Euro. Jnl of Applied Mathematics (2008), *vol.* 19, *pp.* 479–506. © 2008 Cambridge University Press 479
doi:10.1017/S0956792508007547 Printed in the United Kingdom

Online parameter identification in time-dependent differential equations as a non-linear inverse problem

PHILIPP KÜGLER

Industrial Mathematics Institute, University of Linz, and RICAM, Austrian Academy of Sciences, A–4040 Linz, Austria
email: philipp.kuegler@jku.at

(*Received 9 January 2007; revised 24 March 2008; first published online 12 May 2008*)

Online parameter identification in time-dependent differential equations from time course observations related to the physical state can be understood as a non-linear inverse and ill-posed problem and appears in a variety of applications in science and engineering. The feature as well as the challenge of online identification is that sensor data have to be continuously processed during the operation of the real dynamic process in order to support simulation-based decision making. In this paper we present an online parameter identification method that is based on a non-linear parameter-to-output operator and, as opposed to methods available so far, works both for finite- and infinite-dimensional dynamical systems, e.g., both for ordinary differential equations and time-dependent partial differential equations. A further advantage of the method suggested is that it renders typical restrictive assumptions such as full state observability, linear parametrisation of the underlying model and data differentiation or filtering unnecessary. Assuming existence of a solution for exact data, a convergence analysis based on Lyapunov theory is presented. Numerical illustrations given are by means of online identification both of aerodynamic coefficients in a 3DoF-longitudinal aircraft model and of a (distributed) conductivity coefficient in a heat equation.

1 Introduction

Dynamical systems such as systems of ordinary differential equations (ODEs) or time-dependent partial differential equations (PDEs) play an important role in the modelling of instationary processes in science and engineering. Often, these models contain parameters – either constants or functions of dependent and independent variables such as time, spatial coordinates or state variables – that cannot be directly accessed and hence are unknown. Then one faces the (typically) non-linear and ill-posed inverse problem (see [8]) of indirectly determining those parameters from observations in the frequency or time domain related to the dynamical system. In case of time-domain data the problem of parameter identification can be treated either offline or online – depending on the application one has in mind. In offline identification, one first observes the dynamical system over a period of time and only then uses the data collected for the determination of the parameters. In online identification the parameters have to be identified simultaneously to the evolution of the real system and the data-collection process. Initial parameter guesses have to

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

be continuously improved since accurate parameter values are needed for simulation-based decision making while the system still is in operation, e.g., as an input to control algorithms.

Let us consider the abstract time-continuous dynamical system

$$x_t(t) = f(x(t), q, t),\\ x(0) = x_0 \tag{1.1}$$

in state space form for the physical state $x(t)$ in some space X, where the right-hand side f contains (an) unknown parameter(s) q belonging to some space Q. Furthermore, assume that a system output

$$y(t) = h(x(t), q, t) \tag{1.2}$$

with $y(t)$ belonging to some space Y can be observed over time and the corresponding data, possibly affected by measurement errors, are denoted by $z(t) \in Y$. In online identification one aims at inferring and continuously updating the parameter $q \in Q$ simultaneously to the operation of the process modelled by (1.1) and the observation of (1.2). Given an estimate $\hat{q}(t) \in Q$ of q at time t any online update law for the estimate has to satisfy two basic conceptual constraints (see [19]). First, the computation of $\hat{q}(t)$ can only be based on data up to the current time t, i.e., $z(\tau)$ with $\tau \leqslant t$, or in other words, future data $z(\tau)$ with $\tau > t$ cannot be used for calculating the current guess $\hat{q}(t)$. Second, at time t the data $z(\tau)$ with $\tau \leqslant t$ is condensed into $\hat{q}(t)$ and possibly some auxiliary quantity $G(t)$ of fixed dimension and afterwards has to be – with or without loss of relevant information – discarded. Only that way memory space and computation time can be avoided to increase with t. The basic structure of any online algorithm in the time-continuous setting looks like

$$\hat{q}_t(t) = g_1(\hat{q}(t), G(t), z(t)),\\ G_t(t) = g_2(G(t), \hat{q}(t), z(t)).$$

Assuming that a solution q compatible with the data exists, the goal is to construct mappings g_1 and g_2 such that (in case of exact data)

$$\hat{q}(t) \to q \quad \text{as} \quad t \to \infty. \tag{1.3}$$

Online parameter identification, as a central tool for adaptive control of dynamical systems, is mainly investigated in the related literature and also referred to as real-time, recursive, adaptive or sequential identification. So far, available methods backed by a convergence result in the sense of (1.3) are based on one or more of the following assumptions:

- The system (1.1), (1.2) is finite-dimensional.
- The system (1.1), (1.2) is linear time-invariant.
- The system (1.1), (1.2) allows to establish (typically by the use of filtering techniques)

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

the linear relationship

$$y(t) = W(t)q \tag{1.4}$$

between the unknown model parameter $q \in \mathbb{R}^d$ and the model output $y(t) \in \mathbb{R}^m$ via a time-dependent matrix $W(t) \in \mathbb{R}^{m\times d}$.

- The system (1.1), (1.2) shows special structural properties crucial for the construction of (then structure depending) so-called model reference methods.
- The state variable is fully observable, i.e., (1.2) reduces to $y(t) = x(t)$.

For details on the linear and finite-dimensional case we refer to [1, 14, 17, 22, 24]. Model reference techniques relying on the full-state observability can be found in [17] and [20]. In the infinite-dimensional case, so far only a limited number of methods have been suggested for online identification in certain time-dependent PDEs (see [3, 4, 6, 21]). All these techniques require full state observations; in important cases – such as the online identification of heat conduction parameters appearing in the time-dependent heat equation – even observations of spatial derivatives of the data are necessary. In [4], the authors consider 'the elimination of these restrictions' as a 'formidable challenge'.

In this paper, we suggest and analyse an online parameter identification method that works without the assumptions mentioned above. Still it can be applied to general deterministic time-continuous state space models (1.1), (1.2) in Hilbert spaces X, Y and Q of finite and infinite dimensions and also allows for partial state observations. It combines ideas from adaptive control related to the linear and finite-dimensional setup (1.4) and from regularisation theory (see [8]) for offline parameter identification problems. As a consequence, our online method is based on a – only implicitly defined – non-linear parameter-to-output map

$$F(\cdot, t) : \hat{q} \to \hat{y}(t) \tag{1.5}$$

that maps the estimate $\hat{q}$ onto a simulated output $\hat{y}(t) \in Y$ at time t which then is compared to the current data $z(t) \in Y$. Assuming existence of a solution q in case of exact data, i.e., $F(q, t) = z(t)$, for all t, we prove convergence (1.3) based on Lyapunov theory.

The paper is organised as follows: Section 2 briefly reviews basic ideas for online identification in the context of (1.4) and contains preliminary remarks concerning the underlying dynamical system (1.1), (1.2). Section 3 introduces the online method based on (1.5) and analyses its convergence properties for the case of exact data. In practice the data will always be perturbed due to measurement errors and parameter identification schemes (no matter if offline or online) designed for the noise-free case may fail, e.g., lead to parameter drift, if applied to perturbed data due to the ill-posedness of the problem. Hence, Section 4 focuses on techniques in order to counteract negative effects due to data noise. Finally, the method is illustrated in Section 5 by means of online identification of aerodynamic coefficients in a non-linear ODE system describing longitudinal aircraft motions and by means of online parameter identification of a distributed heat conductivity coefficient.

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

We conclude the introductory discussion by pointing out two central concepts in the convergence analysis of online identification methods, common to the methods discussed in [1, 3, 4, 14, 17–20, 22, 24] and also followed in this paper. First, we emphasise that – though online parameter identification algorithms in practice are also used to track time-varying parameters – all available convergence results are based on the assumption that the true model parameter q does not explicitly depend on the time variable. The reason for this assumption is that the time derivative $e_t(t)$ of the parameter error

$$e(t) := \hat{q}(t) - q \tag{1.6}$$

then equals the time-derivative $\hat{q}_t(t)$ of the parameter estimate, which tremendously simplifies theoretical matters (see, for instance, [3, 24]). The use of parameters constant with respect to time in the analysis is often motivated by considering the plant dynamics to be much faster than those of the parameter. Hence, also in this paper we shall assume that the model parameter $q \in Q$ in (1.1), (1.2) to be identified does not depend explicitly on time t. Still, our approach allows for parameter functions q that depend on (parts of) the time-varying state variable, e.g., $q = q(x(t))$, or on space variables. Second, identifiability of unknown model parameters in general can only be expected if the available data contain enough information. In context of online identification such data richness assumptions are typically formulated in terms of so-called persistence of excitation conditions, which, e.g., for (1.4) reads as

$$\exists \gamma, \bar{t} > 0, \ \forall t \in \mathbb{R}^+ \quad \int_t^{t+\bar{t}} W(s)^T W(s)\, ds \geqslant \gamma I \tag{1.7}$$

(see, e.g., [14, 24]). Special variants for the infinite-dimensional case can be found in [3, 5, 21]; the condition used in this paper is introduced in Section 3. Given persistence of excitation, parameter convergence (1.3), or $e(t) \to 0$, is shown by means of Lyapunov theory. The central idea is to define a Lyapunov function $V(e,t)$ such that

$$\begin{aligned} &\kappa_2(\|e\|) \geqslant V(e,t) \geqslant \kappa_1(\|e\|) > 0, \qquad \text{for } e \neq 0, \\ &\frac{d}{dt} V(e(t),t) \leqslant -\kappa_3(\|e(t)\|) < 0, \end{aligned} \tag{1.8}$$

where the κ_i's are class K-functions and (1.8) is considered along the trajectory of the dynamical system for the error $e(t)$. A continuous function $\kappa : \mathbb{R}^+ \to \mathbb{R}+$ is called a class K-function if $\kappa(0) = 0$, $\kappa(s) > 0$ for all $s > 0$ and κ is non-decreasing. See, e.g., [24, 29] for more details on Lyapunov theory in finite- or infinite-dimensional spaces, respectively.

2 Preliminaries

The matrix $W(t)$ in (1.4) serves as a linear parameter-to-output map between finite-dimensional spaces and can formally be understood as a special case of (1.5). Hence, a short review – following [14] and [24] – of online identification based on the setup (1.4) naturally sets up the stage for the method to be presented in Section 3. The linear relationship (1.4) may be derived from applying proper filters to linear time-invariant

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

dynamical systems. For instance, consider the online identification of the parameters a, $b \in \mathbb{R}$ in

$$x_t(t) = -ax(t) + bu(t), \ x(0) = x_0, \tag{2.1}$$

from observations of the system output $y(t) = x(t) \in \mathbb{R}$ with given input $u(t)$. Defining filtered signals $y^f(t)$ and $u^f(t)$ according to

$$y_t^f(f) = -\lambda y^f(t) + y(t) \qquad \text{and} \qquad u_t^f(f) = -\lambda u^f(t) + u(t),$$

with given $\lambda > 0$, one obtains a linear model

$$y(t) = [y^f(t), u^f(t)] \cdot [\lambda - a, b]^T$$

of the form (1.4) with $q = [\lambda - a, b]^T$.

Given – at time t – an estimate $\hat{q}(t)$ of q and (exact) data $z(t) = y(t)$, the *current prediction error* between data $z(t)$ and predicted output $W(t)\hat{q}(t)$ is defined as

$$\|z(t) - W(t)\hat{q}(t)\|^2_{\mathbb{R}^m}. \tag{2.2}$$

As an alternative, the mismatch at time t between data and predicted output can also be expressed in terms of the *total prediction error* at time t with exponential forgetting

$$\int_0^t e^{-\beta(t-s)} \|z(s) - W(s)\hat{q}(t)\|^2_{\mathbb{R}^m} \, ds, \tag{2.3}$$

with $\beta \geqslant 0$. As opposed to (2.2) it also takes all data from the past into account and compares them to outputs made by the current parameter estimate $\hat{q}(t)$, allowing for beneficial averaging effects. For strictly positive β, the exponential term in the integral acts as a data weighting factor. Exponential data forgetting is especially used in practical applications with time-varying parameters since past data are generated by past parameters and hence should be discounted when being used for the estimation of the current parameter. As mentioned earlier, convergence theories for online estimators are only available under the assumption of time-constant parameters to be identified. Even then, the use of exponential data discounting is of advantage (see (2.6) and Section 3). Finally, already with respect to the stability properties of the online method to be derived, a penalty term according to

$$J(\hat{q}(t), t) = \int_0^t e^{-\beta(t-s)} \|z(s) - W(s)\hat{q}(t)\|^2_{\mathbb{R}^m} \, ds + e^{-\beta t} (\hat{q}(t) - q_0)^T G_0^{-1} (\hat{q}(t) - q_0) \tag{2.4}$$

is introduced. The purpose of the second term is to avoid that $\hat{q}$ drifts too far away from q_0 during the initial phase of the identification process where the amount of available data still is limited. For bounded $z(t)$ and $W(t)$, $J(\cdot, t)$ is a convex function over $\mathbb{R}^d$ for all t and any minimum satisfies $\nabla J(\hat{q}(t), t) = 0$, $t \geqslant 0$. This first-order necessary condition is used to derive the non-recursive parameter estimate

$$\hat{q}(t) = G(t) \left[\int_0^t e^{-\beta(t-s)} W^T(s) z(s) \, ds + e^{-\beta t} G_0^{-1} q_0 \right]. \tag{2.5}$$

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

The gain matrix

$$G(t) = \left[\int_0^t e^{-\beta(t-s)} W^T(s) W(s)\, ds + e^{-\beta t} G_0^{-1} \right]^{-1}, \quad G(0) = G_0, \tag{2.6}$$

exists for all t because G_0^{-1} – which is part of (2.6) due to the penalty term in (2.4) – is symmetric positive definite and $W^T(s)W(s)$ is positive semi-definite. Integration and calculation of the inverse in (2.5) can be avoided by exploiting

$$[GG^{-1}]_t = G_t G^{-1} + G G_t^{-1}$$

in order to derive the Riccati-type differential equation

$$G_t(t) = \beta G(t) - G(t) W^T(t) W(t) G(t), \quad G(0) = G_0. \tag{2.7}$$

Then, differentiation of (2.5) with respect to time t yields the recursive update law

$$\hat{q}_t(t) = G(t) W^T(t)(z(t) - W(t)\hat{q}(t)), \quad \hat{q}(0) = q_0. \tag{2.8}$$

The above-mentioned purpose of a strictly positive factor $\beta > 0$ is to ensure that the gain G may not become arbitrarily small or, in other words, $\beta > 0$ avoids that G^{-1} may grow without bound (see (2.6)). In order to also guarantee that G does not grow without bound, (2.7) is modified in [24] to

$$G_t(t) = \beta G(t) - G(t) \left[W^T(t) W(t) + \beta \bar{G}^{-1} \right] G(t), \quad G(0) = G_0, \tag{2.9}$$

where the symmetric positive definite matrix $\bar{G}$ then serves as upper bound for the gain G.

Motivated by (2.8), (2.9), we derive an online parameter identification method in the next section, also applicable to non-linear and infinite-dimensional systems (1.1), (1.2). Instead of the explicitly given matrix $W(t)$ of (1.4) we then have to deal with the non-linear, implicitly defined operator F of (1.5) acting between infinite-dimensional spaces. To this end, we recall that neither existence, uniqueness nor stability of the solution to an inverse problem is guaranteed (see [8]). The focus of this paper is on the design of the method, its analysis and the ill-posedness of the problem due to data error amplification. Hence, we simply assume the existence of a solution q_* for exact data $z(t) \in Y$, i.e., there exists a $q_* \in Q_{\text{ad}}$ such that

$$z(t) = y^{q_*}(t) \tag{2.10}$$

holds.

With respect to the underlying direct problem we make the following assumptions. First, we suppose that the integration of (1.1) and the evaluation of (1.2) with Hilbert spaces X, Y and Q of arbitrary dimension is well posed in the sense that for all parameters q belonging to an admissible set $Q_{\text{ad}} \subseteq Q$ a unique solution x^q of (1.1) with corresponding output y^q exists such that $x^q(t) \in X$ and $y^q(t) \in Y$ hold for all t. In addition, we shall assume that (1.1), (1.2) also admits a unique solution $\hat{x}$ with $\hat{x}(t) \in X$ for all t when

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

integrated with any time-varying parameter $\hat{q}$ that stays in the neighbourhood of the true parameter $q_* \in Q_{\text{ad}}$, i.e.,

$$\hat{q} \in Q_\rho := \{\text{time-varying } q \mid q(t) \in Q \wedge \|q(t) - q_*\| < \rho, \, t \geqslant 0\} \tag{2.11}$$

for some $\rho > 0$. This allows to define for $s \in \mathbb{R}^+$ the non-linear prediction operator

$$F(\cdot, s) : Q_\rho \to Y, \hat{q} \to \hat{y}(s), \tag{2.12}$$

which is evaluated by integration of up to time s

$$\begin{aligned} \hat{x}_t(t) &= f(\hat{x}(t), \hat{q}(t), t), \;\; \hat{x}(0) = x_0, \\ \hat{y}(t) &= h(\hat{x}(t), \hat{q}(t), t). \end{aligned} \tag{2.13}$$

Furthermore, we require that for all $\hat{q} \in Q_\rho$ and all $p \in Q$ the linearised problem

$$\begin{aligned} \hat{v}_t(t) &= f_x(\hat{x}(t), \hat{q}(t), t)\hat{v}(t) + f_q(\hat{x}(t), \hat{q}(t), t)p, \\ \hat{v}(0) &= 0 \end{aligned} \tag{2.14}$$

admits a unique solution $\hat{v}$ with $\hat{v}(t) \in X$ for all t and that

$$\hat{w}(t) = h_x(\hat{x}(t), \hat{q}(t), t)\hat{v}(t) + h_q(\hat{x}(t), \hat{q}(t), t)p$$

is well defined in the sense of $\hat{w}(t) \in Y$. Though these assumptions of course impose certain restrictions on the right-hand side of f in (1.1) (see [7] or [28]), we do not require f to have any special structural properties.

3 Online method and convergence analysis

In this section, we combine ideas from the previous one with the non-linear operator concept from regularisation theory of non-linear inverse problems (see [8]) in order to define and analyse a method for online parameter identification in (1.1), (1.2) with Hilbert spaces X, Y and Q of arbitrary dimension. We consider the total prediction error between the data $z(t)$ and the simulated output $\hat{y}(t)$ and base our method on the minimisation of

$$J(\hat{q}(t), t) = \int_0^t e^{-\beta(t-s)} \|z(s) - F(\hat{q}(t), s)\|_Y^2 \, ds + e^{-\beta t}(\hat{q}(t) - q_0)^T G_0^{-1}(\hat{q}(t) - q_0) \tag{3.1}$$

with data forgetting factor $\beta > 0$. As $W(s)\hat{q}(t)$ in (2.4), $F(\hat{q}(t), s)$ in (3.1) represents the output at time s simulated with a time-constant parameter $q \equiv \hat{q}(t) \in \mathscr{B}_\rho(q_*)$.

Based on the linearised problem (2.14), we define for $\hat{q} \in Q_\rho$ the linear operator

$$F'(\hat{q}, s) : Q \to Y, p \to \hat{w}(s). \tag{3.2}$$

Note that its domain is given only by Q instead of Q_ρ such that only time-constant perturbations $p \in Q$ in (2.14) are considered. This is due to the fact that in order to achieve the online goal (1.3), corrections at time t of past parameter guesses $\hat{q}(\tau)$ with $\tau < t$ are useless and only would mean unnecessary computational burden.

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

We assume that the linear operator $F'(\tilde{q},s)$ is bounded in a neighbourhood of q_*, i.e.,

$$\|F'(\tilde{q},s)\| \leqslant M, \quad \tilde{q} \in Q_\rho, \quad s \in \mathbb{R}^+, \tag{3.3}$$

its Hilbert space adjoint operator will be denoted by

$$F'(\tilde{q},s)^* : Y \to Q.$$

Then, the formal replacement of the matrices $W(t)$ and $W^T(t)$ in (2.8), (2.9) by the operators $F(\hat{q},t)$, $F'(\hat{q},t)$ and $F'(\hat{q},t)^*$ leads to the online parameter identification method

$$\hat{q}_t(t) = G(t)F'(\hat{q},t)^*(z(t) - F(\hat{q},t)), \quad \hat{q}(0) = q_0, \tag{3.4}$$

$$G_t(t) = \beta G(t) - G(t)\left[F'(\hat{q},t)^*F'(\hat{q},t) + \beta\bar{G}^{-1}\right]G(t), \quad G(0) = G_0. \tag{3.5}$$

Here, G_0 and $\bar{G}$ denote linear, self-adjoint positive definite operators on Q with

$$\mu_2\|p\|^2 \geqslant (G_0p,p) \geqslant \mu_1\|p\|^2, \ \eta_2\|p\|^2 \geqslant (\bar{G}p,p) \geqslant \eta_1\|p\|^2, \ (\bar{G}p,p) > (G_0p,p), \quad p \in Q. \tag{3.6}$$

Again, the update law (3.4), (3.5) can be motivated via an Euler equation for solving the first-order necessary condition $\nabla J = 0$ for (3.1) but then ignoring second derivatives of F with respect to the parameter.

The challenge in the analysis of the online method (3.4), (3.5) comes from the fact that the operators $F(\hat{q},t)$ and $F'(\hat{q},t)$ involved are only implicitly defined via the dynamical systems (2.13) and (2.14). Furthermore, the non-linearity of $F(\hat{q},t)$ prevents a straightforward transfer of convergence results from existing theories. At the beginning of our analysis of (3.4), (3.5), we focus on the update law for the gain operator $G(t)$.

Theorem 3.1 *For (arbitrary) $\hat{q} \in Q_\rho$, let (2.13) be well defined and (3.3) hold. If $\beta > 0$, then $G(t) : Q \to Q$ defined by (3.5) with (3.6) is self-adjoint, positive definite and is bounded by $\bar{G}$, i.e.,*

$$g_1(p,p) \leqslant (G(t)p,p) \leqslant (\bar{G}p,p), \quad p \in Q \tag{3.7}$$

and $g_1 = \frac{\mu_1\beta}{\beta+\mu_1M^2}$.

Proof The linear system

$$\begin{pmatrix} G_t^1(t) \\ G_t^2(t) \end{pmatrix} = \begin{pmatrix} 0 & F'(\hat{q},t)^*F'(\hat{q},t) + \beta\bar{G}^{-1} \\ 0 & \beta \end{pmatrix} \begin{pmatrix} G^1(t) \\ G^2(t) \end{pmatrix}, \quad \begin{pmatrix} G^1(0) \\ G^2(0) \end{pmatrix} = \begin{pmatrix} I \\ G_0 \end{pmatrix}$$

has the solution $G^2(t) = G_0e^{\beta t}$ and $G^1(t) = \int_0^t \left[F'(\hat{q},s)^*F'(\hat{q},s) + \beta\bar{G}^{-1}\right]G_0e^{\beta s}\,ds + I$, where I denotes the identity operator. Furthermore, since $F'(\hat{q},s)^*F'(\hat{q},s)$ is self-adjoint and positive semi-definite, $G^1(t)$ is invertible. Then, elementary calculations show that $G(t) = G^2(t)[G^1(t)]^{-1}$ is self-adjoint, solves (3.5) and

$$G^{-1}(t) = (G_0^{-1} - \bar{G}^{-1})e^{-\beta t} + \bar{G}^{-1} + \int_0^t e^{-\beta(t-s)}\left[F'(\hat{q},s)^*F'(\hat{q},s)\right]ds, \tag{3.8}$$

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

which together with (3.6) implies the second inequality in (3.7). In addition, (3.8) gives

$$G^{-1}(t) \leqslant G_0^{-1} + \int_0^t e^{-\beta(t-s)} \left[F'(\hat{q},s)^* F'(\hat{q},s)\right] ds,$$

which together with (3.3) and $\beta > 0$ yields

$$G^{-1}(t) \leqslant G_0^{-1} + \frac{M^2}{\beta} I, \tag{3.9}$$

and hence the first inequality in (3.7). □

While Theorem 3.1 shows that $G^{-1}(t) - \bar{G}^{-1}$ is positive definite for all t, it does not guarantee the existence of a positive lower bound that is independent of t. Existence of such a constant $\tilde{\mu} > 0$ with

$$G^{-1}(t) \geqslant \bar{G}^{-1} + \tilde{\mu} I \tag{3.10}$$

and the assumption

$$\begin{aligned} &([F'(\hat{q},t)^* F'(\hat{q},t) + \beta\tilde{\mu}](\hat{q}(t) - q_*), (\hat{q}(t) - q_*)) \\ &\quad \geqslant 2\left(z(t) - F(\hat{q},t) + F'(\hat{q},t)(\hat{q}(t) - q_*), F'(\hat{q},t)(\hat{q}(t) - q_*)\right), \quad \hat{q} \in Q_\rho,\ t \in \mathbb{R}^+, \end{aligned} \tag{3.11}$$

allow to show that $\hat{q}$ according to (3.4), which exists if

$$G(s)F'(q,s)^*(z(s) - F(q,s)) \text{ is locally Lipschitz,} \tag{3.12}$$

satisfies $\hat{q} \in Q_\rho$ for a sufficiently good initial guess q_0.

Proposition 3.1 *Let the assumptions of Theorem* 3.1, (3.12), (3.10) *and* (3.11) *hold. If the initial guess* q_0 *is sufficiently close to* q_*, *i.e.,*

$$\|q_0 - q_*\| < \sqrt{g_1\left(\frac{1}{\eta_2} + \tilde{\mu}\right)}\rho, \tag{3.13}$$

then $\hat{q}$ *defined by* (3.4) *satisfies* $\hat{q} \in Q_\rho$.

Proof Consider the solution $\hat{q}$ of (3.4) with q_0 according to (3.13), which exists due to (3.12) and the boundedness of G, and assume that it leaves the ball $\mathscr{B}_\rho(q_*)$ at some finite time t_1, i.e.,

$$\|\hat{q}(t_1) - q_*\| = \rho > \|\hat{q} - q_*\| \quad \text{for all } 0 \leqslant t < t_1. \tag{3.14}$$

With $e(t) = \hat{q}(t) - q_*$, one obtains from (3.4), (3.5) the system

$$\begin{aligned} e_t(t) &= G(t)F'(e + q_*, t)^*(z(t) - F(e + q_*, t)), \quad e(0) = q_0 - q_*, \\ G_t(t) &= \beta G(t) - G(t)\left[F'(e + q_*, t)^* F'(e + q_*, t) + \beta\bar{G}^{-1}\right] G(t), \quad G(0) = G_0, \end{aligned} \tag{3.15}$$

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

for the parameter error. Since q_* denotes the true parameter, i.e., $z(t) = y^{q_*}(t)$, 0 is an equilibrium point of (3.15). Next, we define the function

$$V(p,t) = (G^{-1}(t)p, p) \tag{3.16}$$

and get

$$\left(\frac{1}{\eta_2} + \tilde{\mu}\right) \|p\|^2 \leqslant V(p,t) \leqslant \frac{1}{g_1}\|p\|^2,\ p \in Q, \tag{3.17}$$

for $t < t_1$ due to (3.10) and (3.7). Furthermore V satisfies

$$\begin{aligned}\frac{d}{dt}V(e(t),t) &= (G_t^{-1}(t)e(t), e(t)) + 2(G^{-1}(t)e_t(t), e(t))\\ &= -\beta((G^{-1}(t) - \bar{G}^{-1})e(t), e(t)) + (F'(\hat{q},t)^* F'(\hat{q},t)e(t), e(t))\\ &\quad + 2(F'(\hat{q},t)^*(z(t) - F(\hat{q},t)), e(t))\\ &= 2(z(t) - F(\hat{q},t) + F'(\hat{q},t)e(t), F'(\hat{q},t)e(t)) - \beta((G^{-1}(t) - \bar{G}^{-1})e(t), e(t))\\ &\quad - (F'(\hat{q},t)e(t), F'(\hat{q},t)e(t)) \qquad (3.18)\\ &\leqslant 0, \qquad (3.19)\end{aligned}$$

along a trajectory of (3.15) for $t < t_1$ because of (3.11) and (3.10). Finally, (3.13) implies

$$\|e(t)\| \leqslant \sqrt{\frac{1}{\frac{1}{\eta_2} + \tilde{\mu}} V(e(t),t)} \leqslant \sqrt{\frac{1}{\frac{1}{\eta_2} + \tilde{\mu}} V(e(0),0)} \leqslant \sqrt{\frac{1}{(\frac{1}{\eta_2} + \tilde{\mu})g_1}\|e(0)\|^2} < \rho,$$

for $t < t_1$, which for $t \to t_1$ contradicts (3.14). Hence, $\hat{q}(t)$ remains in $\mathscr{B}_\rho(q_*)$ for all t, i.e., $\hat{q} \in Q_\rho$ and V is a Lyapunov function. □

Assumption (3.11) is trivially satisfied in case of $F(\tilde{q},t) = A(t)\tilde{q}(t)$ and $F'(\tilde{q},t) = A(t)$ with a linear operator $A(t)$, then reflecting the infinite-dimensional version of (1.4). Hence, it can be understood as condition that restricts the non-linearity of F locally around the solution q_*. Given the general result (3.8) for the solution of (3.5), (3.10) is a condition on the operator $F'(\hat{q},s)^*F'(\hat{q},s)$ or the excitation of the underlying dynamical system. It certainly is satisfied if persistence of excitation is given in the sense of

$$\exists \gamma, \bar{t} > 0,\ \forall t \in R^+ \int_t^{t+\bar{t}} F'(\tilde{q},s)^* F'(\tilde{q},s)\, ds \geqslant \gamma I, \quad \tilde{q} \in Q_\rho, \tag{3.20}$$

since then, (3.8) implies (3.10) with $\tilde{\mu} = \min\{(\frac{1}{\mu_2} - \frac{1}{\eta_1})e^{-\beta\bar{t}}, \gamma e^{-\beta\bar{t}}\}$. Condition can be understood as an extension of (1.7) but is different from those used in [3], [5] or [21].

Online parameter identification methods typically are based on fitting the simulated output to the measured data, e.g., by minimisation of (3.1), while the real goal is to reduce the error in the parameter space. However, convergence in the output space is in general no guarantee for convergence of the parameter estimate $\hat{q}$ towards the true value (see, e.g., [14] and Section 5 for a numerical example). The following theorem states that the desired parameter convergence can be obtained for (3.4), (3.5) under persistence of

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

excitation (3.20) and a slightly stronger version of assumption (3.11), namely

$$([F'(\hat{q},t)^*F'(\hat{q},t)+\beta\tilde{\mu}](\hat{q}(t)-q_*),(\hat{q}(t)-q_*)) \\ \geqslant 2\left(z(t)-F(\hat{q},t)+F'(\hat{q},t)(\hat{q}(t)-q_*),F'(\hat{q},t)(\hat{q}(t)-q_*)\right)+\kappa(\|\hat{q}(t)-q_*\|), \quad (3.21)$$

for $\hat{q}\in Q_\rho$, $t\in\mathbb{R}^+$ with κ denoting a class-K function. Again, (3.21) (with positive $\tilde{\mu}$) is satisfied in the linear case with $\kappa(\|q\|)=\beta\tilde{\mu}\|q\|^2$.

Theorem 3.2 *Let the assumptions of Proposition* 3.1, (3.20) *and* (3.21) *hold. Then,*

$$\hat{q}(t)\to q_* \qquad \text{as } t\to\infty. \quad (3.22)$$

Proof Following the steps of the proof of Proposition 3.1, (3.20) and (3.21) allow to establish

$$\frac{d}{dt}V(e(t),t)\leqslant-\kappa(\|e(t)\|). \quad (3.23)$$

Since 0 is an equilibrium point of (3.15), Lyapunov theory and (3.17) give $e(t)\to 0$ for $t\to\infty$ (see [29]). □

We emphasise that a positive forgetting factor $\beta>0$ is essential for the convergence analysis presented since only then the positive definiteness of G for all t can be guaranteed (see (3.9)) (which gives the upper bound on V in (3.17)). Still, the choice $\beta=0$ is of special interest since in case of finite dimensions the method (3.4), (3.5) then can be understood as an extended Kalman Bucy filter (see [10]).

Formal similarities of the online method (3.4) are also given with so-called Kaczmarz regularisation techniques for solving a system of N non-linear operator equations (see, e.g., [11]) where the basic idea is to cyclically consider each of the N equations separately. Hence, Kaczmarz methods – as discussed and analysed in the literature – cannot be directly applied to online parameter identification; first, one would have to set $N=\infty$ (with N then considered as discrete time) and abandon the cyclic idea. Motivated by the convergence analysis in [11] one then might try to establish (3.22) under the so-called tangential cone condition

$$\|F(q_*,t)-F(\hat{q},t)+F'(\hat{q},t)(\hat{q}(t)-q_*)\|\leqslant\eta\|F(q_*,t)-F(\hat{q},t)\|, \quad \eta<1/2,\ \hat{q}\in Q_\rho,\ t\geqslant 0.$$

However, this would imply that the data $z(t)$ and the predicted output $\hat{y}(t)$ would have to coincide for all t with $\hat{q}(t)=q_*$ which of course is not realistic (just think of $\hat{q}(t)=q_*\sin(t)$ and $t=(2n+1)\pi/2$, $n=0,1,2,\ldots$). Such implications are not given by (3.11), (3.21). Finally, we mention the idea of asymptotic regularisation for non-linear operator equations, [2, 12, 13, 27]. If applied to parameter identification in dynamical system, a sequence of auxiliary problems has to be integrated over an artificial time interval, hence also these techniques do not allow for online identification in a straightforward manner.

The disadvantage of the persistence of excitation condition (3.20) – and that of its counterparts in the literature – is that they hardly can be verified, especially not during the online computations. Hence, in practical situations studies on how to sufficiently excite

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

the underlying system (1.1) by choosing proper input functions f (see [5] and [24]) have to be conducted in advance before the actual online estimator is started. For ideas for at least getting a feeling about (3.20) during the online computations we refer to Section 5.1.

A direct realisation of the adjoint operator $F'(\hat{q},t)^*$ in (3.4) would require to solve an adjoint state equation backwards in time which is not consistent with the idea of online computation. Hence, (3.4), (3.5) has to be understood as short form of

$$(\hat{q}_t(t),p) = (z(t) - F(\hat{q},t), F'(\hat{q},t)G(t)p), \tag{3.24}$$

$$(G_t(t)p,\tilde{p}) = \beta(G(t)p,\tilde{p}) - (F'(\hat{q},t)G(t)p, F'(\hat{q},t)G(t)\tilde{p}) - \beta(\bar{G}^{-1}G(t)p, G(t)\tilde{p}), \tag{3.25}$$

for $p,\tilde{p} \in Q$. Note that the linearised system (2.14) has to be solved only once, namely simultaneous to the integration of (3.24), (3.25) (see Section 5).

4 Online identification in presence of noisy data

In this section, we turn to the case of noisy data $z^\delta(t) \in Y$ and suppose that a bound δ with

$$\|z^\delta(t) - z(t)\| \leqslant \delta,\ t \geqslant 0, \tag{4.1}$$

for the error between noisy and exact data is available. While we still assume that the exact data are attained by the parameter q_*, existence of a solution in case of noisy data is not assumed and not even an issue, since for $\delta > 0$ only a regularised approximation of q_* is searched for. Due to the ill-posedness of parameter identification, online methods designed for exact data may fail, e.g., the parameter may diverge, if applied to noisy data even if the perturbations are bounded (see [14] and Section 5 for a numerical illustration). Though (3.1) reminds of Tikhonov regularisation (see [8]), it is not sufficient to counteract instabilities due to data noise since there the penalty term is exponentially discounted – hence, additional measures become necessary. Of course, in presence of data noise parameter convergence in the sense of (3.22) can no longer be expected. Instead, the best one can ask for is that the parameter guess $\hat{q}$ converges towards and subsequently stays within a neighbourhood of the solution q_*, whose size is in the range of the noise level, at least asymptotically.

Based on ideas presented in [6], [14] and [26] we first consider a so-called leakage approach

$$\hat{q}_t(t) = G(t)F'(\hat{q},t)^*(z^\delta(t) - F(\hat{q},t)) - \sigma(t)G(t)(\hat{q}(t) - q_0), \quad \hat{q}(0) = q_0, \tag{4.2}$$

$$G_t(t) = \beta G(t) - G(t)\left[F'(\hat{q},t)^*F'(\hat{q},t) + \beta\bar{G}^{-1}\right]G(t), \quad G(0) = G_0. \tag{4.3}$$

The purpose of the scalar, positive design variable $\sigma(t)$ is to ensure that for $V \geqslant \underline{V} > 0$, with V as in (3.16) and some $\underline{V}$ that may depend on δ, the time derivative of V along the trajectory of the parameter error $e(t) = \hat{q} - q_*$ for (4.2) is negative. Then, due to Lyapunov theory, $\|e\|$ decreases exponentially until $\hat{q}$ reaches a neighbourhood of q_* that depends on the noise level δ. That way parameter divergence due to data noise can be avoided by use of a leakage term.

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

Theorem 4.1 *Let the assumptions of Proposition* 3.1, (4.1) *hold and choose*

$$\sigma(t) \equiv \sigma \geqslant M^2 + c \tag{4.4}$$

for some $c > 0$. *If the initial guess* q_0 *satisfies*

$$\frac{1}{\frac{1}{\eta_2} + \tilde{\mu}} \left(\frac{1}{g_1} \left(1 + \frac{\sigma}{c}\right) \|q_0 - q_*\|^2 + \frac{\delta^2}{cg_1} \right) < \rho^2,$$

then $\hat{q}$ *according to* (4.2) *satisfies* $\hat{q} \in Q_\rho$ *and the norm of the parameter error* $e(t) = \hat{q}(t) - q_*$ *decreases exponentially until* $\hat{q}(t)$ *reaches the set*

$$\left\{ q \in Q \mid \|q - q_*\|^2 \leqslant \frac{\delta^2 + \sigma\|q_* - q_0\|^2}{cg_1(\frac{1}{\eta_2} + \tilde{\mu})} \right\}. \tag{4.5}$$

Proof Assume that $\hat{q}$ leaves $\mathscr{B}_\rho(q_*)$ at some finite time t_1, i.e.,

$$\|\hat{q}(t_1) - q_*\| = \rho > \|\hat{q} - q_*\| \quad \text{for all } 0 \leqslant t < t_1. \tag{4.6}$$

With $e(t) = \hat{q}(t) - q_*$ and $V(p,t) = (G^{-1}(t)p, p)$, (4.2), (4.3) yields

$$\begin{aligned} \frac{dV}{dt}(e(t),t) &= (G_t^{-1}(t)e(t), e(t)) + 2(G^{-1}(t)e_t(t), e(t)) \\ &= -\beta((G^{-1}(t) - \bar{G}^{-1})e(t), e(t)) + (F'(\hat{q},t)^* F'(\hat{q},t)e(t), e(t)) \\ &\quad + 2(F'(\hat{q},t)^*(z(t) - F(\hat{q},t)), e(t)) \\ &\quad + 2(F'(\hat{q},t)^*(z^\delta(t) - z(t)), e(t)) - 2\sigma(e(t) + q_* - q_0, e(t)) \\ &\leqslant 2M\delta\|e(t)\| - 2\sigma(e(t) + q_* - q_0, e(t)). \end{aligned} \tag{4.7}$$

Since

$$\begin{aligned} -\sigma(e(t) + q_* - q_0, e(t)) &= -\sigma\|e(t)\|^2 + \sigma\|q_* - q_0\|\|e(t)\| \\ &\leqslant -\frac{\sigma}{2}\|e(t)\|^2 + \frac{\sigma}{2}\|q_* - q_0\|^2, \end{aligned} \tag{4.8}$$

(4.4) and (3.17) lead to

$$\begin{aligned} \frac{dV}{dt}(e(t),t) &\leqslant 2M\delta\|e(t)\| - \sigma\|e(t)\|^2 + \sigma\|q_* - q_0\|^2 \\ &\leqslant -c\|e(t)\|^2 + \delta^2 + \sigma\|q_* - q_0\|^2 \\ &\leqslant -cg_1 V(e(t),t) + \delta^2 + \sigma\|q_* - q_0\|^2, \end{aligned} \tag{4.9}$$

for $t < t_1$. This allows to show

$$V(e(t),t) \leqslant e^{-cg_1 t} V(e(0),0) + \frac{\delta^2 + \sigma\|q_* - q_0\|^2}{cg_1} \tag{4.10}$$

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

and

$$\left(\frac{1}{\eta_2}+\tilde{\mu}\right)\|e(t)\|^2 \leqslant \frac{1}{g_1}\|e(0)\|^2 + \frac{\delta^2+\sigma\|e(0)\|^2}{cg_1} < \left(\frac{1}{\eta_2}+\tilde{\mu}\right)\rho^2,\ t < t_1,$$

which contradicts (4.6) for $t \to t_1$. Hence, the parameter error satisfies $\|e(t)\| < \rho$ for all t and decreases exponentially until $\hat{q}$ reaches the set (4.5) due to (4.10). □

Recall that the scalar ρ introduced in (2.11) represents the size of the parameter domain on which the operator $F(\cdot, t)$ can be defined. For large ρ the statement $\hat{q} \in \mathcal{Q}_\rho$, extreme case $\rho = \infty$, of Theorem 4.1 does not yield a satisfactory bound for the parameter error. The actual stability result comes with 'convergence' of $\hat{q}$ towards the ball (4.5) around q_* whose radius can be much smaller than ρ, especially in case of strong persistence of excitation, i.e., large $\tilde{\mu}$.

The disadvantage of a constant leakage term as in (4.4) is that stability of (4.2) is achieved at the expense of destroying – even under assumption (3.21) – the ideal property (3.22) in case of exact data, i.e., $\delta = 0$. Transferring parameter choice rules from regularisation of general non-linear operators in dependence of the noise level and the data (see [8]) would suggest to choose

$$\sigma = \sigma(\delta) \quad \text{with} \quad \sigma(\delta) \to 0 \quad \text{as } \delta \to 0, \tag{4.11}$$

such that for exact data, i.e., $\delta = 0$ in (4.1), the original algorithm (3.4) is obtained. However, this would require a different proof technique due to the incompatibility of (4.4) and (4.11).

An alternative to a time-constant σ is the switching strategy

$$\sigma(t) = \begin{cases} 0 & \text{if } \|\hat{q}(t) - q_0\| < \xi \\ \bar{\sigma}\left(\frac{\|\hat{q}(t)-q_0\|}{\xi} - 1\right)^n & \text{if } \|\hat{q}(t) - q_0\| \in [\xi, 2\xi] \\ \bar{\sigma} & \text{else} \end{cases} \tag{4.12}$$

with $\bar{\sigma} > 0$ (see [14]), which requires the knowledge of an upper bound

$$\|q_* - q_0\| \leqslant \xi. \tag{4.13}$$

Here, the leakage term is only activated when $\|\hat{q}(t) - q_0\|$ exceeds the bound ξ. Since the actual overshoot over ξ is enforced to stay small, this approach is also referred to as soft projection. The choice (4.12) yields

$$\begin{aligned} \sigma(t)(e(t) + q_* - q_0, e(t)) &= \sigma(t)(\|\hat{q}(t) - q_0\|^2 - (\hat{q}(t) - q_0, q_* - q_0) \\ &\geqslant \sigma(t)\|\hat{q}(t) - q_0\|(\|\hat{q}(t) - q_0\| - \xi + \xi - \|q_* - q_0\|) \\ &\geqslant 0 \end{aligned}$$

because of $\sigma(t)(\|\hat{q}(t) - q_0\| - \xi) \geqslant 0$ and (4.13). Therefore, in contrast to a fixed σ, σ chosen according to (4.12) can only make $\frac{d}{dt}V(e(t), t)$ more negative (see (4.7)). Especially,

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

under the assumptions of Theorem 3.2, one gets

$$\frac{d}{dt}V(e(t),t) \leqslant -\kappa(\|e(t)\|) + 2M\delta\|e(t)\| - 2\sigma(t)(e(t)+q_* - q_0, e(t)),$$

such that in case of exact data, i.e., $\delta = 0$, parameter convergence (3.22) still holds for (4.2), (4.12). As in Theorem 4.1, the choice $\bar{\sigma} \geqslant M + c$ in (4.12) allows to show convergence of e in case of $\delta \neq 0$ to a residual set slightly larger than (4.5) (see [6] and [14]). If (4.13) does not hold, then the switching technique has similar properties as the approach with fixed σ (see [14]).

Finally, we mention the option of so-called dead zone approaches (see [14], [24] and [26]). The basic idea is to pause the integration of the online estimator whenever the output error $\|z^\delta(t) - \hat{y}(t)\|$ is dominated by the noise level δ, i.e., to consider

$$\hat{q}_t(t) = \begin{cases} G(t)F'(\hat{q},t)^*(z^\delta(t) - F(\hat{q},t)) & \text{if } \|z^\delta(t) - F(\hat{q},t)\| \geqslant \tau\delta \\ 0 & \text{else} \end{cases} \tag{4.14}$$

together with (3.5) for some $\tau > 1$. While some theoretical studies of dead-zone approaches for the linear, finite-dimensional case can be found in [14] and [26] (not including the dead-zone analysis of (2.8), (2.9) for $m > 1$) and a similar regularisation strategy is followed in [11] for the cyclic Kaczmarz iteration, a convergence analysis of (4.14) poses additional challenges and will be subject of future investigations.

5 Numerical experiments

This section opens with a simple example in order to illustrate our online parameter identification method (3.4), (3.5) as well as the crucial theoretical assumptions made in Section 3. It will be followed by the identification of aerodynamic coefficients in a longitudinal 3DoF aircraft model before turning the identification of a distributed heat conduction coefficient from full and partial temperature observations.

5.1 Simple example

As many texts on online identification in linear and finite-dimensional dynamical systems, we choose the identification of $q = [a,b]^T \in \mathbb{R}^2$ in (2.1) from full state measurements, i.e., $h(x,q,t) = x$, for a basic illustration of our method (3.4), (3.5). For an estimator $\hat{q}$ with $\hat{q}(t) = [\hat{q}_1(t), \hat{q}_2(t)]^T$ the predicted state $\hat{x}(t)$ satisfies

$$\hat{x}_t(t) = -\hat{q}_1(t)\hat{x}(t) + \hat{q}_2(t)u(t), \quad \hat{x}(0) = x_0,$$

which allows to define the non-linear prediction operator $F(\cdot,t) : \hat{q} \to \hat{x}(t)$. The linearised problem (2.14) reads as

$$\hat{v}_t(t) = -\hat{q}_1(t)\hat{v}(t) + \left(\begin{pmatrix} -\hat{x}(t) \\ u(t) \end{pmatrix}, \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \right), \quad \hat{v}(0) = 0,$$

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

with $p = [p_1, p_2]^T \in \mathbb{R}^2$, which defines the linear operator $F'(\hat{q}, t) : \mathbb{R}^2 \to \mathbb{R}$, $p \to \hat{v}(t)$. Especially, for $p = p_1 e_1 + p_2 e_2$ (with e_i denoting the ith unit vector) we have

$$F'(\hat{q}, t)p = p_1 \hat{v}_1(t) + p_2 \hat{v}_2(t), \quad \text{for all } p = [p_1, p_2]^T \in \mathbb{R}^2, \tag{5.1}$$

with $\hat{v}_1(t) = F'(\hat{q}, t)e_1$ and $\hat{v}_2(t) = F'(\hat{q}, t)e_2$ given by

$$\begin{aligned} \hat{v}_{1_t}(t) &= -\hat{q}_1(t)\hat{v}_1(t) - \hat{x}(t), \quad \hat{v}_1(0) = 0, \\ \hat{v}_{2_t}(t) &= -\hat{q}_1(t)\hat{v}_2(t) + u(t), \quad \hat{v}_2(0) = 0. \end{aligned} \tag{5.2}$$

Next, taking the inner product in $\mathbb{R}^2$ with $p \in \mathbb{R}^2$ at both sides of (3.4), compared to (3.24), leads to

$$\begin{aligned} (\hat{q}_t(t), p) &= (z(t) - F(\hat{q}, t)) \cdot F'(\hat{q}, t)G(t)p \\ &= (z(t) - \hat{x}(t)) \cdot F'(\hat{q}, t) \begin{pmatrix} g_{11}(t)p_1 + g_{12}(t)p_2 \\ g_{21}(t)p_1 + g_{22}(t)p_2 \end{pmatrix} \\ &= (z(t) - \hat{x}(t)) \cdot ((g_{11}(t)p_1 + g_{12}(t)p_2)\hat{v}_1(t) + (g_{21}(t)p_1 + g_{22}(t)p_2)\hat{v}_2(t)) \\ &= \left(\left(G \begin{pmatrix} \hat{v}_1(t) \\ \hat{v}_2(t) \end{pmatrix} \right) (z(t) - \hat{x}(t)), p \right) \end{aligned}$$

because of (5.1) and $G(t)p \in \mathbb{R}^2$. Here, $g_{12}(t) = g_{21}(t)$ due to the symmetry of $G(t) : \mathbb{R}^2 \to \mathbb{R}^2$. Especially, the choices $p = e_1$ and $p = e_2$, respectively, give

$$\begin{aligned} \hat{q}_{1_t}(t) &= (z(t) - \hat{x}(t)) \cdot (g_{11}(t)\hat{v}_1(t) + g_{21}(t)\hat{v}_2(t)), \quad \hat{q}_1(0) = q_1^0, \\ \hat{q}_{2_t}(t) &= (z(t) - \hat{x}(t)) \cdot (g_{12}(t)\hat{v}_1(t) + g_{22}(t)\hat{v}_2(t)), \quad \hat{q}_2(0) = q_2^0. \end{aligned} \tag{5.3}$$

Finally, (3.5), (5.1) and $g_{ij}(t) = \langle G(t)e_i, e_j \rangle$, $i, j \in \{1, 2\}$, yield

$$G_t(t) = \beta G(t) - G(t) \left[\begin{pmatrix} \hat{v}_1^2(t) & \hat{v}_1(t)\hat{v}_2(t) \\ \hat{v}_2(t)\hat{v}_1(t) & \hat{v}_2^2(t) \end{pmatrix} + \beta \bar{G}^{-1} \right] G(t), \quad G(0) = G_0.$$

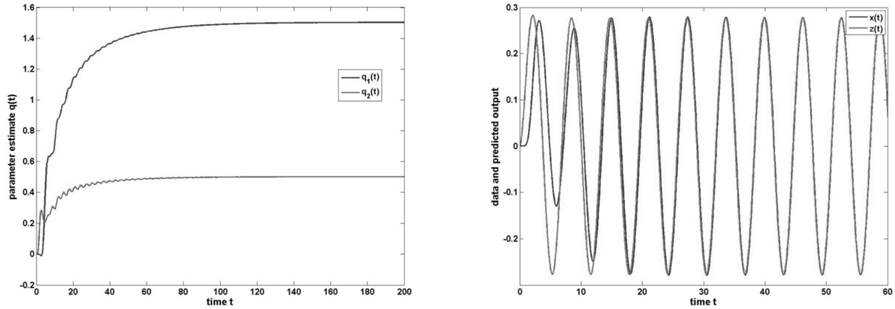
We applied (5.2), (5.3), (5.1) to the online identification of the true parameter $q_* = [1.5, 0.5]^T$ in (2.1), with $x_0 = 0$, input $u(t) = \sin(t)$ and initial guess $q_0 = [0, 0]^T$, from exact observations, i.e., $z(t) = x(t)$. The results obtained with the setting

$$G_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \bar{G} = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}, \ \beta = 0.5,$$

are illustrated in Figures 1–3 (similar results have also been obtained for other settings). Figure 1 shows that the data $z(t)$ are immediately tracked by the predicted output $\hat{x}(t)$ and that the parameter estimate $\hat{q}$ converges towards the true value q_*. The latter suggests that the input $u(t) = \sin(t)$ persistently excites the system (2.1) and hence provides sufficient information for the online parameter identification. Figure 2 supports this conjecture; it indicates that

$$\lambda_{\min} \left\{ \int_0^t \begin{pmatrix} \hat{v}_1(s)^2 & \hat{v}_1(s)\hat{v}_2(s) \\ \hat{v}_1(s)\hat{v}_2(s) & \hat{v}_2^2(s) \end{pmatrix} ds \right\} \to \infty \quad \text{as } t \to \infty,$$

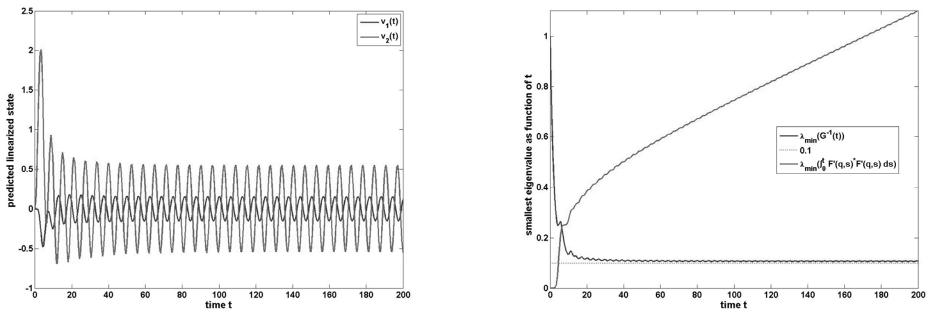
https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press



(a) Parameter estimate $\hat{q}(t)$ converges towards q_*

(b) Predicted output $\hat{x}(t)$ tracks the data $z(t)$

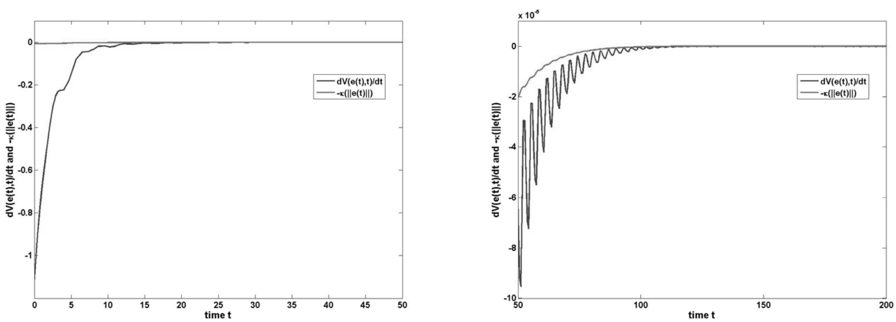
FIGURE 1. Convergence both in output and parameter space.



(a) Linearized state variables $\hat{v}_1(t)$ and $\hat{v}_2(t)$

(b) $\lambda_{\min}(G^{-1}(t))$ and $\lambda_{\min}(\int_0^t F'(\hat{q},s)^* F'(\hat{q},s)\,ds)$

FIGURE 2. Indication of persistence of excitation.

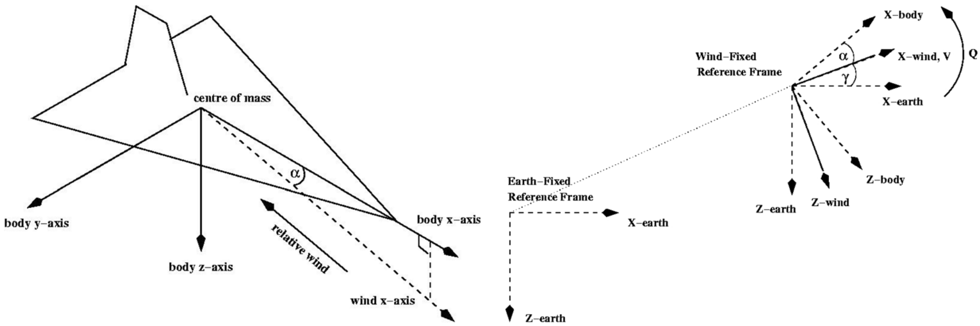


(a) $\frac{dV}{dt}(e(t)t)$ and $-\kappa(\|e(t)\|) = -\beta\tilde{\mu}\|e(t)\|^2$

(b) as (a), only different axes

FIGURE 3. Negative time derivative of the Lyapunov function V.

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press



(a) Definition of axes and angle of attack α (b) Wind axis coordinate system

FIGURE 4. Definition of axes for 3DoF aircraft model.

where $\lambda_{\min}(B)$ denotes the smallest eigenvalue of B, which is a necessary condition for (3.20) to hold. Note, that even after the estimate $\hat{q}$ has converged towards q_*, the linearised states $\hat{v}_1$ and $\hat{v}_2$ stay excited. (A zoom into) Figure 2(b) also suggests that condition (3.10) on $G^{-1}(t)$ with $\bar{G}^{-1} = 0.1I$ is satisfied by $\tilde{\mu} = 0.005$ (also see (3.8)). This allows to compare in Figure 3 the time derivative of the Lyapunov function $V(e(t), t)$ given by (3.18) to the negative of the class-K function $\kappa(\|e\|) = \beta\tilde{\mu}\|e\|^2$, which indicates that conditions (3.21) and (3.23), crucial for proving parameter convergence (3.22), hold.

As mentioned in Section 3, the persistence of excitation condition (3.20) typically is hard to verify, especially during the online identification. Hence, monitoring $F'(\hat{q}, t)^* F'(\hat{q}, t)$ as in Figure 2(a), which has to be computed anyway, might serve as a practicable tool for getting a clue if sufficient excitation is given. In addition, the minimal eigenvalues of $G^{-1}(t)$ and $\int_0^t (F'(\hat{q}, s) * F'(\hat{q}, s)\, ds)$ could be observed as in Figure 2(b); however, such calculations pose additional computational tasks to be carried out online.

Currently, we are aware of at least five different methods that can be used for online parameter identification in (2.1) and, in general, in linear time-invariant systems with full state observations. A comparison will be the subject of future studies. In the next subsection, we apply (3.4), (3.5) to online parameter identification in a non-linear ODE-system with partial state observations. By means of this example, we then also discuss the influence of measurement errors and of a loss of persistent excitation on our method.

5.2 Longitudinal aircraft dynamics

Models for the dynamics of an aircraft are based on a combination of aerodynamic force and moment models with vector equations of motion. Choosing a wind axis coordinate system that has its origin in the rigid aircraft, the x-axis points in the direction of the aircraft velocity V (as opposed to a body fixed system) (see Figure 4(a)). The three translational degrees of freedom then are defined by moving along the x-, y- and z-axis, while the three rotational degrees of freedom are given by the bank angle ϕ around the x-axis, the flight path angle γ around the y-axis and the heading angle χ around the z-axis. If the aircraft is flying under wings-level conditions, i.e., $\phi = 0$, the equations for the lateral-directional motion can be decoupled from those for the longitudinal motion

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

with the latter reading as

$$\begin{aligned} V_t(t) &= \frac{F_T}{m}\cos\alpha - \frac{D}{m} - g\sin(\theta-\alpha)\\ \alpha_t(t) &= -\frac{F_T}{mV}\sin\alpha - \frac{L}{mV} + \frac{g}{V}\cos(\theta-\alpha) + Q\\ \theta_t(t) &= Q\\ Q_t(t) &= M/J_y\\ H_t(t) &= V\sin(\theta-\alpha) \end{aligned} \tag{5.4}$$

with $\gamma = \theta - \alpha$ (see Figure 4(b)). The state vector for these equations is

$$x = [V, \alpha, \theta, Q, H]^T \in \mathbb{R}^5$$

with aircraft velocity V, angle of attack α, pitch angle θ, pitch rate Q and altitude H (see [25] for details). F_T denotes the thrust along the body x-axis, m is the mass of the aircraft, g is the gravitational constant and J_y is the inertia around the y-axis. The typical ansatz for drag D, lift L and pitching moment M is

$$D, L, M \sim V^2 S C_{D,L,M} \tag{5.5}$$

with wing area S and aerodynamic coefficients $C_{D,L,M}$ that may depend on aerodynamic angles, temperature and density of the atmosphere, altitude, thrust coefficients, Mach and Reynolds numbers and control parameters. These aerodynamic coefficients are typically determined through wind tunnel or flight-test data. For a survey about offline methods used in flight vehicle design we refer to [15]. In the context of flight control design, i.e., design of systems that aid or replace the human pilot, modern techniques such as dynamic inversion (see [25]) use the underlying non-linear state-space aircraft model and hence require knowledge of the aerodynamic coefficients during the flight. In order to provide ongoing updates of these coefficients, online parameter identification is necessary.

For testing our online identification method (3.4), (3.5) we considered the model (5.4) with a setup taken from [25] that corresponds to a medium-size transport aircraft and an aerodynamic coefficient ansatz

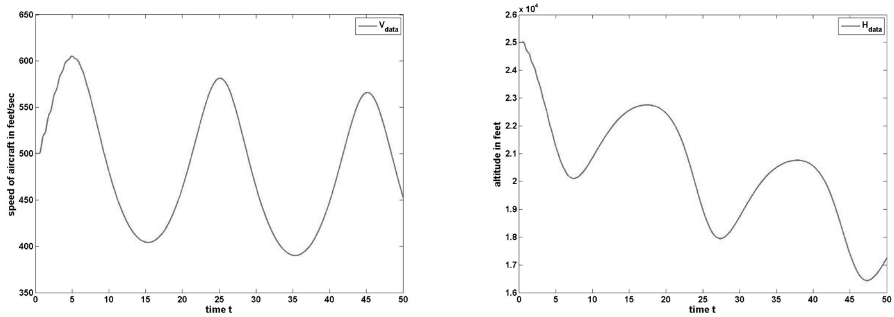
$$\begin{aligned} C_L &= 0.2\cdot q_2 + 0.1\cdot q_1\alpha\\ C_D &= 0.01\cdot q_3 + 0.042\cdot C_L^2\\ C_M &= 0.05 - 0.022\cdot\alpha - 0.016\cdot\delta_e, \end{aligned} \tag{5.6}$$

where δ_e denotes the elevator deflection. The true parameter vector is $q_* = [0.85, 1.0, 1.6]^T \in \mathbb{R}^3$, the initial conditions $x(0)$ represent steady-state flight conditions at $V_0 = 500$ ft/s and at altitude $H_0 = 25{,}000$ ft. Then, choosing a periodic system input

$$\delta_e(t) = -2.3822 + 3.5\sin(0.1\pi t), \tag{5.7}$$

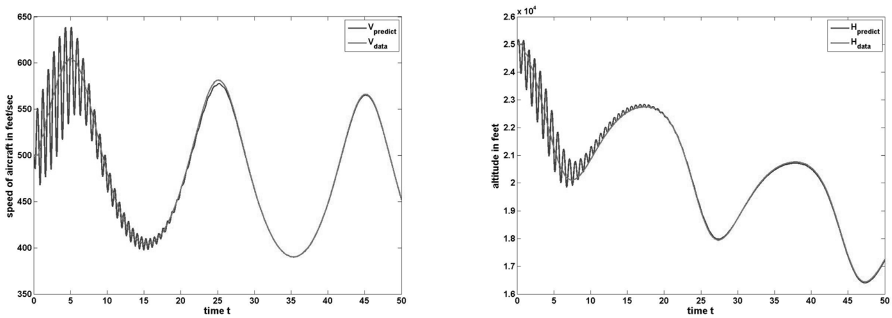
the aircraft falls and climbs with corresponding increase and decrease in velocity as shown

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press



(a) Exact aircraft velocity data $z_V(t)$ (b) Exact altitude data $z_H(t)$

FIGURE 5. Exact flight data.



(a) Predicted output $\hat{V}(t)$ and data $z_V(t)$ (b) Predicted output $\hat{H}(t)$ and data $z_H(t)$

FIGURE 6. Tracking of flight data.

in Figure 5. With altitude and speed measurements taken during the flight, i.e.,

$$h(x, q, t) = [V(t), H(t)], \tag{5.8}$$

Figure 5 also represents the exact data $z(t) = [z_V(t), z_H(t)]$ for the online identification of the parameter q_*. For the initial guess $\hat{q}(0) = [1.0625, 0.75, 2.0]^T$, which corresponds to a 25% deviation from q_* in each component, the results obtained by the online method (3.4), (3.5) with the setting $\beta = 0.001$, $G(0) = 0.05I$ and $\bar{G} = 0.1I$ are shown in Figures 6 and 7. After an initial oscillatory phase, the predicted velocity $\hat{V}$ and altitude $\hat{H}$ track the exact data z_V and z_H; furthermore, the parameter estimate $\hat{q}$ converges to q_*. Hence, the input (5.7) sufficiently excites the system (5.4) for online determination of q_* based on the output (5.8). Actually, the excitation even seems strong enough to yield a bounded parameter error in presence of data perturbations. Figure 8 shows noisy data with

$$\frac{|z_V^\delta(t) - z_V(t)|}{|z_V(t)|} \leqslant 0.0142 \quad \text{and} \quad \frac{|z_H^\delta(t) - z_H(t)|}{|z_H(t)|} \leqslant 0.025,$$

while Figures 9 and 10 illustrate that data tracking and bounded parameter errors (the bounds also hold on a longer time horizon) are obtained by application of (3.4), (3.5)

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

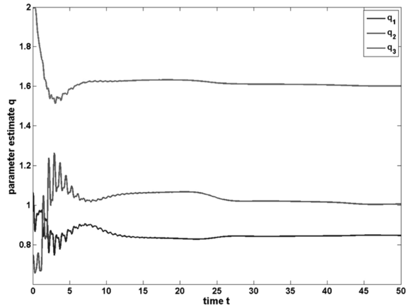
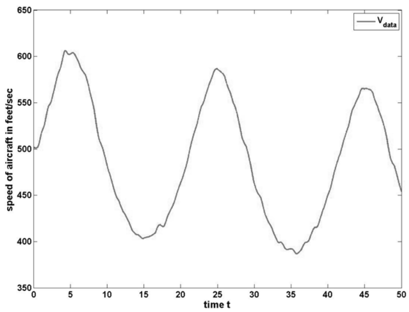
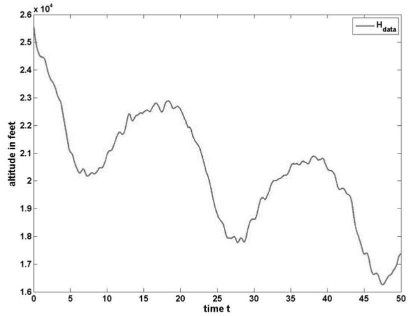


FIGURE 7. Convergence of $\hat{q}$ towards true values.

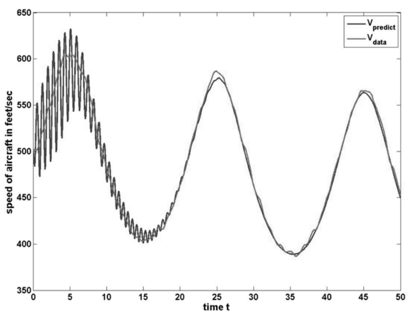


(a) Noisy velocity data $z_V^\delta(t)$

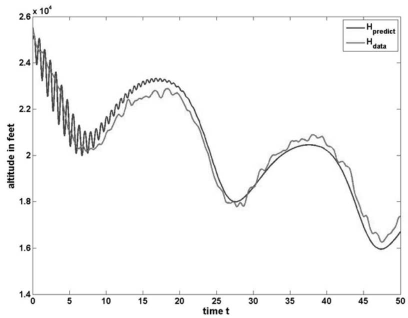


(b) Noisy altitude data $z_H^\delta(t)$

FIGURE 8. Noisy flight data.



(a) Predicted velocity $\hat{V}(t)$ and noisy data $z_V^\delta(t)$



(b) Predicted altitude $\hat{H}(t)$ and noisy data $z_H^\delta(t)$

FIGURE 9. Tracking of noisy data.

without additional measures such as presented in Section 4. This numerical observation would be in line with discussions led in [24], which state that persistence of excitation may guarantee bounded parameter errors in case of bounded data noise.

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

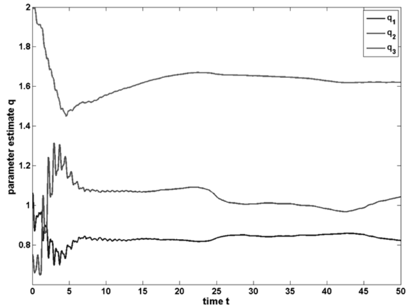
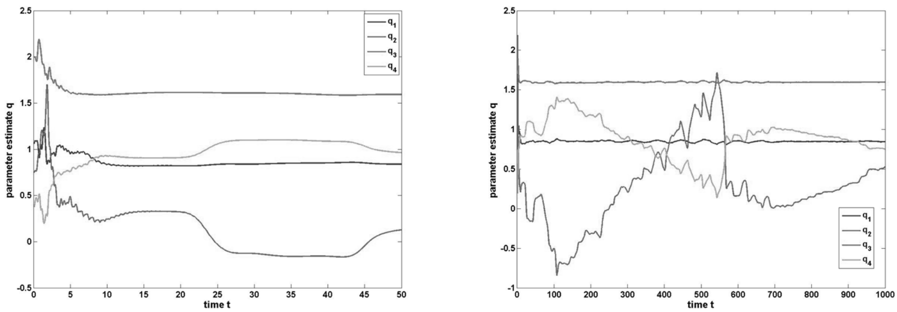


FIGURE 10. Bounded parameter error in presence of data noise due to persistence of excitation.



(a) Parameter estimate $\hat{q}$ (b) Longer time horizon

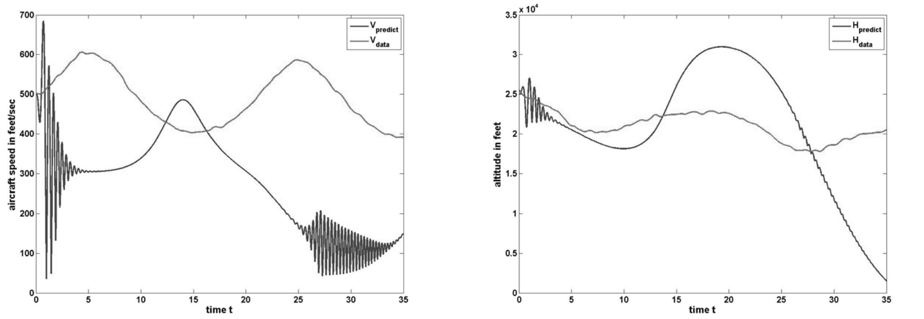
FIGURE 11. Loss of parameter convergence towards q_* due to lack of persistence of excitation.

The situation changes significantly, if we replace the moment coefficient (5.6) by

$$C_M = 0.1 \cdot q_4 - 0.022 \cdot \alpha - 0.016 \cdot \delta_e$$

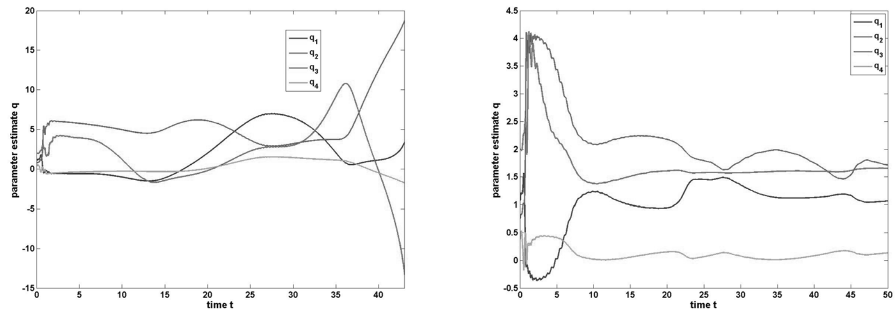
with δ_e as in (5.7) and now address the identification of $q^* = [0.85, 1.0, 1.6, 0.5]^T \in \mathbb{R}^4$ based on the same data as before (i.e., there now is one additional parameter to be determined). Choosing an initial guess $q_0 = [1.0625, 0.75, 2.0, 0.375]$ and solving the corresponding offline problem with data, exact data $z_{\text{offline}} = [z_V([0, 50]), z_H([0, 50])]$ by use of Tikhonov regularisation (see [8]) yields data attainance by means of the parameter $q_{\text{offline}} = [0.85, 0.8312, 1.6, 0.59]^T$. This shows that only the first and the third component of q_* can be uniquely determined. Hence, given non-uniqueness in the offline problem, online parameter convergence to q_* cannot be expected: though perfect tracking of exact data is again obtained by the online method (3.4), (3.5) with settings as before (tracking is not shown since similar plots as in Figure 6), only the first and the third component of the parameter estimate $\hat{q}$ converge, then towards the true values (see Figure 11). The others strongly oscillate but seem to stay bounded. As mentioned in Section 3, convergence in the output space, i.e., data tracking, does not guarantee convergence in the parameter space towards the solution. The latter can only be obtained if the available data contain enough information, i.e., if the dynamical system is persistently excited. The lack of persistence of

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press



(a) Predicted output $\hat{V}$ and noisy data z_V^δ (b) Predicted output $\hat{H}$ and noisy data z_H^δ

FIGURE 12. Loss of data tracking due to noise and no use of regularisation.



(a) Parameters diverge if no regularization is used (b) Parameters stay bounded due to use of leakage term

FIGURE 13. Behaviour of parameter estimate $\hat{q}$ in presence of data noise.

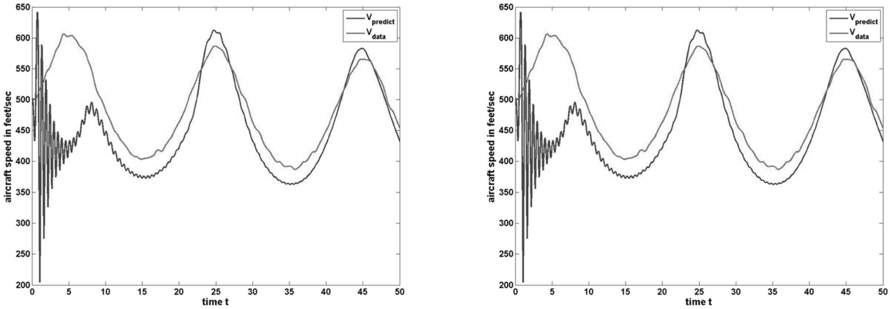
excitation indicated by Figure 11 can have dramatic consequences in case of noisy data; Figures 12 and 13(a) show that if (3.4), (3.5) is applied to noisy data, neither data tracking nor boundedness of the parameter estimate $\hat{q}$ can be guaranteed. Hence, data noise has to be properly taken into account as discussed in Section 4. For instance, the leakage approach (4.2), (4.3) with fixed σ may restore data tracking and parameter boundedness as illustrated in Figures 13(b) and 14 (obtained with $\sigma = 0.5$).

5.3 Heat equation

In our final numerical illustration we consider the infinite-dimensional case. Motivated by an example presented in [3] we focus on the online identification of the space-dependent heat conductivity

$$q_*(x) = 0.1 - 0.05\sin(2\pi(x - 0.25)), \quad x \in \Omega = (0, 1), \tag{5.9}$$

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press



(a) Predicted output $\hat{V}$ and noisy data z_V^δ (b) Predicted output $\hat{H}$ and noisy data z_H^δ

FIGURE 14. Data tracking restored by use of leakage approach.

in the linear heat equation

$$\begin{aligned} u_t(x,t) - \nabla(q(x)\nabla u(x,t)) &= r(x,t), \ x \in \Omega, \ t > 0, \\ u(x,t) &= 0, \ x \in \partial\Omega, \ t > 0, \\ u(x,0) &= u_0(x), \ x \in \Omega, \end{aligned} \tag{5.10}$$

with initial temperature $u_0(x) = 0$ and a heat source

$$r(x,t) = (4\sin(4\pi t) + 0.001t^2)\chi_{[0.215,0.315]}.$$

Here, $\chi_{[0.215,0.315]}$ denotes the characteristic function of the interval $[0.215, 0.315]$. Note that this problem perfectly fits into our abstract setting (1.1) by considering the weak operator formulation of (5.10) (see [23]). The exact data either result from full temperature observations, i.e.,

$$z(x,t) = u_{q_*}(x,t), \ x \in [0,1], \ t > 0, \tag{5.11}$$

or from partial temperature observations on the right half of the domain, i.e.,

$$z(x,t) = u_{q_*}(x,t), \ x \in [0.5,1], \ t > 0, \tag{5.12}$$

where u_{q_*} denotes the unique solution of (5.10) corresponding to the true parameter (5.9). As initial parameter guess we choose the constant function $\hat{q}(x,0) = 0.01$.

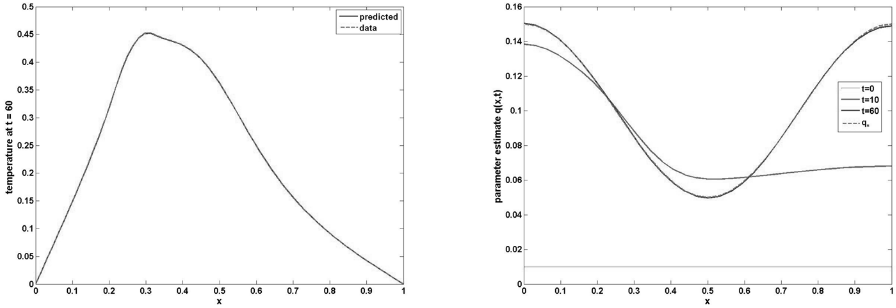
For the numerical realisation the parameter space Q is discretised by linear ansatz functions $\{\varphi_j^m\}_{j=0}^m$ defined on the uniform mesh $\{0, 1/m, ..., 1\}$, i.e.,

$$\varphi_j^m(x) = \begin{cases} 1 - |mx - j| & x \in [\frac{j-1}{m}, \frac{j+1}{m}], \\ 0 & \text{elsewhere on } [0,1]. \end{cases}$$

Similarly, the state space $X = H_0^1(\Omega)$ is discretised by linear ansatz functions $\{\phi_j^n\}_{j=1}^{n-1}$, i.e.,

$$\phi_j^n(x) = \begin{cases} 1 - |nx - j| & x \in [\frac{j-1}{n}, \frac{j+1}{n}], \\ 0 & \text{elsewhere on } [0,1]. \end{cases}$$

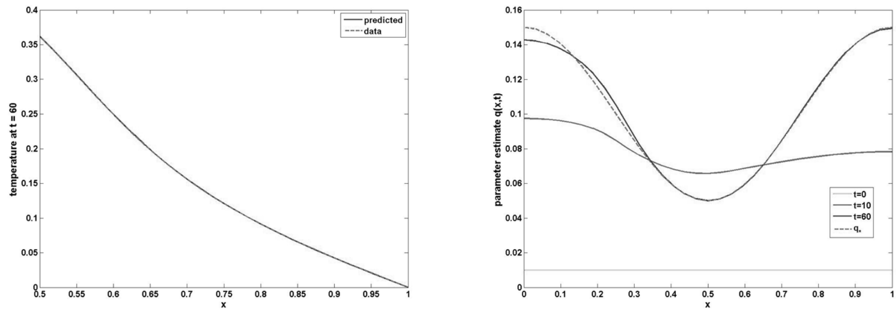
https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press



(a) Data and predicted output at time $t = 60$

(b) Parameter estimate $\hat{q}(x, t_i)$ at several times and true conductivity $q_*(x)$

FIGURE 15. Convergence both in output and parameter space (full state observation).



(a) Convergence in output space (plot for $t = 60$)

(b) Parameter $\hat{q}$ converges towards q_* in region where data are taken

FIGURE 16. Case of partial state observation.

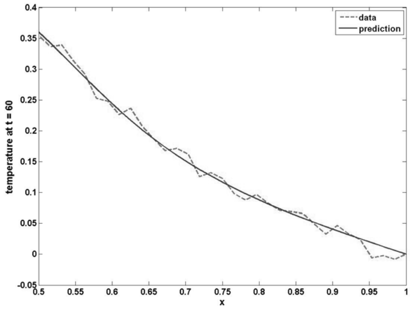
Those functions are also used for the discretisation of the output space $Z = L^2([0,1])$ or $Z = L^2([0.5,1])$, respectively. Figure 15 shows the results obtained by the online method (3.4), (3.5) with

$$\beta = 1,\ G_0 = I,\ \bar{G} = 4I,\ n = 64,\ m = 32 \tag{5.13}$$

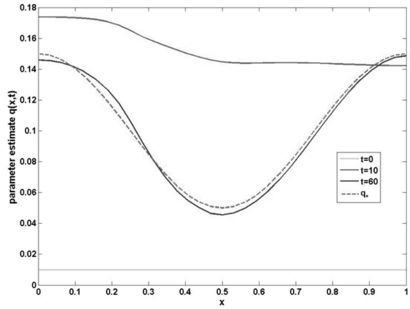
for the case of exact full state observations (5.11). At time $t = 60$ both output and parameter error have vanished. As opposed to the method discussed in [3], (3.4), (3.5) does not require to differentiate the data $z(x,t)$. Our method can also be applied if only partial state observations are available. Figure 16 illustrates that partial data given by (5.12) again are tracked by the predicted output $F(\hat{q}, t)$ while the parameter can at least be uniquely determined on the interval $[0.5, 1]$ (on which the data are taken). Still, $\hat{q}(\cdot, 60)$ is an acceptable online approximation for q_* on all of Ω. Finally, we considered the case of noisy partial data z^δ with

$$\frac{\|z^\delta(\cdot,t) - z(\cdot,t)\|_{L^2(0.5,1)}}{\|z(\cdot,t)\|_{L^2(0.5,1)}} \leqslant 0.05.$$

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press



(a) Noisy data and predicted output at $t = 60$



(b) $\hat{q}(x, t_i)$ and $q_*(x)$

FIGURE 17. Stable result in case of noisy data obtained by use of leakage approach.

Tests for the method (4.2), (4.3) with the setup (5.13) showed that the noise amplification due to the relatively large gain operator G could not be satisfactorily compensated by the leakage term. However, reducing the gain G via the choice $G_0 = 0.1I$ and $\bar{G} = I$ allowed to track the data and recover the parameter even in presence of data noise. Figure 17 shows results obtained with $\sigma = 0.001$.

We mention that we also have run successful tests for the online identification of a temperature-dependent heat conductivity $q = q(u)$ in (5.10). Online identification of such non-linearities poses additional challenges since the domain of the unknown q may not be *a priori* given which then might require to adapt the domain of q during the computations (see [16]). Such problems also may arise in the context of ODEs, for instance think of a general ansatz $C_{D,L,M} = C_{D,L,M}(\alpha)$ for the aerodynamic coefficients in (5.4). That topic will be discussed in a separate paper.

6 Conclusions and outlook

In this paper we discussed the problem of online parameter identification in time-dependent differential equations. Based on ideas from adaptive control and regularisation of non-linear operator equations we suggested a method that can be applied both to finite and infinite non-linear systems and especially allows for partial state observations. Furthermore, data filtering or differentiation is not needed. The method was analysed by means of Lyapunov theory and illustrated by means of three numerical examples.

The list of possible future works is long. From the theoretical point of view emphasis especially has to be put on strategies for choosing the method parameters β, $\bar{G}$ and σ in dependency of the noise level and the data itself. While such parameter choice rules are well developed in context of regularisation of non-linear operator equations, see [8], no corresponding theory based advice currently is – to the best of our knowledge – available for online problems. Time-varying method parameters might accommodate situations in which the quality of measurements changes during the operation of the system. Another theoretical challenge is to ease the assumptions under which a convergence analysis can be performed. One idea is to use so-called averaging techniques (see [17, 19, 22]), where for the analysis the convergence properties of (3.4) would be related to that of an

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

averaged estimate

$$\hat{q}_t^{\mathrm{av}}(t) = \lim_{T\to\infty} \frac{1}{T} \int_0^T G(s)F'(\hat{q}^{\mathrm{av}}, s)^*(z(s) - F(\hat{q}^{\mathrm{av}}, s))\, ds.$$

Furthermore, the dead-zone approach as well as projection methods and ϵ-techniques (see [14]) need to be theoretically backed up for the non-linear case in order to provide reliable alternatives to the leakage approach in presence of noisy data.

From the numerical point of view, the disadvantage of our method is that after discretisation (if necessary) the dimension of the total system to be integrated is $m^2 + mn + m + n$ where m and n denote the dimensions of the parameter and state vectors, respectively. Hence, the integration of (3.4), (3.5) may become very costly due to dimensional aspects. Similar numerical challenges arise in the context of Kalman filtering, where also dynamical systems for covariance matrices of dimension m^2 have to be integrated (see [10]). Techniques based on a reformulation of the Riccati equation such as square root filtering might be transferred to (3.5) in order to decrease the numerical costs.

From the practical point of view, the next step would be to integrate the online identification method into controllers such that the resulting adaptive control method is robust. Especially, we are interested in testing the method by means of auto-pilots based on non-linear 3DoF as well as 6DoF aircraft models. Finally, for any real-world realisation the online identification method needs to be adapted to digital signals, i.e., signals that are discrete and quantised (see [9]) and that can be delayed. The resulting algorithm needs to be implemented such that the online identification can be executed in real time with the available computational resources. The numerical results of Section 5 were obtained in Matlab with 2x Dualcore Intel Xeon CPU (Xeon 5130 2.00 GHz) and 4 GB Ram. For instance, the integration of (5.2), (5.3), (5.4) by the Matlab built in ode45 routine with time interval $[0, 200]$ default settings takes about 6 s. This CPU time could certainly be reduced by the use of especially tailored integration techniques or, e.g., a switch to C + + coding. However, in order to judge if the CPU time needed is short enough for the online application in mind, the true time scale of the underlying real-world process would be needed. Without, the quality of an online method can only be discussed in a relative manner by defining a reference problem to which several techniques are applicable and then comparing them with respect to their speed. Such a comparison is subject of our future work.

Acknowledgements

The author learned about the field of online parameter identification during the special semester on inverse problems at IPAM, UCLA, Los Angeles. The author is thankful to IPAM for the invitation and the financial support. Furthermore, fruitful discussions with H.W. Engl, RICAM and University of Vienna, and E. Lavretsky, Boeing Company, are kindly acknowledged.

References

[1] Åström, K. J. & Wittenmark, B. (1995) *Adaptive Control*, Addison-Wesley, Reaching, Massachusetts.

[2] Banks, H. T. & Kunisch, K. (1989) *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston.

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press

[3] Baumeister, J., Scondo, W., Demetriou, M. A. & Rosen, I. G. (1997) Online parameter estimation for infinite-dimensional dynamical systems. *SIAM J. Control Optim.* **35**, 678–713.

[4] Böhm, M., Demetriou, M. A., Reich, S. & Rosen, I. G. (1998) Model reference adaptive control of distributed parameter systems. *SIAM J. Control Optim.* **36**, 33–81.

[5] Demetriou, M. A. & Rosen, I. G. (1994) On the persistence of excitation in the adaptive identification of distributed parameter systems. *IEEE Trans. Automat. Control* **39**, 1117–1123.

[6] Demetriou, M. A. & Rosen, I. G. (2001) On-line robust parameter identification for parabolic systems. *Int. J. Adapt. Control Signal Process.* **15**, 615–631.

[7] Deuflhard, P. & Bornemann, F. (2002) *Scientific Computing with Ordinary Differential Equations*, Springer, New York.

[8] Engl, H. W., Hanke, M. & Neubauer, A. (1996) *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht.

[9] Franklin, G. F., Powell, J. D. & Workman, M. (1997) *Digital Control of Dynamic Systems*, Addison Wesley Longman, Menlo Park, California.

[10] Grewal, M. S. & Andrews, A. P. (2001) *Kalman Filtering–Theory and Practice Using MATLAB*, John Wiley & Sons, New York.

[11] Haltmeier, M., Leitao, A. & Scherzer, O. (2007) Kaczmarz methods for regularizing nonlinear ill-posed equations I: Convergence analysis. *Inverse Probl. Imaging* **1**, 289–298.

[12] Hoffmann, K. H. & Sprekels, J. (1984–85) On the identification of coefficients of elliptic problems by asymptotic regularization. *Numer. Funct. Anal. Optimiz.* **7**, 157–177.

[13] Hoffmann, K. H. & Sprekels, J. (1986) Inequalities by asymptotic regularization. *SIAM J. Math. Anal.* **17**, 1198–1217.

[14] Ioannou, P. & Sun, J. (1996) *Robust Adaptive Control*, Prentice Hall, Upper Saddle River, New Jersey.

[15] Jategaonkar, R. & Thielecke, F. (2002) ESTIMA–An integrated software tool for nonlinear parameter estimation. *Aerosp. Sci. Technol.* **6**, 565–578.

[16] Kügler, P. & Engl, H. W. (2002) Identification of a temperature-dependent heat conductivity by Tikhonov regularization. *J. Inverse Ill-posed Probl.* **10**, 67–90.

[17] Landau, I. D., Lozano, R. & M'Saad, M. (1998) *Adaptive Control*, Springer, London.

[18] Ljung, L. (1999) *System Identification–Theory for the User*, Prentice Hall, Upper Saddle River, New Jersey.

[19] Ljung, L. & Söderström, T. (1987) *Theory and Practice of Recursive Identification*, MIT Press, Boston.

[20] Narendra, K. S. & Annaswamy, A. M. (1989) *Stable Adaptive Systems*, Dover Publications, New York.

[21] Orlov, Y. & Bentsman, J. (2003) Adaptive distributed parameter systems identification with enforceable identifiability conditions and reduced-order spatial differentiation. *IEEE Trans. Automat. Control* **45**, 203–216.

[22] Sastry, S. & Bodson, M. (1989) *Adaptive Control–Stability, Convergence and Robustness*, Prentice Hall, Englewood Cliffs, New Jersey.

[23] Showalter, R. (1997) *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Mathematical Surveys and Monographs 49, American Mathematical Society, Providence, Rhode Island.

[24] Slotine, J. E. & Li, W. (1991) *Applied Nonlinear Control*, Prentice Hall, Englewood Cliffs, New Jersey.

[25] Stevens, B. L. & Lewis, F. L. (2003) *Aircraft Control and Simulation*, John Wiley & Sons, Hoboken, New Jersey.

[26] Tao, G. (2003) *Adaptive Control, Design and Analysis*, John Wiley & Sons, Hoboken, New Jersey.

[27] Tautenhahn, U. (1994) On the asymptotical regularization of nonlinear ill-posed problems. *Inverse Probl.* **10**, 1405–1418.

[28] Temam, R. (1988) *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Applied Mathematical Sciences 68, Springer, New York.

[29] Xu, G. & Yung, S. P. (2003) Lyapunov stability of abstract nonlinear dynamic in Banach space. *IMA J. Math. Control Inf.* **20**, 105–127.

https://doi.org/10.1017/S0956792508007547 Published online by Cambridge University Press