_{L_2}$, from (54) we obtain

$$\|\bar{y}\|_{L_2} \leq \frac{1}{1-\gamma_{yu}K}\|\mathcal{G}_{yw}w\|_{L_2}. \tag{55}$$

Finally, application of inequality (7) gives

$$\|\bar{y}\|_{L_2} \leq \frac{1}{1-\gamma_{yu}K}\gamma_{yw}\|w\|_{L_2}. \tag{56}$$

Uniting (50), (52), (53), and (56) with the Parseval's equality (3) we obtain (39). □

Remark 4. Using the Parseval's equality (3), in the time domain the accuracy estimate (39) takes the form

$$\|\bar{y} - \bar{y}^N\|_{L_2} \leq \left\{\sup_{|m|>N}|G_{yu}(im\omega)|\gamma_{yw}\frac{K\|w\|_{L_2}}{1-\gamma_{yu}K} + \gamma_{yw}\|w - w^N\|_{L_2}\right\}\frac{1}{1-\gamma_{yu}K}. \tag{57}$$

Note that the algorithm converges for *any* initial guess for the steady-state output solution and hence no a priori knowledge is needed on the steady-state solution. Moreover, note that the algorithm can be made as accurate as we desire. From (57) we see that for a given input function $w(t)$ and a given tolerance $\varepsilon > 0$ one can always choose the truncation parameter N such that $\|\bar{y}^N - \bar{y}\|_{L_2} \leq \varepsilon$. Namely, the transfer function $G_{yu}(s)$ is strictly proper (see (8)–(9)) such that one can always choose N sufficiently large to reduce $\sup_{|m|>N}|G_{yu}(im\omega)|$ to a desired level. Moreover, we can arbitrarily closely approximate a periodic signal $w \in L_2(T)$ by again choosing N sufficiently large, such that $\|w - w^N\|_{L_2}$ is sufficiently small (this follows from the Riesz–Fischer theorem, see for example Beals, 2004).

For a certain choice of N, at each step we only need to store complex-valued $2N + 1$-dimensional vectors (for frequencies ranging from $-N\omega$, up to $N\omega$). Moreover, if the transfer function $G_{yu}(s)$ has good filtering properties and the spectrum W of $w(t)$ has good roll-off for high frequencies, the number N characterizing the dimension of these vectors can be chosen rather small without significant deterioration of the algorithm accuracy. This is definitely a benefit for numerical implementation of this algorithm since a smaller N implies a smaller number of operations at each iteration of the algorithm, which makes the algorithm faster.

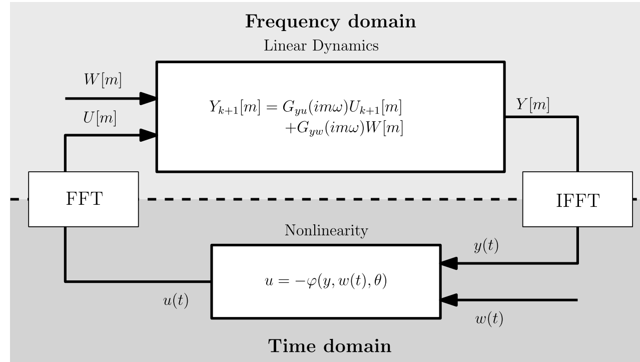

Fig. 1. Mixed-Time–Frequency algorithm to compute the steady-state solutions.

3.2. Numerical implementation of the MTF algorithm

For a numerical implementation of the algorithm (36), (37), we assume that the truncation parameter N is chosen in accordance with (39) to guarantee a desired accuracy of the algorithm. We choose a number $M = 2^b$ for some positive integer b and satisfying $M \geq 2N$, which we will use in the direct and inverse Fast Fourier Transform. In addition to this we will introduce a parameter $\epsilon_{\text{reltol}} > 0$ for stopping the iterative process (36), (37) if, for example,

$$\epsilon_Y := \frac{\|Y_k - Y_{k-1}\|_{l_2}}{\|Y_{k-1}\|_{l_2}} < \epsilon_{\text{reltol}}. \tag{58}$$

The iterative computation of the periodic response can be summarized by the following algorithmic steps, see Fig. 1:

(1) Set iteration index $k = 0$, $\epsilon_Y > \epsilon_{\text{reltol}}$;
(2) Compute the Fourier coefficients $(W)_N[m]$ of $w(t)$ using the Fast Fourier Transform (FFT);
(3) Choose any initial guess $Y_0[m]$ for the first N Fourier coefficients of $\bar{y}(t)$.
(4) Set $Y_0^{ext}[m] = Y_0[m]$ for $|m| \leq N$ and $Y_0^{ext}[m] = 0$ for $N < |m| \leq M/2$;
(5) Compute the time signal $y_0(t)$ corresponding to $Y_0^{ext}[m]$ using the Inverse Fast Fourier Transform (IFFT);
(6) while $\epsilon_Y > \epsilon_{\text{reltol}}$,
   (a) Evaluate the nonlinearity in the time domain:
   $u_{k+1}(t) = -\varphi(y_k(t), w(t), \theta)$; (59)
   (b) Compute the Fourier coefficients $U_{k+1}^{ext}[m]$ of $u_{k+1}(t)$ using the Fast Fourier Transform (FFT);
   (c) Set $U_{k+1}[m] = U_{k+1}^{ext}$ for $|m| \leq N$;
   (d) Evaluate the linear dynamics in the frequency domain:
   $Y_{k+1}[m] = G_{yu}(im\omega)U_{k+1}[m] + G_{yw}(im\omega)W[m]$, (60)
   for $|m| \leq N$;
   (e) Set $Y_{k+1}^{ext}[m] = Y_{k+1}[m]$ for $|m| \leq N$ and $Y_{k+1}^{ext}[m] = 0$ for $N < |m| \leq M/2$;
   (f) Compute the time signal $y_{k+1}(t)$ corresponding to $Y_{k+1}^{ext}[m]$ using the Inverse Fast Fourier Transform (IFFT);
   (g) Check termination criterion, for example (58). If satisfied, terminate algorithm, otherwise go back so step (6)(a);
   (h) set $k = k + 1$.
(7) Calculate the steady-state performance output

$$\bar{E} = G_{eu}(im\omega)U_k[m] + G_{ew}(im\omega)W[m]. \tag{61}$$

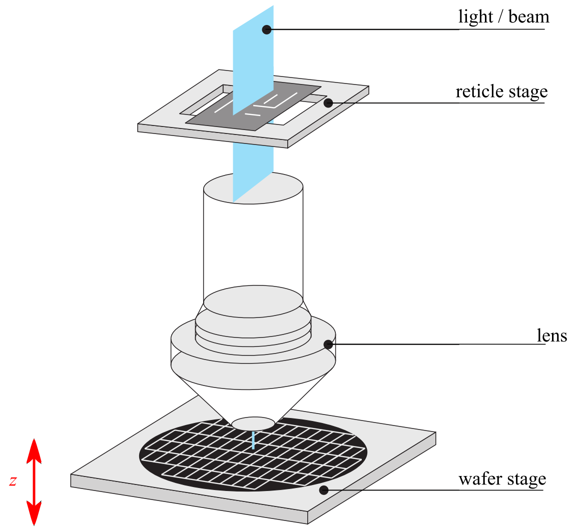

Fig. 2. The z-direction of the wafer stage will be controlled by a variable gain controller.

Note that the computation of the steady-state performance output $\bar{E}$ only has to be carried out once at the end, after convergence of the algorithm.

We approximate the direct and inverse Fourier transforms for continuous signals by discrete Fourier transform for sampled signals. The inaccuracy introduced by this approximation is not accounted for in the analysis in this section, but it can be reduced by increasing the parameter M.

The direct and inverse discrete Fourier transforms can be computed very efficiently using Fast Fourier Transform (FFT) algorithms, while (60) requires only a relatively small number of summations and multiplications. This makes the algorithm very efficient from a computational point of view as we will illustrate in a realistic engineering application in the next section.

4. Variable gain control of wafer stages

In this section, we will use the theory presented in the previous sections to optimize the performance of a variable gain controller for a benchmark system, namely the motion control of the z-direction of a short-stroke wafer stage of a wafer scanner, see Fig. 2, which is disturbed by force disturbances. Wafer scanners are used to produce integrated circuits (IC's). Light, emitted by a laser, falls on a reticle, which contains an image. This image is projected onto a wafer by passing through a lens. Due to this illumination, in combination with a photo-resist, a chemical reaction takes place which results in an image on the wafer, the IC's. This process requires positioning of the wafer stage in three degrees of freedom (x, y and z) with nm-accuracy. High-bandwidth linear controllers are used to achieve this. Due to the waterbed-effect (Freudenberg et al., 2000; Seron et al., 1997), low-frequency performance improvement (i.e. a higher bandwidth) goes hand in hand with high-frequency performance deterioration. Variable gain-control can be used to balance this trade-off in a more desirable manner (Heertjes et al., 2009; Heertjes & van de Wouw, 2006). However, it is a challenging problem to tune the variable gain controller parameters to optimize performance. In this section, we will apply the model-based performance optimization to tune the variable gain controller parameters in order to optimize the closed-loop performance. Exploiting the particular knowledge we have on the disturbances, we compute the steady-state solutions and their sensitivities in a numerically efficient way, using the algorithm developed in Section 3. Subsequent application of a gradient-based Quasi-Newton optimization algorithm lead us to an optimal choice of the controller parameters.

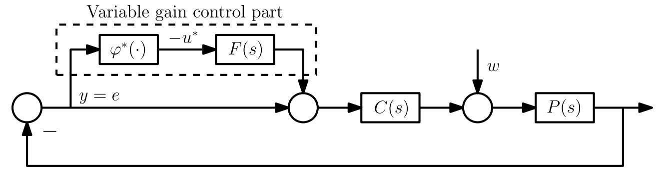

Fig. 3. Closed loop variable gain control scheme.

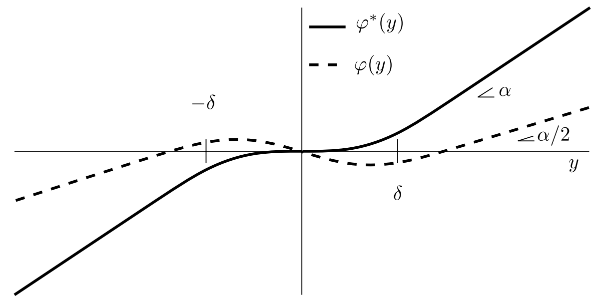

Fig. 4. Nonlinearity $\varphi^*(y)$ discriminating between small errors and large errors and the transformed nonlinearity $\varphi(y)$.

4.1. Variable gain control of linear motion systems

Consider Fig. 3 which shows a closed-loop variable gain control structure with plant $P(s)$, nominal linear controller $C(s)$, force disturbance w, filter $F(s)$, and variable gain element $\varphi^*(y)$. The performance output is the positioning error e, which is also the signal $y = e$ that is used in the nonlinearity $\varphi^*(y)$. The nonlinearity $\varphi^*(y)$ is based on a smooth variant of a dead-zone characteristic:

$$\varphi^*(y) = \alpha y - \delta\alpha \tanh(y/\delta), \tag{62}$$

where α is the additional gain and δ is the dead-zone length ($\theta = [\alpha, \delta]^T$). Note that the nonlinearity satisfies $0 \le \partial\varphi/\partial y \le \alpha$ $\forall y \in \mathbb{R}$, see Fig. 4. The particular choice for the variable gain element $\varphi^*(y)$ is key to our control design. For motion systems, errors induced by low-frequency disturbances are generally larger in amplitude than those induced by high-frequency disturbances. Therefore, if the error signal exceeds some pre-defined dead-zone level δ, an additional controller gain α is induced, yielding superior low-frequency disturbance suppression. If, however, the error signal does not exceed δ, only small additional gain is induced to avoid deterioration of the sensitivity to high-frequency disturbances. Parameters α and δ are the parameters that have to be tuned for performance.

Due to the choice of the variable gain controller structure, see Fig. 3, the closed-loop system inherently falls into the class of Lur'e-type systems, described by (8)–(11), where $y = e \in \mathbb{R}$ is the error (and performance output), $w(t)$ is a bounded scalar piecewise-continuous force disturbance, and the nonlinearity φ^* satisfying $0 \le \partial\varphi^*/\partial y \le \alpha$ is given by (62), see Fig. 4. Equivalently, we can loop-transform (see Khalil, 2002) the dynamics such that the transformed nonlinearity $\varphi = \varphi^* - \alpha y/2$ satisfies the symmetrical bound $|\partial\varphi/\partial y| \le \alpha/2$, see Fig. 4. The transfer functions $G_{yu}(s) = C(sI - A)^{-1}B$ between $u = -\varphi(y, \theta)$ and output y, and the transfer $G_{yw}(s) = C(sI - A)^{-1}H$ between force disturbance w and output y are given in terms of the plant $P(s)$, controller $C(s)$, and filter $F(s)$ by $G_{yw}(s) = -P(s)/(1 + P(s)C(s)(1 + \frac{\alpha}{2}F(s)))$ and $G_{yu}(s) = -C(s)F(s)G_{yw}(s)$, respectively. Because the performance output $y = e$, it holds that $G_{eu} = G_{yu}$ and $G_{ew} = G_{yw}$. Note that we can perform a similar loop-transformation for the sensitivity system (21)–(24). The transfer functions $G_{\lambda_i U_i}(s) = G_{\mu_i U_i}(s) = G_{\lambda_i W_i}(s) = G_{\mu_i W_i}(s) = G_{yu}(s)$.

A typical performance objective J for a motion system is the minimization of the squared steady-state error in a certain important time window $[t_s, t_e]$:

$$J(\bar{e}(\cdot,\theta)) = \frac{1}{t_e - t_s}\int_{t_s}^{t_e} \bar{e}(t,\theta)^2 dt, \tag{63}$$

where t_s and t_e are the starting time and ending time of the time interval, respectively. If (63) is used, the gradients $\partial J/\partial\theta_i$ are given by

$$\frac{\partial J}{\partial \theta_i} = \frac{2}{t_e - t_s}\int_{t_s}^{t_e} \bar{e}(t,\theta)\frac{\partial \bar{e}}{\partial \theta_i} dt. \tag{64}$$

4.2. Model specification and disturbance modeling of the wafer stage

The plant dynamics is modeled by the transfer function

$$P(s) = \frac{m_1 s^2 + bs + k}{s^2(m_1 m_2 s^2 + b(m_1+m_2)s + k(m_1+m_2))}, \tag{65}$$

$s \in \mathbb{C}$, where the following numerical values are used for the plant model (Heertjes et al., 2009): $m_1 = 5$ kg, $m_2 = 17.5$ kg, $k = 7.5 \cdot 10^7$ N/m, $b = 90$ Ns/m. The nominal low-gain ($\alpha = 0$) controller $C(s) = C_{\text{PID}}(s)C_{\text{lp}}(s)C_n(s)$ consists of a PID controller $C_{\text{PID}}(s)$, a second-order low-pass filter $C_{\text{lp}}(s)$ and a notch filter $C_n(s)$ to suppress the plant resonance. The filters are given by: $C_{\text{PID}}(s) = (k_p(s^2 + (\omega_i + \omega_d)s + \omega_i\omega_d))/(\omega_d s)$, where $k_p = 6.9 \cdot 10^6$ N/m is a loop gain, $\omega_d = 3.8 \cdot 10^2$ rad/s is the cutoff frequency of the differential action, and $\omega_i = 3.14 \cdot 10^2$ rad/s is the cutoff frequency of the integral action; $C_{\text{lp}}(s) = \omega_{\text{lp}}^2/(s^2 + 2\beta_{\text{lp}}\omega_{\text{lp}}s + \omega_{\text{lp}}^2)$, where $\omega_{\text{lp}} = 3.04 \cdot 10^3$ rad/s is the cutoff frequency of the low-pass filter, $\beta_{\text{lp}} = 0.08$ is the dimensionless damping coefficient; $C_n(s) = (\omega_p/\omega_z)^2(s^2 + 2\beta_z\omega_z s + \omega_z^2)/(s^2 + 2\beta_p\omega_p s + \omega_p^2)$, where $\omega_p = 5.03 \cdot 10^3$ rad/s is the frequency of the poles of the notch with damping $\beta_p = 0.88$, and $\omega_z = 4.39 \cdot 10^3$ rad/s is the frequency of the zeros of the notch with damping $\beta_z = 2.7 \cdot 10^{-3}$. The loop-shaping filter $F(s)$ is given by $F(s) = (\omega_{p,F}/\omega_{z,F})^2(s^2 + 2\beta_{z,F}\omega_{z,F}s + \omega_{z,F}^2)/(s^2 + 2\beta_{p,F}\omega_{p,F}s + \omega_{p,F}^2)$, with $\omega_{p,F} = \omega_{z,F} = 2000$ rad/s, $\beta_{p,F} = 4.8$, and $\beta_{z,F} = 0.6$. Note that these filters, together with a certain value for α, define the transfer functions G_{yu} and G_{yw}. We aim to optimize the design of the additional gain α and dead-zone length δ of the variable gain controller element $\varphi(y)$, see Fig. 4, in the next section.

The z-direction of the wafer stage should be kept in focus, therefore we need to track a zero-reference signal. Force-disturbances $w(t)$ are a dominant source of disturbances for the z-direction of the wafer stage. These force disturbances can be considered to have two main contributions $u_{FFz}(t)$ and $u_p(t)$, such that

$$w(t) = u_{FFz}(t) + u_p(t), \tag{66}$$

where $u_{FFz}(t)$ is a mainly low-frequency contribution (below the bandwidth), and $u_p(t)$ is a high-frequency contribution (above the bandwidth).

Low-frequency force disturbance u_{FFz}. Because the wafer stage is undergoing large accelerations in the horizontal x- and y-direction (around 28.5 m/s^2), the feed-forward forces acting in the horizontal plane to realize such set-points affect on the z-direction due to unavoidable mechanical cross-talk. Based on a 3rd-order polynomial reference signal $x_d(t)$ in the x-direction, the force-disturbance u_{FFz} in the z-direction is modeled in the following way: the reference trajectory x_d is filtered by a feed-forward filter $FF_x(s)$ which transforms the position x_d to a feed-forward force u_{FFx} in the x-direction, and a static cross-talk factor

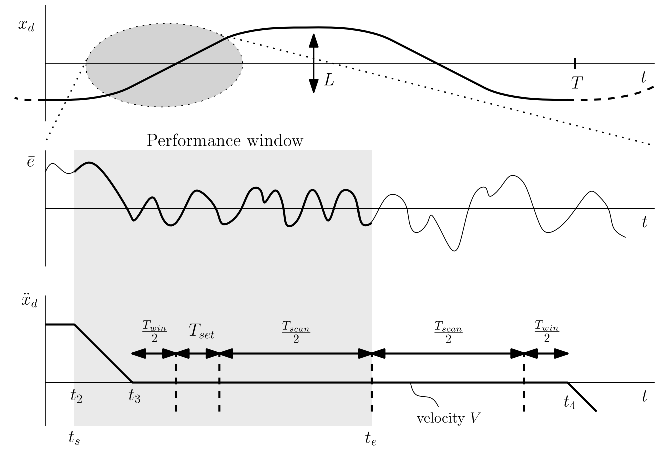

Fig. 5. Weighting interval for the performance objective J_{tot}.

γ_{ct} is used ($\gamma_{ct} = 4.5 \cdot 10^{-2}$, based on experimental data) to link the feed-forward force u_{FFx} in x-direction to the force disturbance u_{FFz} in z-direction. The filter $FF_x(s)$ is given as a 2nd-order high-pass filter $FF_x(s) = (\omega_{hp}^2 s^2)/(s^2 + 2\beta_{hp}\omega_{hp}s + \omega_{hp}^2)$, where $\omega_{hp} = 400\pi$ rad/s, and $\beta_{hp} = 0.5$. All IC's contained on a wafer are illuminated in the same way, over a scanning length L with scanning velocity V. This leads to periodic motion profiles that need to be carried out by the wafer stage in the horizontal plane. The T-periodic 3rd-order reference signal $x_d(t)$ is parameterized corresponding to values for maximum jerk and acceleration used in practice: $j_{\max} = 3000$ m/s^3, and $a_{\max} = 28.35$ m/s^2. The set-point $x_d(t)$, see Fig. 5, is fully determined by the following two variables:

- V, the scanning velocity during the constant velocity part $t_3 \le t \le t_4$;
- L, the scanning length during the scanning part.

Typical ranges for these parameters are $V \in [0.3, 0.6]$ m/s and $L \in [20 \cdot 10^{-3}, 40 \cdot 10^{-3}]$ m. The scanning takes place during the constant velocity part $t_3 \le t \le t_4$, but not this entire time interval is used for scanning. A certain time T_{win} is used to open the entire 'diaphragm' for exposure of the wafer ($T_{\text{win}}/2$ at the beginning and $T_{\text{win}}/2$ at the end of constant velocity part), and a certain settling time T_{set} is allowed for the wafer stage to settle after the acceleration phase. The rest of the time T_{scan} is used for scanning; therefore, the time span $t_4 - t_3 = T_{\text{win}} + T_{\text{set}} + T_{\text{scan}}$ (see Fig. 5), where $T_{\text{scan}} = L/V$ depends on the scanning length L and scanning velocity V, $T_{\text{set}} = 2 \cdot 10^{-3}$ s, and the time T_{win} needed to open the 'diaphragm' depends on the scanning velocity V, namely $T_{\text{win}} = \Delta L_{\text{win}}/V$, where $\Delta L_{\text{win}} = 5.5 \cdot 10^{-3}$ m, in agreement with machine parameters used in practice. Note that the scanning is repeated such that the force disturbance u_{FFz} is periodic with period time T, and this time T depends on the chosen scanning length L and scanning velocity V.

High-frequency force disturbance u_p. The sources of the high-frequency noise-disturbance are amplifier noise, and possibly other high-frequency disturbances such as e.g. perturbations stemming from the immersion process taking place on the wafer stage. In the immersion process, water is used to avoid the transition from lens to air which otherwise disturbs the illumination process. Continuously, water and air is supplied and removed, which causes the force-disturbances. Although a possibly noisy perturbation signal stemming from either of the above sources is in general not periodic, we can very well approximate it as being periodic with period time T of u_{FFz}. Note that this assumption can be justified if the noisy disturbances are of a significantly higher frequency than the frequency $1/T$ of the signal $u_{FFz}(t)$. This is indeed the case in practice and more numerical specifics will be given later. Moreover, it is well worth adopting such a modeling of the high-frequency noise because it allows for an explicit quantification of the steady-state performance. We model the high-frequency noise as a sum of N_p sinusoidal signals of constant amplitude A_p, frequencies $\omega_{p,j}$ and random phase angles $\phi_{p,j}$ such that

$$u_p = \sum_{j=1}^{N_p} A_p \sin(\omega_{p,j} t + \phi_{p,j}), \tag{67}$$

where $N_p = 50$, and $A_p = 0.12$ N, based on experiments, and the phase $\phi_{p,j} \in [0, 2\pi]$ is chosen randomly. Note that different values for $\phi_{p,j}$ lead to different realizations of the high-frequency noise u_p. The N_p frequencies in the signal are chosen as multiples of $1/T$ ($\approx$10 Hz) such that the total force disturbance $w(t)$ is a periodic signal with period time T. We define a vector of evenly distributed frequencies $\tilde{\omega}_{p,j}$ in the range $2\pi[200, 400]$ rad/s, based on experiments, with corresponding period times $\tilde{T}_j = 2\pi/\tilde{\omega}_{p,j}$. Subsequently, the noise frequencies $\omega_{p,j}$ are chosen as $\omega_{p,j} = \text{floor}(T/\tilde{T}_j)/T$, where the operation $\text{floor}(T/\tilde{T}_j)$ is defined as the largest integer smaller or equal to $T/\tilde{T}_j$. This assures that the frequencies $1/(2\pi\omega_{p,j})$ are multiples of the frequency $1/T$ of u_{FFz}. Note that for increasing $\tilde{\omega}_{p,j}$, the relative mismatch $(\omega_{p,j} - \tilde{\omega}_{p,j})/\tilde{\omega}_{p,j}$ becomes smaller, which motivates our choice for modeling the high-frequency noise as being periodic on the long time scale of the set-point.

4.3. Performance optimization of the variable gain controlled wafer stage

We aim to tune the dead-zone length δ and additional gain α of the variable gain controller (so $\theta = [\alpha, \delta]^T$) in order to optimize the performance (related to the error $e = y$, the performance output) for a whole range of disturbance situations, corresponding to different scanning lengths L, scanning velocities V, and high frequency noise realizations u_p. We do this because we aim at optimized performance for a wafer scanner that has to manufacture a wide range of different wafers corresponding to a wide range of set-points. Therefore, we want to use our efficient Mixed-Time–Frequency algorithm, discussed in Section 3, to compute the steady-state error signals $\bar{e}$ and steady-state sensitivities $\partial\bar{e}/\partial\alpha$ and $\partial\bar{e}/\partial\delta$. The following performance objective is used

$$J_{\text{tot}} = c \sum_{j_L=1}^{n_L} \sum_{j_V=1}^{n_V} \sum_{j_p=1}^{n_p} J(\bar{e}(t, \alpha, \delta, L_{j_L}, V_{j_V}, u_{p,j_p})), \tag{68}$$

with $c = 1/n_L n_V n_p$. Here, we use the squared steady-state error performance objective J defined in (63), $n_L = 10$ is the number of scanning lengths considered, evenly distributed such that $L_{j_L} \in [20 \cdot 10^{-3}, 40 \cdot 10^{-3}]$ m, $n_V = 10$ is the number of scanning velocities considered, evenly distributed such that $V_{j_V} \in [0.3, 0.6]$ m/s, and $n_p = 20$ is the number of realizations of u_p considered, leading to different realizations of the phase $\phi_{p,j}$ of the high-frequency disturbance u_p, see (67) (the random numbers are chosen from a uniform distribution). The time interval $t_s \le t \le t_e$ in (63) is indicated in Fig. 5, and is located around the time instance where scanning starts, and hence represents a key performance window in practice (Heertjes et al., 2009).

Note that we consider a finite set of disturbances for which we evaluate the performance. This set is representative for the range of tasks that are performed by the wafer scanner, allowing us to draw conclusions, in a practical sense, about the systems' performance.

We choose to optimize the variable-gain controller in the range $\alpha \in [\underline{\alpha}, \overline{\alpha}] = [0, 3]$, and $\delta \in [\underline{\delta}, \overline{\delta}] = [1 \cdot 10^{-10}, 1 \cdot 10^{-4}]$, which defines Θ. For these values, the conditions of Theorems 1 and 2

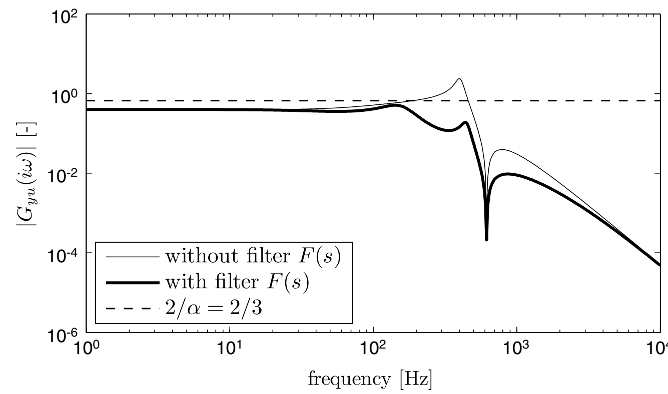

Fig. 6. Frequency-domain condition which shows $|G_{yu}(i\omega)| < 1/K = 2/\alpha \; \forall\omega \in \mathbb{R}$.

are all satisfied: we can verify that matrix A (or equivalently, the transfer function $G_{yu}(s)$) is Hurwitz $\forall\theta \in \Theta$, such that condition A1 of Theorem 1 is satisfied. The nonlinearity $\varphi(y, \alpha, \delta)$ satisfies $|\partial\varphi/\partial y| \le \alpha/2$, see Fig. 4. Moreover, $\varphi(0) = 0$, such that condition A2 of Theorem 1 holds. Condition A3 is satisfied for all $\alpha \in [0, 3]$ ($K = \alpha/2$ in (17)) as can be concluded from the limiting case $\alpha = 3$ shown in Fig. 6 (note that this figure shows the reason why the loop-shaping filter $F(s)$ is included). Because $\varphi(y, \alpha, \delta)$ is C^1 in the parameters, the conditions in Theorem 2 are satisfied. Thus, all conditions of Theorems 1 and 2 are satisfied such that the original system (8)–(11) and the sensitivity system (21)–(24) both exhibit unique bounded globally exponentially stable T-periodic steady-state solutions. Now $\bar{e}$ and $\partial\bar{e}/\partial\theta$ can be efficiently computed using the Mixed-Time–Frequency (MTF) algorithm presented in Section 3 (recall that the algorithm *always* converges to the unique steady-state solution for *any* initial guess).

The MTF algorithm has been implemented in Matlab (The MathWorks, Inc., 2010) and we use the following parameter values for our calculations: the number of points used to describe the response equals $M = 2N = 2^{13} = 8192$, the tolerance criterion (58) is used with $\epsilon_{\text{reltol}} = 1 \cdot 10^{-8}$, which together guarantee sufficient accuracy of the calculated responses. Furthermore, the model, controllers, nonlinearity, and disturbances as discussed in Section 4.2 are used. The steady-state error signals $\bar{e}$ (defining the performance objective J_{tot} in (68)) and steady-state sensitivities $\partial\bar{e}/\partial\alpha$ and $\partial\bar{e}/\partial\delta$ (defining $\partial J_{\text{tot}}/\partial\alpha$ and $\partial J_{\text{tot}}/\partial\delta$) are supplied to a gradient-based Quasi-Newton optimization algorithm to minimize the performance objective J_{tot} given by (68). The Quasi-Newton optimization routine is a second-order optimization routine which uses subsequent gradient information to build up curvature information on the Hessian using a BFGS update (Papalambros & Wilde, 2000). As an initial guess we choose $\alpha = 0.4$, and $\delta = 5 \cdot 10^{-8}$ m. The optimization converged in 15 iterations to the optimal variable gain controller with $\alpha = 3.000$, $\delta = 2.405 \cdot 10^{-8}$ m, and $J_{\text{tot}} = 4.857 \cdot 10^{-16}$ m^2, see Fig. 7. In this figure, for validation, the iteration history of the 2nd-order optimization is plotted together with the performance objective J_{tot} for a grid of values of α and δ. For $\alpha = 0$, we see a straight line of the performance objective J_{tot}; the setting $(\alpha, \delta) = (0, \delta)$, for arbitrary δ, corresponds to the low-gain control setting, because no additional gain is applied. Finally, for $\delta > 1 \cdot 10^{-6}$ the objective function J hardly changes anymore. This is due to the fact that in such a case the complete error stays within the dead-zone length δ. Therefore, hardly any additional gain is applied, such that for large δ the response will equal the response of a low-gain controller setting. The optimal variable gain controller outperforms the linear control limits by approximately 25% in terms of the performance objective J_{tot}, which is a very significant performance increase for this type of application in which nm-accuracy is required.

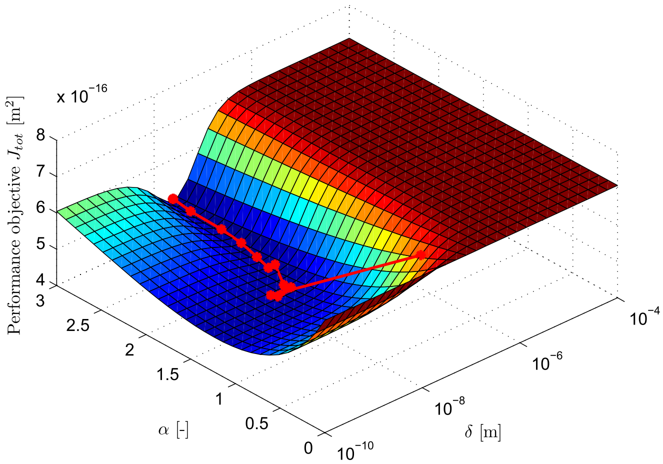

Fig. 7. Quasi-Newton optimization of the performance objective J_{tot} with the optimization surface for validation.

To emphasize the computational efficiency of the algorithm, note that if forward integration is used to compute a single steady-state response with similar accuracy, this takes approximately 20 s (on an Intel Core 2 Duo, 3 GHz processor). Opposed to a computational time of 0.1s for a single steady-state response using the Mixed-Time–Frequency algorithm, this is a factor 200 difference in computational time. Note that each combination of α and δ requires the calculation of $n_L n_V n_p = 10 \cdot 10 \cdot 20 = 2000$ steady-state solutions. The calculation of *a single point* out of 900 points in Fig. 7 therefore roughly takes $2000 \cdot 0.1$ s $\approx$ 3.3 min using our novel approach, while it would, hence, take 3.3 min $\cdot$ 200 $\approx$ 11 h using forward integration. Computing J_{tot} for all 900 points and then choosing optimal α and δ would take more than a year of computation time using forward integration. The same (brute force) optimization method, but based on the MTF algorithm for computing steady-state solutions reduces this time to approximately 3.3 min $\cdot$ 900 $\approx$ 50 h. Furthermore, if instead we apply the developed gradient-based algorithm, it takes only 2 h of computational time to arrive at the optimal δ and α. This comparison clearly demonstrates the benefits of the developed method.

5. Conclusions

In this paper, we have developed a method for the steady-state performance optimization for nonlinear control systems of Lur'e type. Accurate and efficient calculation of steady-state responses of periodically excited Lur'e systems using the Mixed-Time–Frequency algorithm has led to an efficient and non-conservative performance assessment of the nonlinear control system. Moreover, a gradient-based optimization strategy has been presented which can be used to tune system parameters of the nonlinear system to optimize the closed-loop performance. Remarkably, the same Mixed-Time–Frequency algorithm can be used for calculating the necessary gradients, by which an accurate and efficient performance optimization strategy is obtained. The results are applied to a variable gain controlled motion stage of a wafer scanner.

Appendix. Proof of Theorem 2

In this proof, we will use the following additional property of Lur'e systems (8)–(11) satisfying the conditions of Theorem 2.

Property 1. *Under the conditions of Theorem 1, if $\theta_1(h)$ converges to θ_2, as $h \to 0$, and T-periodic $w_1(t,h)$ converges to T-periodic $w_2(t)$ uniformly in $t \in [0,T]$, then the corresponding steady-state solution $\bar{x}_{w1(h)}(t,\theta_1(h))$ converges to $\bar{x}_{w2}(t,\theta_2)$ uniformly in $t \in [0,T]$ as $h \to 0$.*

This property follows from the fact that under the conditions of Theorem 1, system (8)–(11) is input-to-state convergent (Pavlov et al., 2005), which implies a continuous dependence of the steady-state solutions on the inputs in the uniform metric.

Now we will prove the theorem for the case of scalar θ. If θ is vector-valued, the proof can be repeated for each scalar component. In the proof we assume that $w(t)$ is fixed. For this reason we will denote the steady-state solution as $\bar{x}(t,\theta)$, without subscript w.

Consider θ in the interior of Θ and all sufficiently small h such that $\theta+h$ lies in Θ. Let us show that for $z(t,h) := \frac{1}{h}(\bar{x}(t,\theta+h) - \bar{x}(t,\theta))$ there exists the limit $\lim_{h\to 0} z(t,h)$, i.e. that $\bar{x}(t,\theta)$ is C^1 in θ.

As follows from the definition of the steady-state solution and from system Eqs. (8)–(11), for $h \neq 0, z(t,h)$ is a T-periodic function satisfying

$$\frac{dz(t,h)}{dt} = Az(t,h) + B\Delta_h\varphi \tag{A.1}$$

where $\Delta_h\varphi := -\frac{1}{h}(\varphi(\bar{y}(t,\theta+h), w(t), \theta+h) - \varphi(\bar{y}(t,\theta), w(t), \theta))$. Notice that $\Delta_h\varphi$ can be rewritten as

$$\begin{aligned}\Delta_h\varphi = &-\frac{1}{h}(\varphi(\bar{y}(t,\theta+h), w(t), \theta+h)\\ &-\varphi(\bar{y}(t,\theta), w(t), \theta+h))\\ &-\frac{1}{h}(\varphi(\bar{y}(t,\theta), w(t), \theta+h) - \varphi(\bar{y}(t,\theta), w(t), \theta)).\end{aligned} \tag{A.2}$$

Applying the mean value theorem to (A.2) and using the fact that

$$\frac{1}{h}(\bar{y}(t,\theta+h) - \bar{y}(t,\theta)) = Cz(t,h),$$

we obtain

$$\begin{aligned}\Delta_h\varphi = &-\frac{\partial\varphi}{\partial y}(\zeta(t,h), w(t), \theta+h)\cdot Cz(t,h)\\ &-\frac{\partial\varphi}{\partial\theta}(\bar{y}(t,\theta), w(t), \xi(t,h)),\end{aligned}$$

for some $\zeta(t,h) \in (\bar{y}(t,\theta), \bar{y}(t,\theta+h))$ and $\xi(t,h) \in (\theta,\theta+h)$, both of which can be chosen T-periodic. Combining this with (A.1), we conclude that for all sufficiently small $h \neq 0, z(t,h)$ is a T-periodic solution of the system

$$\begin{aligned}\dot{\Psi} &= A\Psi + BU + B\tilde{W}\\ \lambda &= C\Psi,\\ U &= -\frac{\partial\varphi}{\partial y}(\zeta(t,h), w(t), \theta+h)\lambda,\\ \tilde{W} &= \tilde{W}(t,h) := -\frac{\partial\varphi}{\partial\theta}(\bar{y}(t,\theta), w(t), \xi(t,h)).\end{aligned} \tag{A.3}$$

For fixed θ, we will consider $\zeta(t,h), h$ and $\tilde{W}(t,h)$ as inputs to system (A.3).

Let us show that system (A.3) satisfies the conditions of Theorem 1. Matrices A, B and C of the linear part of the system are the same as for system (8)–(11). Thus condition A1 is satisfied. Condition A2 with the same K as in (16) holds since $|\partial\varphi/\partial y(y,w,\theta)| \le K$ for all y, w and $\theta \in \Theta$. The latter inequality directly follows from (16) and the condition that $\varphi(y,w,\theta)$ is C^1 in y. Condition A3 holds automatically. Applying Theorem 1 to system (A.3), we conclude that for the T-periodic input $[\zeta(t,h), h, \tilde{W}(t,h)]$, system (A.3) has a unique T-periodic solution $\bar{\Psi}(t,h)$. Since $z(t,h)$ is a T-periodic solution of the same system, we conclude that $\bar{\Psi}(t,h) \equiv z(t,h)$ for all small $h \neq 0$.

As follows from Property 1, $\bar{y}(t,\theta+h)$ converges to $\bar{y}(t,\theta)$ uniformly in t, as $h \to 0$. By the definition of $\zeta(t,h)$ and $\xi(t,h)$, this implies that, as $h \rightarrow 0$, $\zeta(t,h)$ converges to $\bar{y}(t,\theta)$ and $\xi(t,h)$ converges to θ uniformly in t. Thus, all the inputs of system (A.3)—$\zeta(t,h)$, h and $\tilde{W}(t,h)$—converge, respectively, to $\bar{y}(t,\theta)$, 0 and $W(t,\theta) = -\frac{\partial\varphi}{\partial\theta}(\bar{y}(t,\theta), w(t), \theta)$ uniformly in t. Applying Property 1 to system (A.3), we conclude that $\bar{\Psi}(t,h)$ converges to $\bar{\Psi}(t,0)$, uniformly in t. Hence, $\lim_{h\to 0} z(t,h) = \lim_{h\to 0} \bar{\Psi}(t,h) = \bar{\Psi}(t,0)$. This proves that $\bar{x}(t,\theta)$ is C^1 in θ. Moreover, $\partial\bar{x}/\partial\theta(t,\theta) = \bar{\Psi}(t,0)$, which is the unique T-periodic solution of system (21)–(24). From this it is straightforward to show that $\partial\bar{e}/\partial\theta(t,\theta)$ is the corresponding T-periodic output $\bar{\mu}(t)$ of that system. □

References

Angeli, D. (2002). A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, *47*, 410–421.

Ascher, U., Mattheij, R., & Russell, D. (1995). *Numerical solution of boundary value problems for ordinary differential equations*. Philidelphia: SIAM.

Beals, R. (2004). *Analysis: an introduction*. New York: Cambridge University Press.

Bonsel, J., Fey, R., & Nijmeijer, H. (2004). Application of a dynamic vibration absorber to a piecewise linear beam system. *Nonlinear Dynamics*, *37*, 227–243.

Brigham, E. O. (1976). *The fast Fourier transform*. Englewood Cliffs: Prentice-Hall.

Cardona, A., Lerusse, A., & Géradin, M. (1998). Fast fourier nonlinear vibration analysis. *Computational Mechanics*, *22*(2), 128–142.

Demidovich, B. P. (1961). Dissipativity of a nonlinear system of differential equations. *Vestnik Moskovskogo Universiteta. Seriya I. Matematika, Mekhanika*, *Part 1–6*, 19–27.

Forni, F., & Sepulchre, R. (2012). A differential Lyapunov framework for contraction analysis. arXiv:1208.2943 [cs.SY], August.

Freudenberg, J., Middleton, R., & Stefanopoulou, A. (2000). A survey of inherent design limitations. In *Proc. of the American control conference* (pp. 2987–3001). Chicago, Illinois.

Fromion, V., & Scorletti, G. (2002). Performance and robustness analysis of nonlinear closed loop systems using μ_{nl} analysis: applications to nonlinear PI controllers. In *Proc. of 41st IEEE conf. on decision and control* (pp. 2480–2485). Las Vegas, Nevada, USA.

Heertjes, M. F., Schuurbiers, X. G. P., & Nijmeijer, H. (2009). Performance-improved design of N-PID controlled motion systems with application to wafer stages. *IEEE Transactions on Industrial Electronics*, *56*(5), 1347–1355.

Heertjes, M.F., & van de Wouw, N. (2006). Variable control design and its applications to wafer scanners. In *Proc. of the 45th IEEE conference on decision and control*, CDC2006. (pp. 3765–3770). San Diego, USA.

Huang, W.-Y., Chao, C.-P., Kang, J.-R., & Sung, C.-K. (2002). The application of ball-type balancers for radial vibration reduction of high-speed optic drives. *Journal of Sound and Vibration*, *250*(3), 415–430.

Jiang, F., & Gao, Z. (2002). An application of nonlinear PID control to a class of truck ABS problems. In *Proc. of 40th IEEE conf. on decision and control* (pp. 516–521). Orlando, Florida, USA.

Khalil, H. (2002). *Nonlinear systems*. Upper Saddle River, New Jersey: Prentice Hall.

Kreyszig, E. (1978). *Introductory functional analysis with applications*. New York: John Wiley & Sons.

Krstić, M., & Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, *36*(4), 595–601.

Lohmiller, W., & Slotine, J.-J. E. (1998). On contraction analysis for nonlinear systems. *Automatica*, *34*, 683–696.

Papalambros, P. Y., & Wilde, D. J. (2000). *Principles of optimal design: modeling and optimization*. Cambridge University Press.

Parker, T. S., & Chua, L. O. (1988). *Practical numerical algorithms for chaotic systems*. Springer-Verlag.

Pavlov, A., & van de Wouw, N. (2008). Fast computation of frequency response functions for a class of nonlinear systems. In *Proc. of IEEE conf. decision and control* (pp. 1180–1186). Cancun, Mexico.

Pavlov, A., van de Wouw, N., & Nijmeijer, H. (2005). *Uniform output regulation of nonlinear systems: a convergent dynamics approach*. Boston: Birkhäuser.

Rüffer, B. S., van de Wouw, N., & Mueller, M. (2013). Convergent systems vs. incremental stability. *Systems & Control Letters*, *62*(3), 277–285.

Russo, G., di Bernardo, M., & Sontag, E. D. (2010). Global entrainment of transcriptional systems to periodic inputs. *PLoS Computational Biology*, *6*(2).

Semlyen, A., & Medina, A. (1995). Computation of the periodic steady state in systems with nonlinear components using a hybrid time and frequency domain methodology. *IEEE Transactions on Power Systems*, *10*(3), 1498–1504.

Seron, M. M., Braslavsky, J. H., & Goodwin, G. C. (1997). *Fundamental limitations in filtering and control*. London: Springer-Verlag.

Sontag, E. D. (2007). Input to state stability: basic concepts and results. In *Nonlinear and optimal control theory* (pp. 163–220). Springer-Verlag.

Tan, Y., Nešić, D., & Mareels, I. M. Y. (2006). On non-local stability properties of extremum seeking control. *Automatica*, *42*(6), 889–903.

Telang, N. P., & Hunt, L. R. (2001). Frequency domain computations for nonlinear steady-state solutions. *IEEE Transactions on Signal Processing*, *49*(8), 1728–1733.

The MathWorks, Inc., (2010). http://www.mathworks.com.

van de Wouw, N., Pastink, H. A., Heertjes, M. F., Pavlov, A. V., & Nijmeijer, H. (2008). Performance of convergence-based variable-gain control of optical storage drives. *Automatica*, *44*, 15–27.

van der Schaft, A. J. (2000). *Springer communications and control engineering series, L2-gain and passivity techniques in nonlinear control* (2nd ed.). London: Springer-Verlag.

Yakubovich, V. A. (1964). The matrix inequality method in the theory of the stability of nonlinear control systems—I. The absolute stability of forced vibrations. *Automation and Remote Control*, *7*, 905–917.

Zheng, J., Guo, G., & Wang, Y. (2005). Nonlinear PID control of linear plants for improved disturbance rejection. In *16th IFAC world congress*, IFAC 2005. Prague.

A. Pavlov received M.Sc. degree (cum laude) in Applied Mathematics from St. Petersburg State University, Russia (1998), and Ph.D. degree in Mechanical Engineering from Eindhoven University of Technology, The Netherlands (2004). He held researcher positions at the Institute for Problems of Mechanical Engineering, Russian Academy of Sciences (1999), Ford Research Laboratories, USA (2000), Eindhoven University of Technology (2005) and Norwegian University of Science and Technology (2005–2009). Since 2009 he has been a principal researcher in Statoil Research Centre in Porsgrunn, Norway, working on automation of oil and gas production and drilling.

Alexey has published more than 60 refereed papers, 2 patents and a book "Uniform Output Regulation of Nonlinear Systems: A Convergent Dynamics Approach" co-authored with N.van de Wouw and H. Nijmeijer (Birkhauser, 2005). In 2004 and 2007 he was selected as outstanding reviewer for Automatica. In 2011 he received IFAC World Congress application paper prize (with J.-M. Godhavn, G.-O. Kaasa and N.L. Rolland). His current research interests are focused on automatic control technologies in oil and gas industry.

B.G.B. Hunnekens received the M.Sc. degree (cum laude) in Dynamics and Control from Eindhoven University of Technology, Eindhoven, the Netherlands, in 2011. He is currently pursuing the Ph.D. degree at the Department of Mechanical Engineering, Eindhoven University. His current research interests are nonlinear control of linear motion systems and performance-optimal nonlinear controller synthesis.

N.v.d. Wouw was born in 1970 and obtained his M.Sc.-degree (with honors) and Ph.D.-degree in Mechanical Engineering from the Eindhoven University of Technology, Eindhoven, the Netherlands, in 1994 and 1999, respectively. From 1999 until 2013 he has been affiliated with the Department of Mechanical Engineering of the Eindhoven University of Technology in the group of Dynamics and Control as an assistant/associate professor. In 2000, Nathan van de Wouw has been working at Philips Applied Technologies, Eindhoven, The Netherlands, and, in 2001, he has been working at the Netherlands Organisation for Applied Scientific Research (TNO), Delft, The Netherlands. He has held positions as a visiting professor at the University of California Santa Barbara, USA, in 2006/2007, at the University of Melbourne, Australia, in 2009/2010 and at the University of Minnesota, USA, in 2012 and 2013. Nathan van de Wouw has published a large number of journal and conference papers and the books 'Uniform Output Regulation of Nonlinear Systems: A Convergent Dynamics Approach' with A.V. Pavlov and H. Nijmeijer (Birkhauser, 2005) and 'Stability and Convergence of Mechanical Systems with Unilateral Constraints' with R.I. Leine (Springer-Verlag, 2008). He currently is an Associate Editor for the journal "Automatica". His current research interests are the analysis and control of nonlinear/non-smooth systems and networked control systems.

H. Nijmeijer (1955) obtained his M.Sc.-degree and Ph.D.-degree in Mathematics from the University of Groningen, Groningen, the Netherlands, in 1979 and 1983, respectively. From 1983 until 2000 he was affiliated with the Department of Applied Mathematics of the University of Twente, Enschede, the Netherlands. Since 2000, he is a full professor at Eindhoven, and chairs the Dynamics and Control group. He was editor in chief of the Journal of Applied Mathematics until 2009, corresponding editor of the SIAM Journal on Control and Optimization, and board member of the International Journal of Control, Automatica, Journal of Dynamical Control Systems, International Journal of Bifurcation and Chaos, International Journal of Robust and Nonlinear Control, Journal of Nonlinear Dynamics and the Journal of Applied Mathematics and Computer Science, and Hybrid Systems. He is a fellow of the IEEE since 2000 and was awarded in 1990 the IEE Heaviside premium. In the 2008 research evaluation of the Dutch Mechanical Engineering Departments the Dynamics and Control group was evaluated as excellent regarding all aspects (quality, productivity, relevance and viability). His research interests are in the broad scope of dynamics and control and their applications.