An online parameter identification method for time dependent partial differential equations

R Boiger and B Kaltenbacher

Institute of Mathematics, Alpen-Adria-Universität Klagenfurt, Austria

E-mail: romana.boiger@aau.at and barbara.kaltenbacher@aau.at

Received 1 June 2015, revised 24 November 2015
Accepted for publication 9 December 2015
Published 15 March 2016

Abstract

Online parameter identification is of importance, e.g., for model predictive control. Since the parameters have to be identified simultaneously to the process of the modeled system, dynamical update laws are used for state and parameter estimates. Most of the existing methods for infinite dimensional systems either impose strong assumptions on the model or cannot handle partial observations. Therefore we propose and analyse an online parameter identification method that is less restrictive concerning the underlying model and allows for partial observations and noisy data. The performance of our approach is illustrated by some numerical experiments.

Keywords: parameter identification, online estimation, convergence analysis

( Some figures may appear in colour only in the online journal)

1. Introduction

Dynamical systems like ordinary differential equations or time dependent partial differential equations play an important role for modelling instationary processes in science and technology. Such models often contain parameters that cannot be accessed directly and therefore must be determined from measurements, which leads to inverse problems. In many applications, e.g., in model predictive control, the parameter identification has to take place during the operation of the considered system. Hence online methods become necessary. Examples of applications range from heating ventilation airconditioning (HVAC) systems via battery charge estimation to aircraft dynamics, see e.g. [7, 10, 12].
In many applications we face the additional problem of having only partial and noisy observations of the state. Motivated by these facts, in this paper we propose an online parameter identification method that is also applicable in case of indirect partial observations and takes into account noisy data. For this purpose we employ a dynamic update law for both the estimated parameters and the state estimate that is strongly inspired by the schemes from [1] and [9]. Online parameter identification has been extensively studied in the finite dimensional setting, e.g. [6, 11] or [13]. The literature becomes much more scarce when dealing with infinite dimensional models as arising in the context of partial differential equations. We refer to the extensive literature review in [1] and [9]. More recent work on this topic can e.g. be found in [2].

The paper is organised as follows: in section 2 we state the underlying differential equation with the according assumptions and define the online parameter identification method. In the next section the convergence analysis of the method is discussed for the exact data case, the case with noisy data and also the one with smooth noisy data. An application example and numerical experiments illustrate the performance of the method in section 4. We conclude with some remarks and an outlook in section 5.

2. Online parameter identification method

In this section we present the underlying general model and the corresponding assumptions. Furtheron we introduce an online parameter identification method.

Let \( Q, X \) and \( Z \) be Hilbert spaces. We consider the abstract ordinary differential equation

\[
\begin{align*}
  \dot{u}_t(t) + C(q, u(t)) &= f(t) \quad t > 0 \\
  u(0) &= u_0, \\
\end{align*}
\]

where \( C : Q \times \mathcal{D}(C) (\subseteq Q \times X) \to X \), \( f : [0, \infty) \times X \to X \) and the initial value \( u_0 \) are given. The inverse problem we are interested in is to find the stationary parameter \( q \) from given observations of the state \( u \) over time, \( Gu(t, x) = z(t, x) \), where \( G : X \to Z \) is the observation operator and \( Z \) the observation space. For simplicity of exposition we consider a linear observation operator here. Most of what follows can be carried over to the case of nonlinear observations.

We will denote the exact solution by \( q^* \) and \( u^* \). To define an evolution system for identifying \( q^* \) from measurements \( z \) we split \( u^* \) in its ‘observed part’ \( Ru^* = G^* z \in \mathcal{N}(G)^\perp \subseteq \tilde{V} \) and its ‘unobserved’ part \( Pu^* = u^* - Ru^* \in \mathcal{N}(G) \subseteq \tilde{V} \) by appropriate projections \( R \) and \( P \). Here \( \tilde{V} \subseteq \tilde{X} \subseteq X \) and \( \tilde{V} \subseteq \tilde{X} \subseteq X \) with the corresponding embedding constants \( C_{\tilde{V} \tilde{X}} \), \( C_{\tilde{V}X} \), \( C_{\tilde{X} \tilde{X}} \) and the operator \( G^* : Z \to X \) is the Moore-Penrose Inverse of \( G \). Hence the projection \( R \) for the ‘observed’ part is the projection on the orthogonal complement of the nullspace of \( G \), i.e., \( R : X \to \mathcal{N}(G)^\perp \), \( R = G^* G \), and \( P \) is the orthogonal projection onto the nullspace of \( G \), that is \( P : X \to \mathcal{N}(G) \), \( P = I - R \).

Assumption 2.1. For the abstract ODE (1) we assume that

(i) the exact solution \( u^* \) exists and stays bounded, i.e. there exists \( \rho > 0 \) such that for all times \( t > 0 \) we have \( u^*(t) \in B_{\rho}(u_0) \subseteq \mathcal{D}(C) \), where \( B_{\rho}(u_0) = \{ v + w \in \tilde{V} + \tilde{V} | \| v - Ru_0 \| + \| w - Pu_0 \| \leq \rho \} \);
(ii) the operator \( C \) satisfies a Lipschitz condition with respect to the second variable, i.e. for all times \( t > 0 \) and for all \( v + w \in \tilde{V} + \tilde{V} \)

\[
\| C(q^*, u^*(t) + v + w) - C(q^*, u^*(t)) \|_X \leq L_C (\| v \|_X + \| w \|_X) \quad (2)
\]

holds;
(iii) the operator $C$ can be split in a part that is dependent of $q$ and the rest:

$$C(q, u) = A(u)q + B(u),$$

(iv) for all $u \in B^{\nu}(u_0)$ the operator $A(u): Q \to X$ is linear and bounded and there exists $C_A > 0$ such that

$$\|A(u^* + v)\|_{Q \to X} \leq C_A (1 + \|v\|_\psi) \quad \forall \ v \in \tilde{V}$$

or

$$\|A(u^* + v)\|_{Q \to X} \leq C_A (1 + \|v\|_X) \quad \forall \ v \in X$$

or

$$\|A(u^* + v + w)\|_{Q \to X} \leq C_A (1 + \|v\|_\psi + \|w\|_\psi) \quad \forall \ v \in \tilde{V}, \ w \in \tilde{V};$$

(v) there exist coercive and bounded operators $M: \tilde{V} \to X$ and $N: \tilde{V} \to X$ i.e., there exist positive constants $c_M, c_N, C_M, C_N$ such that for all $v \in \tilde{V}, \ w \in \tilde{V}$

$$(Mv, v)_X \geq c_M \|v\|_{\tilde{V}}^2 \text{ and } \|RMv\|_X \leq C_M \|v\|_{\psi}$$

$$(Nw, w)_X \geq c_N \|w\|_{\tilde{V}}^2 \text{ and } \|PNw\|_X \leq C_N \|w\|_{\psi}.$$ 

Note that by continuity of the embeddings $\tilde{V} \hookrightarrow X, \tilde{V} + \tilde{V} \hookrightarrow X$, (4) is sufficient for (3), (5). Conditions (ii), (iii), and (iv) are similar to assumptions 1 and 2 in [9].

The method we will propose here is strongly motivated by the methods proposed by Kügler [9] and by Baumeister et al [1]. The main difference compared to [1] is that we also allow for partial observations, which often occur in applications. This is also to some extent possible with the method from [9], however in contrast to [9] we do not assume monotonicity of the operator $C$.

Online identification means that the parameter identification, the data collection process and the operation of the system are taking place at the same time. Accurate parameter values are needed for making decisions while the system is in operation. Therefore our online parameter identification method includes a dynamical update law for both the parameter and the state estimate. To derive such an update law, we first of all consider $q$ as a function of time as well (however, a constant one, taking the value $q^*$ for all time instances) to rewrite (1) as

$$q_i(t) = 0 \quad t > 0,$$

$$u_i(t) + C(q(t), u(t)) = f(t) \quad t > 0,$$

$$(q, u)(0) = (q^*, u_0).$$

Notationally, in the following we will identify the exact parameter $q^*$ with the time constant function with value $q^*$ and skip the time variable in the differential equations. Now the dynamical system for the estimators $\hat{q}, \hat{u}$

$$\hat{q}_t - A(Ru^* + P\hat{u})^*(R\tilde{u} - Ru^*) = 0,$$

$$\hat{u}_t + C(\hat{q}, Ru^* + P\hat{u}) + \mu R - Ru^* = f,$$

$$\mu PNP\hat{u} = f.$$
\[(q, \hat{u})(0) = (\hat{q}_0, \hat{u}_0)\] (11)
is supposed to mimic (6)–(8) and at the same time drive the modelled data \(\hat{R}u\) towards the ‘observed’ one \(Ru^* = G^1z\) as time tends to infinity. The latter is achieved by the term 
\[\muRM \frac{R\hat{u} - Ru^*}{\|R\hat{u} - Ru^*\|_\Omega},\]
where \(M\) satisfies assumption 2.1(v), with an appropriate choice of \(\mu > 0\). Similarly, the term \(\nuPN\hat{u}\) with \(N\) according to assumption 2.1(v) and \(\nu > 0\) will have a stabilising effect. (Since we have no a priori information on the ‘unobserved’ part of the solution, we just prevent it from blowup here.) The rest of the terms in (10) just corresponds to the original equation (7). The term 
\[-A(Ru^* + \hat{P}\hat{u})^*(R\hat{u} - Ru^*)\]
in equation (9) for \(\hat{q}\), provides us with some stability, due to skew symmetry to the corresponding part of \(C(\hat{q}, Ru^* + \hat{P}) = A(Ru^* + \hat{P})\hat{q} + B(Ru^* + \hat{P})\) in (10). Moreover, since the measurement error \(R\hat{u} - Ru^*\) should tend to zero as \(t \to \infty\), this term should vanish asymptotically, so we expect this term to approach the zero right-hand side of (6). In (11), \(\hat{u}_0\) need not necessarily coincide with \(u_0\).

Like in [1], we use a separate coercive operator \(M\) to be independent of whether or not \(C\) is monotone (as assumed in [9]). Additionally, introduction of the term \(\nuPN\hat{u}\) also enables us to control the ‘unobserved’ part of the state and thus deal with partial observations.

We will also consider the case of noisy data, where instead of \(z\) only a noisy version \(z^\delta\) is available and therefore in place of \(Ru^*\) a possibly regularised version \(u^\delta_0 = G_\delta z^\delta \approx G^1z\) in (9)–(11) has to be used
\[\hat{q}_t - A(u^\delta_0 + \hat{P}\hat{u})^*(R\hat{u} - u^\delta_0) = 0,\] (12)
\[\hat{u}_t + C(\hat{q}_t, u^\delta_0 + \hat{P}\hat{u}) + \muRM \frac{R\hat{u} - u^\delta_0}{\|R\hat{u} - u^\delta_0\|_\Omega} + \nuPN\hat{u} = f,\] (13)
\[(\hat{q}, \hat{u})(0) = (\hat{q}_0, \hat{u}_0),\] (14)
where \(\alpha = \alpha(t), \mu = \mu(t)\) and \(\nu = \nu(t)\) are chosen properly dependent on the noise level \(\delta(t)\) in
\[\delta(t) \geq \|z^\delta(t) - z(t)\|_Z.\] (15)

In the following section we are going to prove convergence of the ‘observed’ part of the state as \(t \to \infty\). Under a persistence of excitation condition this will also imply convergence of the parameter. Here certain link conditions between ‘observed’ and ‘unobserved’ part of the state are imposed to guarantee control over the ‘unobserved’ part of the state.

3. Convergence analysis

In this section we analyse convergence of the estimator in case of smooth noisy data (including the exact data case) as well as in case of arbitrary noisy data, respectively.

To do so we take a look at the errors between the exact parameter and projected state \((q^*, Ru^*, Pu^*)\) and the estimated counterparts \((\hat{q}, R\hat{u}, \hat{P}\hat{u})\)
\[e = \hat{q} - q^*, \quad r = R\hat{u} - Ru^*, \quad p = \hat{P}\hat{u} - Pu^*\] (16)
as well as the errors including the regularised version of the ‘observed’ part
\[r^\delta_0 = R\hat{u} - u^\delta_0 = r - d^\delta_0\]
where \(d^\delta_0 = u^\delta_0 - R\hat{u}\) (17)
which we will assume to be bounded

$$\sup_{t>0} \|d_{0t}^b\|_X < \infty.$$  \hfill (18)

For these quantities the identities

$$u_{0t}^b + P\hat{u} = u_{0t}^b + P\hat{u} + u^* - Ru^* - Pu^* = u^* + p + d_{0t}^b$$  \hfill (19)

hold.

In case noisy data \(z^b\) are given instead of \(z\) and the range of \(G\) is non-closed, the quantity \(G^*z^b\) might not be well-defined, and even if it is well-defined it will not depend on \(z^b\) in a stable manner. Thus we define a regularised version of the ‘observed’ part of \(u^*\)

$$u_{0t}^b = G_0z^b$$

with \(G_0\) a regularised version of \(G^\dagger\) with regularisation parameter \(\alpha\), defined, e.g., by the Tikhonov–Philips method

$$G_0 = (G^*G + \alpha I)^{-1}G^*: Z \to \mathcal{N}(G)^\perp \subseteq X$$

with \(G^\dagger: Z \to X\) the Hilbert space adjoint of \(G: X \to Z\), and \(\alpha > 0\) appropriately chosen.

 Therewith the online parameter identification method (12)–(14) can be rewritten as

$$\dot{q}_t - A(u^* + p + d_{0t}^b)(R\hat{u} - u_{0t}^b) = 0,$$  \hfill (20)

$$\dot{u}_t + C(q_t, u^* + p + d_{0t}^b) + \mu RM\frac{R\hat{u} - u_{0t}^b}{\|R\hat{u} - u_{0t}^b\|_Y} + \nu PNP\hat{u} = f,$$  \hfill (21)

$$(\hat{q}, \hat{u})(0) = (\hat{q}_0, \hat{u}_0).$$  \hfill (22)

### 3.1. Convergence with exact or smooth noisy data

With smooth(ed) noisy data we denote \(z^b\) that is differentiable with respect to time (which can be achieved by averaging over sufficiently large time intervals), i.e.

$$\|z^b - z(t)\|_Z = \|z^b(t) - z(t)\|_Z \leq \hat{\delta}(t).$$

For the error in the ‘observed’ part of the state this yields

$$r_t = [r_{0t}^b + d_{0t}^b] = r_{0t}^b + [G_0z^b - G^*z_t] = r_{0t}^b + \tilde{d}_{0t}^b$$

with

$$\tilde{d}_{0t}^b := G_0(z^b - z_t) + (G_0 - G^h)z_t + \alpha t\frac{d}{dt}G_0(z^b).$$  \hfill (23)

By subtracting (6)–(8) from (20)–(22) we see that the error components satisfy the following system of differential equations, where we split up the differential equation for the state in the ‘observed’ and the ‘unobserved’ part

$$e_t = A(u^* + d_{0t}^b + p)^b r_{0t}^b = 0,$$  \hfill (24)
\[ r_c^\delta + \hat{d}_\alpha^\delta + RC(q^\delta, u^\delta + p + d_\alpha^\delta) - RC(q^\delta, u^\delta) \]
\[ + RA(u^\delta + p + d_\alpha^\delta) e + \mu RM \frac{r_c^\delta}{\|r_c^\delta\|} = 0, \tag{25} \]

\[ p_0 + PC(q^\delta, u^\delta + d_\alpha^\delta + p) - PC(q^\delta, u^\delta) + PA(u^\delta + p + d_\alpha^\delta) e + \nu PNP\hat{u} = 0 \]

\((e, r_c^\delta, p)(0) = (\hat{q}_0 - q^\delta, R\hat{u}(0) - G_0 z^0(0), P(\hat{u}_0 - u_0)).\)

Here we have used the identities (19) and

\[ C(\hat{q}_0, u_0^\delta + GR) - C(q^\delta, u^\delta) \]
\[ = C(q^\delta, u^\delta + p + d_\alpha^\delta) - C(q^\delta, u^\delta) + A(u^\delta + p + d_\alpha^\delta)q^* \tag{27} \]

as well as assumption 2.1(iii).

And the following analysis includes the case of exact data \( z^0 = z \) with \( G_0 = G^\dagger, \)
\( u_0^\delta = Ru^s, \) and \( d_\alpha^\delta = \hat{d}_\alpha^\delta. \) We mention in passing that in the exact data case, instead of
\((5) \) only \((3) \) will be needed.

### 3.1.1. Well-definedness

To obtain existence and boundedness of the solutions according to our method (12)--(14), we first multiply (24) and (25) with \( e \) and \( r_c^\delta \) respectively to get, using assumptions 2.1(ii), (v)

\[ \frac{d}{dt} \left[ \|e\|^2_Q + \|r_c^\delta\|^2_X \right] = (e_t, e)_Q + (r_c^\delta, r_c^\delta)_X \]
\[ = -((\hat{d}_\alpha^\delta + RC(q^\delta, u^\delta + p + d_\alpha^\delta) - RC(q^\delta, u^\delta) + \mu RM \frac{r_c^\delta}{\|r_c^\delta\|}, r_c^\delta)_X \]
\[ \leq (L_C(\|d_\alpha^\delta\|_V + \|p\|_V) + \|\hat{d}_\alpha^\delta\|_X) \|r_c^\delta\|_X - C_M \mu \frac{\|r_c^\delta\|^2_X}{\|r_c^\delta\|_V}. \tag{28} \]

We see that the equation for \( \hat{q}_0 \) was designed such that the terms containing \( A \) cancel out. The above estimate leads us to choose \( \mu \) according to

**Assumption 3.1.** For all \( t > 0 \)

\[ \mu(t) \geq \frac{2}{c_M} (L_C(\|d_\alpha^\delta(t)\|_V + \|p(t)\|_V) + \|\hat{d}_\alpha^\delta(t)\|_X) \frac{\|r_c^\delta(t)\|_X}{\|r_c^\delta(t)\|_V}. \]

Therewith we obtain

\[ \frac{d}{dt} \left[ \|e\|^2_Q + \|r_c^\delta\|^2_X \right] \leq - \frac{c_M}{2} \mu \frac{\|r_c^\delta\|^2_X}{\|r_c^\delta\|_V} \]
\[ \leq - (L_C(\|d_\alpha^\delta\|_V + \|p\|_V) + \|\hat{d}_\alpha^\delta\|_X) \|r_c^\delta\|_X < 0. \tag{29} \]

This particularly implies boundedness

\[ \forall t > 0: \|e(t)\|^2_Q + \|r_c^\delta(t)\|^2_X \leq \|e(0)\|^2_Q + \|r_c^\delta(0)\|^2_X, \]
and finiteness of the integral
\[ \forall T > 0 : \int_0^T 2(L_C(d_u^t\|\nu + \| p \|_p + ||d_u^t||_X) \times \|c_u^t\|_X \, dt \leq \| \epsilon(0) \|^2_{\mathcal{H}} + \| r(0) \|^2_{\mathcal{H}} < \infty. \]

Now it remains to find an appropriate bound for the error of the ‘unobserved’ part of the state, which can be done in a quite similar manner. For this purpose we multiply (26) with \( p \) and use assumptions 2.1(ii), (iv) with (5) as well as (26) to obtain

\[
\frac{d}{dt} \| [p]_X \|^2 = (p, p)_X
\]

\[= -(P_C(q^*, u^* + d_u^t + p) - P_C(q^*, u^*), p)_X
\]

\[+ (P_A(u^* + d_u^t + p)e, p)_X - (\nu PN P\tilde{u}, p)_X
\]

\[\leq (L_C(\|d_u^t\|_X + \| p \|_p) + ||d_u^t||_X)\| p \|_X + C_A(1 + \| p \|_p + ||d_u^t||_X)\| e \|_Q \| p \|_X
\]

\[- \nu (PN (p + P u^*), p)_X .
\]

For the second and the last term we use Young’s inequality and the embedding inequalities to get

\[C_A\| e \|_Q \| p \|_X \leq \frac{C_A^2}{2} \| e \|_Q^2 + \| p \|_X^2 \leq \frac{C_A}{2} \| e \|_Q^2 + \frac{C_A}{2} C_{\bar{\psi}X} C_{\bar{\psi}X} \| p \|_X \| p \|_X ,
\]

and

\[- \nu (PN P u^*, p)_X \leq \nu C_N \| P u^* \|_Q \| p \|_X \leq \nu \left( \frac{C_N^2 C_{\bar{\psi}X}^2 \| P u^* \|_Q^2}{2C_N} + \frac{C_N}{2} \| p \|_{\bar{\psi}X}^2 \right). \quad (30)
\]

So altogether we have

\[
\frac{d}{dt} \| [p]_X \|^2 = (L_C + C_A\| e \|_Q + ||d_u^t||_X)\| p \|_X
\]

\[+ \frac{C_A}{2} C_{\bar{\psi}X} C_{\bar{\psi}X} \| p \|_X \| p \|_X + C_A \| e \|_Q^2 - \nu \frac{C_N}{2} \| p \|_{\bar{\psi}X}^2 + \nu \frac{C_N^2 C_{\bar{\psi}X}^2 \| P u^* \|_Q^2}{2C_N}.
\]

This leads us to choose \( \nu \) according to

**Assumption 3.2.** For all \( t > 0 \)

\[\nu(t) \geq \max \left\{ L_C \left( \frac{4(L_C + C_A\| e(t) \|_Q)}{c_N} (\| p(t) \|_p + \| d_u^t(t) \|_X) + \frac{2C_A C_{\bar{\psi}X} C_{\bar{\psi}X}}{c_N} \| p(t) \|_Q \right) \frac{\| p(t) \|_X}{\| p(t) \|_{\bar{\psi}X}} \right\}.
\]
With this we obtain
\[
\frac{d}{dt} \frac{1}{2}\|p\|^2_{\mathcal{H}} \leq -\frac{\nu C_N}{4}\|p\|^2_{\mathcal{H}} + \frac{C_N^2 C_{\mathcal{H}}}{2\nu} \|Pu^*\|^2_{\mathcal{V}} + \frac{c_N}{2}\|e\|^2_{\mathcal{Q}}.
\]
We now define \(\tilde{\mathcal{V}}(\tau(t)) = \mathcal{V}(t) = \frac{1}{2}\|p(t)\|^2_{\mathcal{H}}\) and \(\mathcal{V}(t) := c_N^2 \int_0^t \nu(\xi) \, d\xi\) and hence
\[
\frac{d}{dt} \mathcal{V}(t) = \frac{d}{dt} \frac{1}{2}\|p(t)\|^2_{\mathcal{H}} = \frac{d}{dt} \frac{1}{2}\|p(t)\|^2_{\mathcal{H}} - \frac{4}{c_N\nu(t)}.
\]
We have that for any differentiable non-negative function \(\eta: [0, T] \rightarrow \mathbb{R}_{+}\) and \(a, b > 0\) and for all \(t \in [0, T]\) the following implication holds:
\[
\eta'(t) \leq -a\eta(t) + b \Rightarrow \eta(t) \leq \frac{b}{a} + (\eta(0) - \frac{b}{a})e^{-at} \leq \max\left\{\frac{b}{a}, \eta(0)\right\}.
\]
So with \(a = 1\) and \(b = \frac{c_N}{c_N^2}\left(\|e(0)\|^2_{\mathcal{Q}} + \|r^a_0(0)\|^2_{\mathcal{X}}\right)\) we get:

**Proposition 3.3.** Let assumptions 2.1 with (5), 3.1, and 3.2 hold and let \((\bar{q}, \bar{a}) \in Q \times (\bar{V} + \bar{V})\). Then there exists a solution \((\hat{q}(t), \hat{a}(t)) \in Q \times (\hat{V} + \hat{V})\) for all \(t > 0\) and the following estimates on the parameter and state errors (see (16), (17)) hold.

(i) For all \(t > 0\):
\[
\|p(t)\|^2_{\mathcal{Q}} + \|r^a_0(t)\|^2_{\mathcal{X}} \leq \|e(0)\|^2_{\mathcal{Q}} + \|r^a_0(0)\|^2_{\mathcal{X}}.
\]

(ii) For all \(t > 0\):
\[
\|p(t)\|^2_{\mathcal{X}} \leq \max\left\{\|e(0)\|^2_{\mathcal{Q}} + \|r^a_0(0)\|^2_{\mathcal{X}} + \frac{2C_{\mathcal{Q}}C_{\mathcal{H}}}{c_N}\sup_{t > 0}\|Pu^*(t)\|^2_{\mathcal{V}}\right\};
\]

(iii) \(\int_0^t 2(L_C(\|d^a_0(t)\|_\mathcal{V} + \|p(t)\|_\mathcal{V} + \|\tilde{d}^a_0(t)\|_\mathcal{X})\|r^a_0(t)\|_\mathcal{X}) \, dt \leq \|e(0)\|^2_{\mathcal{Q}} + \|r^a_0(0)\|^2_{\mathcal{X}} < \infty\).

**3.1.2 State convergence.** In this section we will show that the estimated ‘observed’ state converges towards the ‘observed’ part of the exact solution. For improving the state convergence we impose an additional lower bound on \(\mu\) as compared to assumption 3.1 (note that therewith proposition 3.3 still remains valid).

**Assumption 3.4.** There exists a constant \(c_1 > 0\) such that for all \(t > 0\)
\[
\mu(t) \geq \max\left\{\frac{2}{c_M}(L_C(\|d^a_0(t)\|_\mathcal{V} + \|p(t)\|_\mathcal{V} + \|\tilde{d}^a_0(t)\|_\mathcal{X})\|r^a_0(t)\|_\mathcal{X}), c_1\|r^a_0(t)\|_\mathcal{X}\right\} 
\times \frac{\|r^a_0(t)\|_\mathcal{X}^2}{\|r^a_0(t)\|_\mathcal{X}^2}.
\]
Theorem 3.5 (State convergence). Under assumptions 2.1 with (4), 3.2, 3.4 and (18) we have that \( \| \tilde{R} u(t) - G_n z^n(t) \|_X = \| r^n(t) \|_X \to 0 \) as \( t \to \infty \).

Proof. We first take a look at the ‘observed’ state error for \( t_2 > t_1 > 0 \), for which we get from (25)

\[
\| r^n(t_2) \|_X^2 - \| r^n(t_1) \|_X^2 = \int_{t_1}^{t_2} \frac{d}{dt} \| r^n(t) \|_X^2 dt = \int_{t_1}^{t_2} (r^n, r^n)_X dt
\]

\[
= \int_{t_1}^{t_2} (-\bar{d}_n^* + RC(q^*, u^*) - RC(q^*, u^* + p + d^n) - RA(u^* + p + d^n) e, r^n)_X
\]

\[
= -(\mu M \frac{r^n}{\| r^n \|_Y}, r^n)_X dt.
\]

We make use of the fact that by assumption 2.1 with (4) we get

\[
| (RA(u^* + p + d^n) e, r^n)_X | \leq \frac{L_A}{2} (\| e(0) \|_Q + \| r^n \|_X)
\]

with

\[
L_A := C_A (1 + \sup_{t > 0} \| d^n(t) \|_X + \sup_{t > 0} \| p(t) \|_X),
\]

which is finite by proposition 3.3 and (18). With that, an estimate similarly to (28) and assumption 3.4 we have

\[
\| r^n(t_2) \|_X^2 - \| r^n(t_1) \|_X^2 \leq \int_{t_1}^{t_2} \left\{ - (L_C (\| d^n \|_Y + \| p \|_Y) + \| d^n \|_X) \| r^n \|_X + \frac{L_A}{2} (\| e(t) \|_Q^2 + \| r^n(t) \|_X^2) \right\} dt
\]

\[
\leq \int_{t_1}^{t_2} \frac{L_A}{2} (\| e(t) \|_Q^2 + \| r^n(t) \|_X^2) dt \leq c_2 (t_2 - t_1)
\]

with \( c_2 := \frac{L_A}{2} (\| e(0) \|_Q^2 + \| r^n(0) \|_X^2) \), which follows from proposition 3.3. Using this estimate we get for any \( t > \gamma > 0 \) fixed

\[
\gamma \| r^n(t) \|_X^2 = \int_{t-\gamma}^{t} (\| r^n(\tau) \|_X^2 + (\| r^n(\tau) \|_X^2 - \| r^n(\tau) \|_X^2)) d\tau
\]

\[
\leq \int_{t-\gamma}^{t} \| r^n(\tau) \|_X^2 d\tau + c_2 \int_{t-\gamma}^{t} (t - \tau) d\tau = \int_{t-\gamma}^{t} \| r^n(\tau) \|_X^2 d\tau + c_2 \gamma^2.
\]

Hence we have for all \( t > \gamma > 0 \) that

\[
\int_{t-\gamma}^{t} \| r^n(\tau) \|_X^2 d\tau \geq \gamma \| r^n(t) \|_X^2 - \frac{c_2 \gamma^2}{2}.
\]

From (29) and assumption 3.4 we get

\[
\frac{1}{2} \| e(t) \|_Q^2 + \| r^n(t) \|_X^2 \leq - \int_{t_1}^{t_2} \frac{c_M}{2} \frac{\| r^n(t) \|_Y^2}{\| r^n(t) \|_Y} dt
\]

\[
\leq - \frac{c_M c_1}{2} \int_{t_1}^{t_2} \| r^n(t) \|_X^2 dt,
\]
hence
\[
\int_0^\infty \| r_0^s (t) \|_X^2 \, dt \leq \frac{\| e(0) \|_X^2 + \| r_0^s(0) \|_X^2}{c_M c_1} < \infty. \tag{36}
\]

We want to show that \( \lim_{t \to \infty} \| r_0^s(t) \| = 0 \). So we suppose that \( \lim_{t \to \infty} \| r_0^s(t) \| \neq 0 \). If this is the case then there exists a sequence \( (t_i)_{i \in \mathbb{N}} \) with \( t_i \to \infty \) for \( i \to \infty \), and an \( \varepsilon > 0 \) such that for all \( i \in \mathbb{N} \) \( \| r_0^s(t_i) \|_X^2 \geq \varepsilon \). Now we select a subsequence \( (t_{i_j})_{j \in \mathbb{N}} \) such that for all \( j \in \mathbb{N} \) we additionally have \( t_{i_j} - t_{i_j - 1} > \frac{\varepsilon}{c_2} \). Because of inequality (34), choosing \( \gamma = \frac{\varepsilon}{c_2} \) we have
\[
\frac{\varepsilon^2}{2c_2} \leq \int_{t_{i_j} - \gamma}^{t_{i_j}} \| r_0^s(\tau) \|_X^2 \, d\tau.
\]
By summing up on both sides and using \( t_{i_j} - \gamma = t_{i_j} - \frac{\varepsilon}{c_2} \geq t_{i_j - 1} \), we get for all \( n \in \mathbb{N} \)
\[
n \frac{\varepsilon^2}{2c_2} \leq \sum_{j=1}^{n} \int_{t_{i_j} - \gamma}^{t_{i_j}} \| r_0^s(\tau) \|_X^2 \, d\tau \leq \int_0^{\infty} \| r_0^s(\tau) \|_X^2 \, d\tau \leq \int_0^\infty \| r_0^s(\tau) \|_X^2 \, d\tau,
\]
which gives a contradiction to (36). \( \square \)

3.1.3. Parameter convergence. The proofs in this section are to some extent similar to those in section 3 of [9]. Note however, that the lemma quantifying the relation between state error and parameter error can be stated in a stronger manner (see lemma 3.11 below), which enables to considerably simplify the final convergence proof, see theorem 3.12 below. In order to show that the parameter error converges to zero we start with some preparatory results. First we prove an estimate on the norm of the ‘observed’ state error.

Lemma 3.6. Under assumption 2.1 with (4) and (18), the projected state errors \( r_0^s = \tilde{R}u - G_0 \tilde{z}^t \) and \( p \) satisfy the following relation for all \( 0 < t_a \leq t_b \leq t_c \):
\[
\| r_0^s(t_a) \|_X \geq \int_{t_a}^{t_b} R(\dot{u}(\tau) + p(\tau) + d_0^s(\tau))e(t_a)\, d\tau - \| r_0^s(t_b) \|_X - \sum_{j=1}^{n} \int_{t_{i_j} - \gamma}^{t_{i_j}} \| r_0^s(\tau) \|_X \, d\tau
\]

\[
- L_2 \int_{t_a}^{t_b} \int_{t_a}^{\tau} \| r_0^s(s) \|_X \, ds \, d\tau
\]

\[
- L C \int_{t_a}^{t_b} (\| d_0^s(\tau) \|_\psi + \| p(\tau) \|_\psi) \, d\tau
\]

\[
- C_M \int_{t_a}^{t_b} \mu(\tau) \, d\tau.
\]

Proof. Integrating (25) with respect to time we obtain
\[
r_0^s(t_c) - r_0^s(t_a) = \int_{t_a}^{t_c} r_0^s(\tau) \, d\tau
\]

\[
= \int_{t_a}^{t_c} \left( -d_0^s + R(C(q^*, u^*) - C(q^*, u^* + p + d_0^s)) \right)
\]

\[
- R(\dot{u}^* + p + d_0^s)e - \mu RM \frac{r_0^s}{\| r'_{0} \|_\psi} \, d\tau.
\]
Taking the norm we get, using the triangle inequality and the reverse triangle inequality
\[ \|r^\delta(x, t)\|_X + \|r^\delta(x, t)\|_X \geq \|r^\delta(x, t) - r^\delta(x, t)\|_X \]
\[ \geq \left| \int_{t_a}^{t} \left[ RA(u^\gamma + p + d^\delta(x, \tau))e(\tau) - \int_{t_a}^{t} ||d^\delta(x, \tau)||_X \ d\tau \right. \right. \]
\[ \left. \left. - \int_{t_a}^{t} RC(q^\delta, u^\gamma) - RC(q^\delta, u^\gamma + p + d^\delta(x, \tau))\right|_X \ d\tau \right. \]
\[ \geq \left| \int_{t_a}^{t} \left[ RA(u^\gamma + p + d^\delta(x, \tau))e(\tau) - \int_{t_a}^{t} L_{C}(||p||_\psi + \|d^\delta(x, \tau)||_X)\right]_X \ d\tau \right. \]
\[ \left. \left. - \int_{t_a}^{t} C_{M} d\tau \right. \right. \]

where we have used assumption 2.1. Now we have to estimate the remaining first term on the right-hand side. With assumption 2.1 as well as \( L_A \) as in (32) we get
\[ \left| \int_{t_a}^{t} RA(u^\gamma(\tau) + p(\tau) + d^\delta(\tau))e(\tau) \right|_X \]
\[ \geq \left| \int_{t_a}^{t} RA(u^\gamma(\tau) + p(\tau) + d^\delta(\tau))(e(t_a) + e(\tau) - e(t_a))\right|_X \]
\[ \geq \left| \int_{t_a}^{t} RA(u^\gamma(\tau) + p(\tau) + d^\delta(\tau))e(\tau)\right|_X \]
\[ \geq \left| \int_{t_a}^{t} \left[ L_{C}(||p||_\psi + \|d^\delta(x, \tau)||_X)\right]_X \ d\sigma \ d\tau \]

where we used the fact that with \( \tau \geq t_a \) and (24)
\[ \|e(\tau) - e(t_a)\|_Q = \| \int_{t_a}^{\tau} A(u^\gamma + p + d^\delta(\tau))d\sigma \|_Q \leq L_A \int_{t_a}^{\tau} ||d^\delta(x, \tau)||_X \ d\sigma. \]

Combining everything yields the assertion. \( \square \)

Consider the right-hand side in the estimate of lemma 3.6. While by theorem 3.5, the negative terms containing \( r^\delta \) will tend to zero as time tends to infinity, the first (positive) term enables us to enforce parameter convergence by means of a so-called persistence of excitation condition.

**Assumption 3.7** (Persistence of excitation). There are \( T_0, \varepsilon_0, \gamma_0, l > 0 \) such that for all \( t_a \geq t, \xi \in \partial B^l(0) \) there exists a time instance \( t_b \in [t_a, t_a + T_0] \) such that
\[ \left| \int_{t_b}^{t_b + \gamma_0} RA(u^\gamma(\tau) + p(\tau) + d^\delta(\tau))\xi d\tau \right|_X \geq \varepsilon_0. \]

**Remark 3.8.** This condition hinges on a proper choice of the right-hand side \( f \) (‘excitation’) and its verification seems to be hardly possible in an \( a \ priori \) fashion. However, \( a \ posteriori \) control of \( f \) can be viewed as a practical tool for achieving such a relation. On the other hand, note that without such a condition, parameter convergence cannot be expected to be achievable by an online method, that relies on computations forward in time only, while in classical parameter identification for time dependent problems, forward and backward computations over time are fully coupled.
To control the remaining terms $-\int_b^t ||d^2_\nu(\tau)||_X \, d\tau - L_\nu \int_b^t (||d^2_\nu(\tau)||_Y + ||p(\tau)||_Y) \, d\tau$, and $-C_M \int_b^t \mu(\tau) \, d\tau$ on the right-hand side of the estimate in lemma 3.6, we will combine the estimate

$$\frac{1}{2} \||e||_0^2 + ||r^\alpha_\nu||^2_X \leq -\frac{c_M}{2} \mu \frac{\||r^\alpha_\nu||_Y^2}{||r^\alpha_\nu||_Y} \, d\tau = -\frac{c_M}{2} \int_b^t \theta(\tau) \, d\tau,$$

where

$$\theta = \frac{||r^\alpha_\nu||_Y^2}{||r^\alpha_\nu||_Y},$$

that results from (29), with some link conditions.

**Assumption 3.9 (Link conditions).** There exist $\lambda, \kappa \in [1, \infty), T_\beta, T_\gamma > 0$ and $C_\alpha, C_\kappa > 0$ such that for $\gamma_0$ as in assumption 3.7 the following holds. For all $t \geq T_\lambda$

$$C_\alpha \geq \left\{ \begin{array}{ll}
\left( \int_t^{t+\gamma_0} \left( \frac{||p(\tau)||_Y + ||d^2_\nu(\tau)||_Y}{\theta(\tau)} \right)^{\frac{1}{\alpha}} \, d\tau \right)^{\frac{1}{1-\alpha}} & \text{if } \lambda > 1 \\
\sup_{\tau \in [t,t+\gamma_0]} \frac{||p(\tau)||_Y + ||d^2_\nu(\tau)||_Y}{\theta(\tau)} & \text{if } \lambda = 1.
\end{array} \right.$$

For all $t \geq T_\kappa$

$$C_\kappa \geq \left\{ \begin{array}{ll}
\left( \int_t^{t+\gamma_0} \left( \frac{\mu(\tau)^k}{\theta(\tau)} \right)^{\frac{1}{\kappa}} \, d\tau \right)^{\frac{1}{1-\kappa}} & \text{if } \kappa > 1 \\
\sup_{\tau \in [t,t+\gamma_0]} \frac{\mu(\tau)^k}{\theta(\tau)} & \text{if } \kappa = 1.
\end{array} \right.$$

**Remark 3.10.** Sufficient for assumption 3.9 is the existence of some $\rho > 0$ and a constant $C_\rho$ such that for all $t > 0$

$$||p||_Y + ||d^2_\nu||_Y \leq C_\rho ||r^\rho_\nu||_Y$$

and existence of constants $c_{\mu \nu}$ and $C_{\mu \nu}$ respectively $c_\mu$ and $C_\mu$ such that for all $t > 0$ the following interpolation estimate

$$c_{\mu \nu} ||r^\alpha_\nu||_Y ||r^\mu_\nu||_X \leq ||r^\nu_\nu||_Y^2 \leq C_{\mu \nu} ||r^\nu_\nu||_Y ||r^\nu_\nu||_X$$

and also the connecting estimate of $r^\rho_\nu$ and $\mu$

$$c_{\mu \nu} \mu \leq ||r^{\nu \kappa}_\nu||^\nu_{X^\kappa} \leq C_{\mu \nu} \mu$$

holds. This can be seen as follows.

Since we want to estimate the integral $\left( \int_t^{t+\gamma_0} \left( \frac{\mu(\tau)^k}{\theta(\tau)} \right)^{\frac{1}{\kappa}} \, d\tau \right)^{\frac{1}{1-\kappa}}$ we first take a look at the integrand. Using the definition of $\theta$ and the stated interpolation estimate for the state error as well as the connecting estimate of $r^\nu_\nu$ and $\mu$ we get
and so the integral is
\[
\left( \int_1^{t+\gamma} \frac{\mu}{\widehat{\theta}(\tau)} \mathrm{d}\tau \right)^{\frac{1}{\kappa-1}} \leq \left( \int_1^{t+\gamma_0} \frac{1}{c_{int} c_{\mu}} \mathrm{d}\tau \right)^{\frac{1}{\kappa-1}} \leq \frac{1}{c_{int} c_{\mu}}
\]

The second integral \( \left( \int_1^{t+\gamma} \left( \frac{\|p(\tau)\| + \|d^h(\tau)\|}{\widehat{\theta}(\tau)} \right)^{\frac{1}{\lambda}} \mathrm{d}\tau \right)^{\frac{1}{\lambda-1}} \) can be estimated similarly. Again using the definition of \( \widehat{\theta} \) and the estimates stated in the remark yields
\[
\frac{\|p\|}{\widehat{\theta}} + \frac{\|d^h\|}{\widehat{\theta}} \leq \frac{1}{c_{int}} \frac{\|p\|}{\mu \|d^h\|} \leq C_P^\lambda \left\| \frac{\|p\|}{\mu \|d^h\|} \right\|^{\frac{1}{\lambda}} = C_P^\lambda \frac{c_{\lambda}}{c_{int}} \|p\|^{-\frac{1}{\lambda}} \mathrm{d}\tau.
\]

Therewith the integral is bounded, using proposition 3.3
\[
\left( \int_1^{t+\gamma} \left( \frac{\|p(\tau)\| + \|d^h(\tau)\|}{\widehat{\theta}(\tau)} \right)^{\frac{1}{\lambda}} \mathrm{d}\tau \right)^{\frac{1}{\lambda-1}} \leq \left( \int_1^{t+\gamma_0} \frac{1}{c_{int} c_{\mu}} \mathrm{d}\tau \right)^{\frac{1}{\lambda-1}} \leq \text{const}
\]
provided \( \lambda \geq \frac{\kappa}{\rho (\kappa - 1)} \).

A possible choice for \( \kappa \) and \( \lambda \) is to take \( \lambda = \frac{-\kappa}{(\kappa - 1)\rho} \) and \( \kappa = \max\{1 + \frac{1}{\rho}, 2\} \), which arises from assumption 3.4.

Practical verifiability of these link conditions depends on the type of PDE and observations. E.g., in case of observations on a subdomain, infinite speed of propagation (as valid for certain parabolic equations) is expected to enable such an instantaneous link between ‘observed’ and ‘unobserved’ part of the state. On the other hand, in view of absence of \textit{a priori} information on this part of the state, some kind of link condition to the ‘observed’ seems to be indispensable.

With these assumptions we can state the next lemma.

**Lemma 3.11.** Let assumptions 2.1 with (4), 3.2, 3.4, 3.7, 3.9, and (18) hold.

Then, for any given \( \gamma > \frac{2\gamma_0}{\|p\|_{L_\infty(x)}} \sup_{t \geq 0} \|d^h(t)\|_x \), there are \( \varepsilon > 0 \), \( T > 0 \) and \( T_1 > 0 \) such that for all \( t_1 \geq T_1 \) the following holds true:

If the parameter error \( \|e(t_1)\|_Q \geq \gamma \), then there exists a \( t_2 \in [t_1, t_1 + T] \) such that the state error \( \|r^h(t_2)\|_X \geq \varepsilon \).

**Proof.** We choose \( T_0, \varepsilon_0, \gamma_0, T > 0 \) according to assumption 3.7, fix \( \gamma > 0 \) arbitrarily, set \( T_1 = \max \{T, T_1\} \) and assume that \( t_1 > T_1 \) and \( \|e(t_1)\|_Q > \gamma \). (Here \( T \) will be chosen sufficiently large below.) Setting \( \xi = \frac{\gamma(t_0)}{\|e(t_0)\|_Q} \) we can choose \( t_b \) according to assumption 3.7. Now we use lemma 3.6 with \( t_a = t_1, t_c = t_b + \gamma_0 \), and set \( t_2 = t_c \) and \( T = T_0 + \gamma_0 \) (i.e.
\[ l_0 = t_1 \leq t_2 \leq t_3 \leq t_4 \leq t_5 + T \to \text{obtain} \]
\[
\| r_\psi(t_2) \|_X = \| r_\psi(t_0 + \gamma_0) \|_X \geq \int_{t_0}^{t_0 + \gamma_0} R A (u^a(\tau) + p(\tau) + d^\psi(\tau)) \frac{e(t_0)}{\| e(t_0) \|_Q} d\tau \| e(t_0) \|_Q \\
- \| r_\psi(t_0) \|_X - \int_{t_0}^{t_0 + \gamma_0} \| d^\psi(\tau) \|_X d\tau - L_A^2 \int_{t_0}^{t_0 + \gamma_0} \int_{t_0}^{t_0 + \gamma_0} \| r_\psi(\sigma) \|_X d\sigma d\tau \\
- L_c \int_{t_0}^{t_0 + \gamma_0} (|p(\tau)|\varphi + |d^\psi(\tau)|\varphi) d\tau - C_M \int_{t_0}^{t_0 + \gamma_0} \mu(\tau) d\tau \\
\geq \varepsilon_0 \| e(t_0) \|_Q - \| r_\psi(t_0) \|_X - \int_{t_0}^{t_0 + \gamma_0} \| d^\psi(\tau) \|_X d\tau - L_A^2 \int_{t_0}^{t_0 + \gamma_0} \int_{t_0}^{t_0 + \gamma_0} \| r(\sigma) \|_X d\sigma d\tau \\
- L_c \int_{t_0}^{t_0 + \gamma_0} (|p(\tau)|\varphi + |d^\psi(\tau)|\varphi) d\tau - C_M \int_{t_0}^{t_0 + \gamma_0} \mu(\tau) d\tau.
\]

The last three terms remain to be estimated.
\[
\int_{t_0}^{t_0 + \gamma_0} \int_{t_0}^{t_0 + \gamma_0} \| r_\psi(\sigma) \|_X d\sigma d\tau \leq \int_{t_0}^{t_0 + \gamma_0} \int_{t_0}^{t_0 + \gamma_0} \| r_\psi(\sigma) \|_X d\sigma d\tau \leq \gamma_0 T \sup | r_\psi(\sigma) |_X .
\]

Estimating by Hölder’s inequality and using the link conditions results in
\[
\int_{t_0}^{t_0 + \gamma_0} \| p(\tau) \|_Q + \| d^\psi(\tau) \|_Q d\tau = \int_{t_0}^{t_0 + \gamma_0} \| p(\tau) \|_Q + \| d^\psi(\tau) \|_Q d\tau \leq \left( \int_{t_0}^{t_0 + \gamma_0} \theta(\tau) d\tau \right)^{\frac{1}{2}}
\]
and analogously for the last term
\[
\int_{t_0}^{t_0 + \gamma_0} \mu(\tau) d\tau = \int_{t_0}^{t_0 + \gamma_0} \frac{\mu(\tau)}{\varphi \theta(\tau)^2} \theta^2(\tau) d\tau \leq C_i \left( \int_{t_0}^{t_0 + \gamma_0} \theta(\tau) d\tau \right)^{\frac{1}{2}}.
\]

Now using (37) we can estimate the term \( \int_{t_0}^{t_0 + \gamma_0} \theta(\tau) d\tau \) as follows
\[
\int_{t_0}^{t_0 + \gamma_0} \theta(\tau) d\tau \leq - \frac{1}{c_M} \left[ | e(t_0) |_Q^2 + \| r_\psi(\tau) \|_X^2 \right]_{t_0}^{t_0 + \gamma_0} \\
\leq \frac{1}{c_M} \left[ | e(t_0) |_Q^2 - | e(t_0 + \gamma_0) |_Q^2 \right] + \frac{1}{c_M} | r_\psi(t_0) |_X^2.
\]

At this point we utilise (24), (32), assumption 2.1 as well as proposition 3.3
\[
| e(t_0) |_Q^2 - | e(t_0 + \gamma_0) |_Q^2 = | e(t) |_Q^2 | e(t) |_Q^2 = -2 \int_{t_0}^{t_0 + \gamma_0} (e, e)_Q d\tau \\
\leq 2 \int_{t_0}^{t_0 + \gamma_0} \| e \|_Q \| e \|_Q d\tau = 2 \int_{t_0}^{t_0 + \gamma_0} A (u^a + p + d^\psi)^2 \| e \|_Q d\tau \\
\leq 2 T \int_{t_0}^{t_0 + \gamma_0} \| r_\psi(\tau) \|_X d\tau \leq 2 L_A \int_{t_0}^{t_0 + \gamma_0} \| r_\psi(\tau) \|_X d\tau \sqrt{ \| e(0) \|_Q^2 + \| r_\psi(0) \|_X^2 }.\]
Hence altogether we end up with
\[
\|r_o^e(t_2)\|_X \geq \varepsilon_0 \gamma - \|r_o^e(t_0)\|_X - \int_{t_0}^{t_0 + \gamma_0} \|d_{\alpha}^e\|_X \, d\tau - L_1^2 \gamma_0 T \sup_{\sigma \geq t_1} \|r_o^e(\sigma)\|_X \\
- L_1 C_1 \left( \frac{2L_1}{C_M} \int_{t_0}^{t_0 + \gamma_0} \|r_o^e\|_X \, d\tau \sqrt{\|e(0)\|_Q^2 + \|r_o^e(0)\|_X^2} + \frac{1}{C_M} \|r_o^e(t_0)\|_X^2 \right)^{\frac{1}{2}} - C_M C_1 \left( \frac{2L_1}{C_M} \int_{t_0}^{t_0 + \gamma_0} \|r_o^e\|_X \, d\tau \sqrt{\|e(0)\|_Q^2 + \|r_o^e(0)\|_X^2} + \frac{1}{C_M} \|r_o^e(t_0)\|_X^2 \right)^{\frac{1}{2}}. 
\]

By theorem 3.5 for \( T \) sufficiently large, \( t_0, t_1, T_1 \geq T \) the sum of all negative terms on the right-hand side will be contained in the interval \( \left[ -\frac{\varepsilon_0 \gamma}{2} - \gamma_0 \sup_{t>0} \|d_{\alpha}^e(t)\|_X, 0 \right] \), so that we get
\[
\|r_o^e(t_2)\|_X \geq \varepsilon_0 \gamma - \frac{\varepsilon_0 \gamma}{2} - \gamma_0 \sup_{t>0} \|d_{\alpha}^e(t)\|_X = \frac{\varepsilon_0 \gamma}{2} - \gamma_0 \sup_{t>0} \|d_{\alpha}^e(t)\|_X =: \varepsilon > 0.
\]

This implies the assertion. \( \square \)

**Theorem 3.12**: (Parameter convergence). Under assumptions 2.1 with (4), 3.2, 3.4, 3.7, 3.9 and (18) we have that
\[
\lim_{t \to \infty} \sup_{t>0} \|\hat{q}(t) - q^*\|_Q \leq \frac{2\gamma_0}{\varepsilon_0} \sup_{t>0} \|d_{\alpha}^e(t)\|_X,
\]
i.e., in the exact data case \( \varepsilon^b = z \), we have convergence \( \hat{q}(t) \to q^* \) as \( t \to \infty \).

**Proof.** By contraposition in lemma 3.11, for all \( \gamma > \frac{2\gamma_0}{\varepsilon_0} \sup_{t>0} \|d_{\alpha}^e(t)\|_X \) there exists \( \varepsilon > 0 \), \( T > 0 \), \( T_1 > 0 \) such that for all \( t_1 \geq T_1 \) and for all \( t_2 \in [T_1, t_1 + T] \): \( \|r_o^e(t_2)\|_X < \varepsilon \) implies \( \|e(t)\|_Q < \gamma \).

So for given \( \gamma > \frac{2\gamma_0}{\varepsilon_0} \sup_{t>0} \|d_{\alpha}^e(t)\|_X \) we choose \( \varepsilon \) and \( T_1 > 0 \) according to lemma 3.11. Then from theorem 3.5 it follows that there exists \( t_2 \geq T_1 \) such that for all \( t \geq t_2 \) we have \( \|r_o^e(t_2)\|_X < \varepsilon \), hence by the above \( \|e(t)\|_Q < \gamma \) for all \( t_1 \geq T_1 \). Since \( \gamma \) can be chosen arbitarily close to \( \frac{2\gamma_0}{\varepsilon_0} \sup_{t>0} \|d_{\alpha}^e(t)\|_X \) the assertion follows. \( \square \)

### 3.2. Convergence with arbitrary noisy data

We now consider the situation of possibly non-smooth(ed) noise, so that \( \hat{d}_{\alpha}^e \) according to (23) is not well-defined and hence we cannot use the representation (25). Additionally to regularising \( G^T \) one might add a stabilising term defined by another parameter \( \sigma = \sigma(t) \geq 0 \), see e.g. [6]. Note that also the case \( \sigma \equiv 0 \) is included in our analysis. As a matter of fact, it turns out that this term is not really needed. For the sake of completeness to some extent we will also consider the case of strictly positive \( \sigma \). The case of partially vanishing, partially positive \( \sigma \) is not included here (but could be approximated by some positive \( \sigma \) which partially gets arbitrarily small). Therewith, we redefine the estimators \( \hat{q}, \hat{u} \) by
\[
\hat{q}_t - A(u^\alpha + P\hat{u}^\alpha)(R\hat{u} - \hat{u}^\alpha) = -\sigma \hat{q},
\]
\[\hat{u}_t + C(\hat{q}, u^\delta + P\hat{u}) + \mu RM \frac{R\hat{u} - u^\delta}{\|R\hat{u} - u^\delta\|_\psi} + \nu PNP\hat{u} = f, \quad (40)\]

\[(\hat{q}, \hat{u})(0) = (\hat{q}_0, \hat{u}_0). \quad (41)\]

Then the differential equations for the errors are

\[e_t - A(a^\delta + p + d^\delta)\sigma e = -\sigma \hat{q}, \quad (42)\]

\[r_t + RC(q^* - u^\delta + p + d^\delta) - RC(q^*, u^\delta) + RA(a^\delta + p + d^\delta)e + \mu RM \frac{r^\delta}{\|r^\delta\|_\psi} = 0, \quad (43)\]

\[p_t + PC(q^* - u^\delta + p + d^\delta) - PC(q^*, u^\delta) + PA(a^\delta + p + d^\delta)e + \nu PNP\hat{u} = 0, \quad (44)\]

\[(e, r, p)(0) = (\hat{q}_0 - q^*, R(\hat{u}_0 - u_0), P(\hat{u}_0 - u_0)). \quad (45)\]

### 3.2.1. Well-definedness

For showing well-definedness we again take a look at the error components \(e = \hat{q} - q^*, \ r = R\hat{u} - Ru^\delta, \ r^\delta = R\hat{u} - u^\delta, \ p = P\hat{u} - Pu^\delta, \) (see identities (16), (17)).

As in the previous section some assumptions concerning the parameters \(\mu\) and \(\nu\) are required to prove well-posedness of the system. For \(\nu\) we remain with assumption 3.2, whereas the condition for \(\mu\) slightly changes to

**Assumption 3.13.** For all \(t > 0\)

\[\begin{align*}
\mu(t) \geq & \max \left\{ \frac{4Lc}{c_M} (\|d^\delta(t)\|_\psi + \|p(t)\|_\psi) \|r(t)\|_X \\
+ & \frac{4CA}{c_M} (1 + \|d^\delta(t)\|_\psi + \|p(t)\|_\psi) \|e(t)\|_0 \|d^\delta(t)\|_X, \quad \frac{2\sigma(t)}{c_M} \|r(t)\|_X^2 \right\} \|r(t)\|_X^2/\|d^\delta(t)\|_\psi. \end{align*} \]

A condition on the error between the regularised version of the ‘observed’ part and the exact state is also needed, namely for all considered time instances \(t\)

\[\|d^\delta(t)\|_\psi \leq \frac{c_M}{2CM} \|r(t)\|_X^2/\|d^\delta(t)\|_X^2 \quad (46)\]

should hold. This condition on smallness can be further accessed using the fact that \(d^\delta = G\delta - G^T z\) and (15), based on results of regularisation theory and an appropriate choice of \(\alpha(t)\) in dependence of \(\delta(t)\) and \(z^\delta(t)\), see, e.g. [3]. We now prove that \(\hat{q}\) and \(\hat{u}\) according to (39) and (40) are well defined at least up to a certain time.

**Proposition 3.14.** Let assumptions 2.1 with (5), 3.2 and 3.13 hold and let \((\hat{q}_0 - q^*, \hat{u}_0 - u_0) \in Q \times (\bar{V} + \bar{V})\). Then there exists a solution \((\hat{q}(t), \hat{u}(t)) \in Q \times (\bar{V} + \bar{V})\) of (39)–(41) for all times \(0 < t < T^*\) where
\[ T^* = \min \left\{ t > 0 : \| d^t(t) \|_p > \frac{c_M}{2CM} \frac{\| r(t) \|_{2,\infty}^2}{\| r(t) \|_\infty} \right\} \] (47)

(i.e. the first time, when condition (46) is violated) and satisfies the following error bounds.

(i) Case \( \sigma \equiv 0 \): for all \( 0 < t < T^* \) : \( \| e(t) \|_Q^2 + \| r(t) \|_{\infty}^2 \leq \| e(0) \|_Q^2 + \| r(0) \|_{\infty}^2 \);

Case : \( \sigma > 0 \) : for all \( 0 < t < T^* \) : \( \| e(t) \|_Q^2 + \| r(t) \|_{\infty}^2 \leq \max \{ \| q^* \|_Q^2, \| e(0) \|_Q^2 + \| r(0) \|_{\infty}^2 \} \);

(ii) For all \( 0 < t < T^* \): \( \| p(t) \|_X \leq \max \left\{ \| p(0) \|_X, \frac{2c^2}{c^2 + \| e(0) \|_Q^2 + \| r(0) \|_{\infty}^2} \right\} \sup_{t > 0} \| Pu^*(t) \|_Y^2 \} \);

(iii) If \( T^*_C = \infty \) (see (16)) and \( \sigma \equiv 0 \) then \[ \int_0^T (2L_C(\| p(\tau) \|_Y + \| d^\alpha(\tau) \|_Y)) \| r(\tau) \|_X \, d\tau \leq \| e(0) \|_Q^2 + \| r(0) \|_{\infty}^2 < \infty. \]

Proof. For proving the proposition, like in the previous section we take a look at the norms of the squared errors

\[
\frac{d}{dt} \left[ \| e \|_Q^2 + \| r \|_{\infty}^2 \right] = (e, e)_Q + (n, r)_X
\]

\[
= (A(u^* + p + d^\alpha)^\mu, e)_Q - (RA(u^* + p + d^\alpha), e)_X
\]

\[- (\sigma \dot{q}, e)_Q + (RC(q^*, u^*) - RC(q^*, u^* + p + d^\alpha), r)_X - \mu \left( \frac{RM \dot{r}^\alpha}{\| r \|_X}, r \right)_X
\]

\[
\leq C_A (1 + \| d^\alpha \|_X + \| p \|_Y) \| e \|_Q \| d^\alpha \|_X + \frac{\sigma}{2} \| q^* \|_Q^2 - \frac{\sigma}{2} \| e \|_Q^2
\]

\[+ L_C(\| d^\alpha \|_Y + \| p \|_Y) \| r \|_X - \mu \left( \frac{\| r \|_X^2}{\| r \|_X^2}, \frac{\| r \|_X^2}{\| r \|_X^2} \right), \]

(48)

where we have used the identity \( r = r^\alpha + d^\alpha \), assumption 2.1, the identity

\[- (\sigma \dot{q}, e)_Q = -\sigma (\dot{q} \pm q^*, e)_Q = \sigma (q^*, e)_Q - \sigma (e, e)_Q
\]

\[\leq -\sigma \| e \|_Q^2 + \sigma \| q^* \|_Q \| e \|_Q \leq -\sigma \| e \|_Q^2 + \frac{\sigma}{2} (\| q^* \|_Q^2 + \| e \|_Q^2) = \frac{\sigma}{2} \| q^* \|_Q^2 - \frac{\sigma}{2} \| e \|_Q^2, \]

and the fact that coercivity and boundedness of \( M \) (assumption 2.1) and \( t \leq T^* \) with \( T^* \) as in (47) results in

\[- \mu \left( \frac{RM \dot{r}^\alpha}{\| r \|_X}, r \right)_X = - \frac{\mu}{\| r \|_X^2} (RM \dot{r}^\alpha, r)_X + \frac{\mu}{\| r \|_X^2} (RMd^\alpha, r)_X
\]

\[\leq - \mu c_M \| e \|_X^2 + \mu c_M \| d^\alpha \|_X \| r \|_X \leq - \mu \left( \frac{\| r \|_X^2}{\| r \|_X^2}, \frac{\| r \|_X^2}{\| r \|_X^2} \right). \]

(49)

Using assumption 3.13 on \( \mu \) we get

\[
\frac{d}{dt} \left[ \| e \|_Q^2 + \| r \|_{\infty}^2 \right] \leq \frac{\sigma}{2} \| q^* \|_Q^2 - \frac{\sigma}{2} \| e \|_Q^2 - \frac{\mu c_M \| r \|_X^2}{4 \| r \|_X^2}, \]

\[
\leq \frac{\sigma}{2} (\| q^* \|_Q^2 - (\| e \|_Q^2 + \| r \|_{\infty}^2)). \]

(50)

(51)

Now we distinguish between the two cases \( \sigma \equiv 0 \) and \( \sigma > 0 \).
In the first case $\sigma = 0$ we have
\[
\frac{d}{dt} \frac{1}{2} \|e(t)\|_Q^2 + \|r(t)\|_X^2 \leq 0.
\]
For the second case $\sigma > 0$ we define $\tau(t) := \int_0^t \sigma(\xi) \, d\xi$, $V(t) := \frac{1}{2}[\|e(t)\|_Q^2 + \|r(t)\|_X^2]$ and $\nabla(\tau(t)) := V(t)$. Differentiating $V$ with respect to $\tau$ leads to
\[
\frac{d}{d\tau} V(\tau(t)) = \frac{1}{2}[\|e(\xi)\|_Q^2 + \|r(\xi)\|_X^2] - \frac{1}{\sigma(t)} \leq \frac{1}{2}\|q^+\|_Q^2 - \frac{1}{2}[\|e(0)\|_Q^2 + \|r(0)\|_X^2] = \frac{1}{2}\|q^+\|_Q^2 - V(\tau(t)).
\]
So we have for all $t > 0$
\[
\frac{1}{2}[\|e(t)\|_Q^2 + \|r(t)\|_X^2] \leq \max \left\{ \frac{1}{2}[\|q^+\|_Q^2], \frac{1}{2}[\|e(0)\|_Q^2 + \|r(0)\|_X^2] \right\} < \infty.
\]

The proof of (ii) is exactly the same as the one of the corresponding part of proposition 3.3.

In case $T^* = \infty$ and $\sigma \equiv 0$, (iii) is a consequence of inequality (50) and assumption 3.13.

3.2.2. State convergence. As in the previous section we introduce an additional lower bound on $\mu$ for proving convergence of the ‘observed’ part of the state estimate.

Assumption 3.15. There exists a constant $\tilde{c}_i > 0$ such that for all $t > 0$
\[
\mu(t) \geq \max \left\{ \frac{4L_C}{c_M} \|d_i(u(t))\|_X + \|p(t)\|_V \|r(t)\|_X \right. \\
+ \left. \frac{4C_M}{c_M} (1 + \|d_i(u(t))\|_X + \|p(t)\|_V) \|e(t)\|_Q \|d_i(u(t))\|_X, \tilde{c}_i \|r(t)\|_X^2 \right\} \frac{\|r_i(u(t))\|_V}{\|r(t)\|_X^2}.
\]

Theorem 3.16. (State convergence) Under assumptions 2.1 with (4), 3.2, 3.15 and (18) and if $T^* = \infty$ and $\sigma \equiv 0$ we have that $\|R(u(t) - u^*(t))\|_X = \|r(t)\|_X \to 0$ as $t \to \infty$.

Proof. The proof is quite similar to the one of theorem 3.5. In place of (31) and (33) we get
\[
\|r(t_2)\|_X^2 - \|r(t_1)\|_X^2 = \int_{t_1}^{t_2} (\xi, r)_X \\
= \int_{0}^{t_2} (R_C(q^+, u^p) - R_C(q^+, u^* + d_i^p + p) \\
- RA(u^p + d_i^p)e - \frac{\mu}{\|r_i(u\|_V, r)_X d\tau
\]

18
with $\tilde{c}_2 := \frac{L_a}{2}(\|e(0)\|_X^2 + \|r(0)\|_X^2)$, which follows from proposition 3.14. As in the previous section (see (34)) we get for any fixed $t > \gamma > 0$

$$\int_{t-\gamma}^{t} \|r(\tau)\|_X^2 d\tau \geq \gamma \|r(\tau)\|_X^2 - \frac{\tilde{c}_2}{2} \gamma^2.$$ 

For $\sigma \equiv 0$ the proof from now on is exactly the same as the one of theorem 3.5. □

**Remark 3.17.** If $T^* < \infty$ we cannot expect convergence of the state error to zero if $\delta > 0$. However in this case the definition (47) of $T^*$ implies

$$\|d_\delta^*(T^*)\|_X > \frac{c_M}{2C_M \|r(T^*)\|_X}$$

and therefore that $r(T^*)$ is small, namely in case the interpolation inequality (38) holds we even have that at time $T^*$ the ‘observed’ state error is (up to a constant factor $\frac{c_M}{2C_M}$) as small as the error in the ‘observed’ state, both of them in the $V$-norm.

### 3.2.3. Parameter convergence

Like lemma 3.6 one obtains

**Lemma 3.18.** Under assumption 2.1 with (4) and (18) the projected state errors $r$ and $p$ satisfy the following relation for all $0 < t_0 \leq t_0 \leq t_1$:

$$\|r(t)\|_X \geq \int_{t_0}^{t_1} RA(u^0(\tau)) + p(\tau) + d_\delta^*(\tau) \mu(t_0)d\tau_0$$

$$- \|r(t_0)\|_X - L_a \int_{t_0}^{t_1} \int_{t_0}^{\tau} \sigma \|q(s)\|_Q d\sigma d\tau$$

$$- L_a \int_{t_0}^{t_1} \int_{t_0}^{\tau} \|q(s)\|_X d\sigma d\tau - L_c \int_{t_0}^{t_1} \||\| d_\delta^*(\tau)\|_Y + ||p(\tau)||_Y d\tau - C_M \int_{t_0}^{t_1} \mu(\tau) d\tau.$$ 

The persistence of excitation and link conditions are exactly as the in previous section. Furtheron we just consider the case $\sigma = 0$. In the other case $\sigma > 0$ we cannot prove parameter convergence. The second lemma that is needed for parameter convergence is exactly the same as in the exact data case. (see lemma 3.11).

**Lemma 3.19.** Let assumptions 2.1 with (4), 3.2, 3.15, 3.7, 3.9 and (18) hold and $\sigma \equiv 0$. Then, for any given $\gamma > 0$, there are $\varepsilon > 0$, $T > 0$ and $T_i > 0$ such that for all $t_i \geq T_i$ the following holds true:
If the parameter error \( ||e(t_1)||_Q \geq \gamma \), then there exists a \( t_2 \in [t_1, t_1 + T] \) such that the state error \( ||r(t_2)||_X \geq \varepsilon \).

**Proof.** In case \( \sigma \equiv 0 \) Lemma 3.18 with assumption 3.7 gives the same estimate as in the exact and smoothed noisy data case with the only difference that we can replace \( \tilde{a}_n \) by zero. Thus the proof obviously goes through like the one of Lemma 3.11. \( \square \)

For \( T^* = \infty \) (see (47), which by proposition 3.14 implies (18)) we can prove parameter convergence analogously to theorem 3.12.

**Theorem 3.20** (Parameter convergence). Under assumptions 2.1 with (4), 3.2, 3.15, 3.7, 3.9 and if \( T^* = \infty, \sigma \equiv 0 \) we have that
\[
\| \hat{q}(t) - q^* \|_Q \to 0 \text{ as } t \to \infty.
\]

**Remark 3.21.** In case \( T^* < \infty \) we cannot prove parameter convergence, because in the persistence of excitation assumption we need to have \( t \to \infty \).

4. **Application and numerical tests**

4.1. **Optimal choice of \( \mu, \nu \)**

Assumptions 3.1, 3.2, 3.4, 3.13, 3.15 contain technical assumptions on the parameters \( \mu \) and \( \nu \) that we need in our analysis to guarantee convergence. The practical use of these conditions for actually choosing \( \mu \) and \( \nu \) at each time instance is questionable, since they contain the unknown error terms \( p \) and \( r \) and heavily depend on the choice of certain constants such as \( \nu \) and \( c_t \). Alternatively, a practically more feasible approach to choose \( \mu \) and \( \nu \) is an optimisation approach, in which, at each time step, we minimise some cost functional \( J \), e.g., a regularised variant of the data misfit, with respect to the values of \( \mu \) and \( \nu \) at the new time step. Implementation of such an approach depends on the choice of the time stepping scheme applied to the dynamical system (12)–(14). We exemplarily demonstrate it for the simple setting of an implicit Euler method with time step size \( h \). Thus, starting from the values \( \tilde{q}_n, \tilde{u}_n \) of \( \tilde{q}, \tilde{u} \) at some current time instance \( t = t_n \) and given the regularised observations \( u^*_{n+1} = u^*_n(t_{n+1}) \), excitation \( f = f(t_{n+1}) \), and data \( z^* = z(t_{n+1}) \), the problem of determining the values \( \hat{q}^+, \hat{u}^+, \hat{\mu}, \hat{\nu} \) of \( \hat{q}, \hat{u}, \hat{\mu} := \frac{1}{2} \mu \hat{u}_0 - \frac{1}{4} \mu \hat{u}_0^2 \) and \( \nu \) at time \( t_{n+1} = t_n + h \) amounts to solving

\[
\min_{(\hat{q}^+, \hat{u}^+, \hat{\mu}, \hat{\nu}) \in Q \times V \times R^2} J(\hat{q}^+, \hat{u}^+)
\]

(e.g., \( J(\hat{q}^+, \hat{u}^+) = \frac{1}{2} \|G\hat{q}^+ - z^+\|^2_Q + \frac{\mu}{2} \|\hat{q}\|^2_Q \)) with some regularisation parameter \( \beta > 0 \), under the equality constraints
\[
\hat{q}^+ = \hat{q}_n - hA(u^*_n + \tilde{P}\hat{u}^+)(R\hat{u}^+ - u^*_n) = 0
\]
\[
\hat{u}^+ = \hat{u}_n + hC(\tilde{q}_n, u^*_n + \tilde{P}\hat{u}^+) + \tilde{\mu}\hat{u}^+ hM(R\hat{u}^+ - u^*_n) + \nu^i hP\hat{u}^+ = hf^*.
\]
The Lagrange functional reads as
\[
L(q^+, \hat{u}^+, \tilde{\mu}^+, \nu^+, o, w) = J(q^+, \hat{u}^+) + (\tilde{\mu}^+ - \hat{\mu}_n, o)
- h(A(u^+_n + \tilde{P}^+)w)(R\hat{u}^+ - u^+_n), o)_Q
+ (\hat{u}^+ - \tilde{u}_n + hC(q^+, u^+_n + \tilde{P}^+) + \tilde{\nu}^+ hRM(R\hat{u}^+ - u^+_n)
+ \tilde{\nu}^+ hPNP\tilde{\nu}^+ - h_f, w)_X
\]
for \((q^+, \hat{u}^+, \tilde{\mu}^+, \nu^+, o, w) \in Q \times X \times \mathbb{R}^2 \times Q \times X\), where
\[
((A(u^+_n + \tilde{P}^+)w)(R\hat{u}^+ - u^+_n), o)_Q = (R\hat{u}^+ - u^+_n, A(u^+_n + \tilde{P}^+)w)_Q
\]
and
\[
(C(q^+, u^+_n + \tilde{P}^+), w)_X = (B(u^+_n + \tilde{P}^+), w)_X + (q^+, A(u^+_n + \tilde{P}^+)w)_Q.
\]
Thus, by differentiation with respect to all variables, the first order optimality conditions for this minimisation problem can be easily derived as
\[
J_q(q^+, \hat{u}^+) + o + hA(u^+_n + \tilde{P}^+)w = 0, \quad (52)
\]
\[
J_u(q^+, \hat{u}^+) + w + hR(-A(u^+_n + \tilde{P}^+)o + \tilde{\mu}^+ M\tilde{R}w)
+ hP(-D(u^+_n + \tilde{P}w)(o, R\hat{u}^+ - u^+_n) + C_u(q^+, u^+_n + \tilde{P}^+\nu w + \nu^\nu N^\nu Pw) = 0,
\]
\[
w \perp RM(R\hat{u}^+ - u^+_n), \quad w \perp PNP\tilde{\nu}^+, \quad (53)
\]
\[
\tilde{\nu}^+ - \hat{\mu}_n - hA(u^+_n + \tilde{P}^+)w(R\hat{u}^+ - u^+_n) = 0, \quad (54)
\]
\[
\hat{\nu}^+ - \hat{u}_n + hC(q^+, u^+_n + \tilde{P}^+) + \tilde{\nu}^+ hRM(R\hat{u}^+ - u^+_n) + \nu^\nu hPNP\tilde{\nu}^+ = h_f, \quad (55)
\]
where by Riesz’ Representation theorem we identify the derivatives of \(J\) with respect to \(q\) and \(u\) with elements in \(Q\) and \(X\), respectively, the used adjoints are just defined as Hilbert space adjoints, except for \((\text{recall } \tilde{M} : \tilde{V} \rightarrow X, N : \tilde{V} \rightarrow X)\)
\[
(M\tilde{v}_1, \tilde{v}_2)_X = (\tilde{v}_1, M^*\tilde{v}_2)_X, \quad (N\tilde{v}_1, \tilde{v}_2)_X
= (\tilde{v}_1, N^*\tilde{v}_1)_X, \quad \forall \tilde{v}_1 \in X, \tilde{v}_1 \in \tilde{V}, \tilde{v}_1 \in \tilde{V}
\]
and for any \(u \in X\), the bilinear operator \(D(u) : Q \times X \rightarrow X\) is defined by
\[
(D(u)(q_1, x_1), x_2)_X = (x_1, A'(u)x_2)_X \quad \forall q_1 \in Q, x_1, x_2 \in X.
\]
This is a system of coupled nonlinear equations in the space \(Q \times X \times \mathbb{R}^2 \times Q \times X\), which can, e.g. be solved by a fixed point iteration. To this end, we take the inner product of \((53)\) with \(Rw\) and \(Pw\), respectively, to obtain expressions for \(\tilde{\nu}^+\) and \(\nu^\nu\)
\[
\tilde{\nu}^+ = \frac{-(J_u(q^+, \hat{u}^+) + w + h(-A(u^+_n + \tilde{P}^+)o), Rw)_X}{h(M\tilde{R}w, Rw)_X}, \quad (57)
\]
\[
\nu^\nu = \frac{-(J_u(q^+, \hat{u}^+) + w + h(-D(u^+_n + \tilde{P}w)(o, R\hat{u}^+ - u^+_n) + C_u(q^+, u^+_n + \tilde{P}^+\nu w), Pw)_X}{h(N\tilde{P}w, Pw)_X}, \quad (58)
\]
So we end up with the following algorithm to compute $\hat{q}^+ = \hat{q}(t_{n+1})$, $\hat{u}^+ = \hat{u}(t_{n+1})$, $\tilde{\mu}^+ = \tilde{\mu}(t_{n+1})$, $\nu^+ = \nu(t_{n+1})$.

Algorithm 1.

1: Set $\tilde{\mu}^+ = \tilde{\mu}(t_0)$, $\nu^0 = \nu(t_0)$.
2: while stopping criterion violated do
3: Solve state equations (55), (56) with fixed parameters $\tilde{\mu}^+\,\nu^+$, to obtain $\hat{q}^+\,\hat{u}^+$.
4: Solve adjoint equations (52), (53) in the space $\tilde{X}(\tilde{\mu}^+) := \text{span}\{RM(\tilde{\mu}^+ - \tilde{\mu})_n, PNP\tilde{\mu}^+\}$
   with fixed parameters $\tilde{\mu}^+\,\nu^+$ and states $\hat{q}^+, \hat{u}^+$, to obtain $\alpha, \omega$.
5: Compute updated parameters $\tilde{\mu}^{n+1}$, $\nu^{n+1}$ from (57), (58) with fixed primal and dual states $\hat{q}^+, \hat{u}^+, \alpha, \omega$
6: end while

Investigation of convergence of this algorithm as well as design of appropriate stopping rules will be subject of future research. So far in our implementation we just carry out one such iteration per time step to keep the numerical effort compatible with the online identification paradigm. More precisely, we only fully optimise $\tilde{\mu}(t_0)$, $\nu(t_0)$ before the beginning of the online computations. Then in each time step we just carry out one step of algorithm 1 to obtain $\tilde{\mu}^+, \nu^+$ and set $\tilde{\mu}(t_{n+1}) = \max\{\tilde{\mu}^+, \tilde{\mu}(t_{n+1})\}$, $\nu(t_{n+1}) = \max\{\nu^+, \nu(t_{n+1})\}$ to guarantee that $\tilde{\mu}, \nu$ stay bounded away from zero. (Note that only in case the $\tilde{\mu}$ value or the $\nu$ value has been reset to $\tilde{\mu}(t_0)$ or $\nu(t_0)$ by this procedure, (55), (56) has to be solved again to obtain the correct states at time $t_{n+1}$, otherwise we can just use the state values obtained from algorithm 1.)

4.2. Identification of a coefficient in a degenerate diffusion equation

Consider the problem of identifying $q = q(x)$ on a domain $\Omega \subseteq \mathbb{R}^d$ in the (possibly degenerate) parabolic initial boundary value problem

$$
\begin{align*}
&u_t(t,x) - \nabla \cdot (D(x)\nabla u(t,x)) + q(x)u(t,x) = f(t,x) \quad \text{in } \Omega \\
&u(t,x) = g(t,x) \quad \text{on } \partial \Omega \\
&u(0,x) = u_0(x)
\end{align*}
$$

from measurements of the state $u$ on a subdomain $\omega \subseteq \Omega$.

$$
\begin{align*}
z(t,x) = Gu(t,x) = u(t,x)|_\omega.
\end{align*}
$$

Here $f(t) \in L^2(\Omega)$, $g(t) \in H^2(\partial \Omega)$, $D \in L^\infty(\Omega)$, are assumed to be known and chosen such that for $q = q^*$ a solution $u(t) = u^*(t) \in H^2(\Omega)$ to (59) exists for all times $t > 0$.

With the spaces

$$
Q = H^s(\Omega), \quad X = L^2(\Omega), \quad Z = L^2(\omega),
$$

where $s > \frac{d}{2}$ so that $Q$ is continuously embedded in $L^\infty(\Omega)$, and the operators defined by

$$
\begin{align*}
&\quad C(q,u) = B(u) + A(u)q, \\
&(B(u),v)_X = \int_\Omega (D\nabla u)^T \nabla v \, dx, \quad A(u)q = qu, \quad Gu = u|_\omega,
\end{align*}
$$

this fits into the framework of the previous sections with an appropriate choice of the spaces $V, \hat{V}, \tilde{V}, \tilde{X}, X$ and the operators $M, N$, see below. Note that this formulation corresponds to the standard semigroup formulation for parabolic problems in case $D > 0$ (see, e.g., [4]).
However we do not assume $D$ to be positive, not even non-negative, hence the monotonicity assumption from [9] fails even if we use the setting there with the problem adapted spaces $V = \{ v \in L^2(\Omega) | \nabla v \in L^2(\Omega) \}$ with the norm \( ||v||_V = (\| \sqrt{D} \nabla v \|^2_{L^2(\Omega)} + \|v\|^2_{L^2(\Omega)})^{1/2} \), $H = L^2(\Omega)$ (with the notation $V$ and $H$ from [9]). The case $D < 0$, often denoted as antithetic diffusion, for example occurs in certain models of pattern formation, see e.g. [5].

For this problem, e.g., in [8, 14] efficient numerical identification of the parameter $q$ (in the latter paper even simultaneously with the initial temperature) has been studied. Here we aim at showing that also the online parameter identification method proposed and analysed in the previous sections is applicable, even in case of partial observations and antidiffusion.

We first of all define the spaces $\tilde{V}$, $\hat{V}$ such that the Lipschitz condition on $C$ from assumption 2.1 holds:

\[
\tilde{V} = \{ v \in L^2(\Omega) | \text{supp}(D\nabla v) \subseteq \omega, \ \text{supp} v \subseteq \omega \text{ and } \nabla \cdot (D\nabla v) \in L^2(\omega) \},
\]

\[
\hat{V} = \{ v \in L^2(\Omega) | \text{supp}(D\nabla v) \subseteq \Omega \setminus \omega, \ \text{supp} v \subseteq \Omega \setminus \omega \text{ and } \nabla \cdot (D\nabla v) \in L^2(\Omega \setminus \omega) \}
\]

with norms

\[
\|v\|_{\tilde{V}} = \|\nabla \cdot (D\nabla v)\|_{L^2(\omega)} + \|v\|_{L^2(\Omega)},
\]

\[
\|v\|_{\hat{V}} = \|\nabla \cdot (D\nabla v)\|_{L^2(\Omega \setminus \omega)} + \|v\|_{L^2(\Omega \setminus \omega)}.
\]

And their smooth counterparts

\[
\tilde{V}^\infty = \{ \phi \in C_0^\infty(\Omega) | \text{supp}(\phi) \text{ is a compact subset of } \omega \}
\]

\[
\hat{V}^\infty = \{ \psi \in C_0^\infty(\Omega) | \text{supp}(\psi) \text{ is a compact subset of } \Omega \setminus \omega \}
\]

which are dense in $\tilde{V}$ and $\hat{V}$, and whose sum $\tilde{V} + \hat{V}$ is dense in $L^2(\Omega)$. Therewith the Lipschitz condition on $C(q^*, \cdot)$ is obtained as follows. For any $\phi + \psi \in \tilde{V} + \hat{V}$ we have

\[
(C(q^*, u^\delta(t) + v + w) - C(q^*, u^\delta(t)), \phi + \psi)_{\Omega}
\]

\[
= \int_{\Omega} ((D\nabla (v + w))^T \nabla (\phi + \psi) + q^*(v + w)(\phi + \psi))dx
\]

\[
= \int_{\Omega} q^*(v + w)(\phi + \psi)dx + \int_{\Omega} (D\nabla v)^T \nabla \phi dx + \int_{\Omega \setminus \omega} (D\nabla w|_{\Omega \setminus \omega})^T \nabla \psi |_{\Omega \setminus \omega} dx
\]

\[
= \int_{\Omega} q^*(v + w)(\phi + \psi)dx - \int_{\Omega} \nabla \cdot (D\nabla v)\phi dx - \int_{\Omega \setminus \omega} \nabla \cdot (D\nabla w|_{\Omega \setminus \omega})\psi |_{\Omega \setminus \omega} dx
\]

\[
\leq \max \{1, \|q^*\|_{L^\infty(\Omega)} \} (\|v\|_V + \|w\|_V)\|\phi + \psi\|_{L^2(\Omega)}.
\]

where $\chi_\omega : L^2(\omega) \rightarrow L^2(\Omega), \chi_{\Omega \setminus \omega} : L^2(\Omega \setminus \omega) \rightarrow L^2(\Omega)$ denote the operators defined by the respective extension by zero to all of $\Omega$. 

23
The operator $A(u)$ can be estimated as follows: For all $v \in X$ we get
\[
\|A(u^* + v)\|_{Q \to X} = \|u^* + v\|_{Q \to X} = \sup_{q \in Q, q \neq 0} \frac{\|A(u^* + v)q\|_X}{\|q\|_Q} \leq \sup_{q \in Q, q \neq 0} \frac{\|u^* + v\|_{L^2(\Omega)} \|q\|_{L^\infty(\Omega)}}{\|q\|_{H^1(\Omega)}} \leq C \frac{\|u^*\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}}{\|q\|_{H^1(\Omega)}},
\]
which by continuity of the embeddings $\tilde{V} \hookrightarrow X$ and $\tilde{V} + \tilde{V} \hookrightarrow X$ implies (3) and (5).

The nullspace of $G$ and its orthogonal complement are given by
\[
\mathcal{N}(G) = \{w \in L^2(\Omega) \mid w|\partial \Omega = 0\} = \{w \in L^2(\Omega) \mid \text{supp}(w) \subseteq \Omega \setminus \omega\},
\]
\[
\mathcal{N}(G)^\perp = \{v \in L^2(\Omega) \mid v|\partial \Omega = 0\} = \{v \in L^2(\Omega) \mid \text{supp}(v) \subseteq \omega\},
\]
and the respective projections are defined by
\[
Ru = \chi_{\omega}[u|_{\partial \Omega}], \quad Pu = u - Ru = \chi_{\Omega \setminus \omega}[u|_{\partial \Omega}].
\]

We define the operators $M, N$ and the spaces $\tilde{V}, \tilde{X}$ as follows.
\[
(Mv, \phi)_X = \int_{\omega} |D| \nabla \phi|_{\partial \Omega} \cdot \nabla \phi \, dx + \int_{\Omega} v \phi \, dx,
\]
\[
(Nw, \psi)_X = \int_{\Omega \setminus \omega} |D| \nabla w|_{\partial \Omega} \cdot \nabla \psi|_{\partial \Omega} \, dx + \int_{\Omega} w \psi \, dx
\]
for $v \in \tilde{V}, \phi \in \tilde{V}^\infty, w \in \tilde{V}, \psi \in \tilde{V}^\infty$ (extending by density of the spaces $\tilde{V}^\infty, \tilde{V}^\infty$ in $\tilde{V}, \tilde{X}$, respectively).

Hence assuming that $D$ does not change its sign on $\omega$ and on $\Omega \setminus \omega$
\[
D \geq 0 \text{ a.e. on } \Sigma \text{ or } D \leq 0 \text{ a.e. on } \Sigma \text{ for } \Sigma \in \{\omega, \Omega \setminus \omega\}
\]
we get
\[
\|Mv\|_X = \sup_{\phi \in \tilde{V}^\infty, \phi \neq 0} \frac{(Mv, \phi)_X}{\|\phi\|_{L^2(\Omega)}} \leq \|v\|_X \quad \forall v \in \tilde{V}
\]
\[
\|Nw\|_X = \sup_{\psi \in \tilde{V}^\infty, \psi \neq 0} \frac{(Mw, \psi)_X}{\|\psi\|_{L^2(\Omega)}} \leq \|w\|_X \quad \forall w \in \tilde{V}
\]
and
\[
(Mv, v)_X = \|D| \nabla v|_{\partial \Omega}\|_{L^2(\omega)} + \|v\|_{L^2(\Omega)}^2 = \|v\|_{\tilde{V}_X}^2 \quad \forall v \in \tilde{V}
\]
\[
(Nw, w)_X = \|D| \nabla w|_{\partial \Omega}\|_{L^2(\Omega \setminus \omega)} + \|w\|_{L^2(\Omega \setminus \omega)}^2 = \|w\|_{\tilde{V}_{X}}^2 \quad \forall w \in \tilde{V}
\]
Since in this case $G$ has closed range we need not regularise in case of noisy data, i.e., we can set $\alpha = 0$:
\[
u^\delta = G^\delta z^\delta = \chi_{\omega}[z^\delta], \quad d^\delta = G_\alpha(z^\delta - z) = \chi_{\omega}[z^\delta - z] = \chi_{\omega}[z^\delta - u^\omega].
\]
Thus we have proven...
Proposition 4.1. Consider the setting (60)–(65) with $D \in L^\infty(\Omega)$ satisfying (66). Then assumption 2.1 is satisfied for any $q^* \in Q$. $(f, g, u_0) \in (\partial_t - C(q^*, \cdot, \cdot), \tau_{(0, (0, \infty))}, \tau_{(0, (0, \infty))})(H^1((0, \infty); L^2(\Omega)) \cap L^\infty((0, \infty); V + \dot{V})).$

Note that here the first part (i) of assumption 2.1 is enforced by an explicit solvability condition on the data $(f, g, u_0)$. Indeed, well-posedness of (59) for arbitrary $(f, g, u_0)$ cannot be expected to hold since in regions with negative diffusion, (59) exhibits the ill-posed behaviour of the backwards heat equation.

4.3. Implementation and tests

In our implementation we consider the one-dimensional case of the problem from section 4.2 with domain $\Omega = (0, 1)$. The right-hand side $f$ is given by

$$f(t, x) = \frac{1}{1 + t} \left( D \pi^2 - \frac{1}{1 + t} + q^*(x) \sin(\pi x) \right),$$

where the exact parameter $q^*$ is a quadratic polynomial, $q^* = 0.025x^2 - 0.025x$. For simplicity the diffusion coefficient is chosen to be constant, $D = 1$.

For solving the system of partial differential equations (12)–(14) we derive its variational formulation and discretise the spaces $Q$ and $X$ by cubic Hermite basis functions $\phi_j$ and $\psi_j$ for $j = 2, \ldots, N - 1$ on a uniform mesh $0 = x_1 < x_2 < \ldots < x_N = 1$, where $N = 31$ in our case.

$$\phi_j(x) = \begin{cases} -2 \left( \frac{x - x_{j-1}}{h} \right)^3 + 3 \left( \frac{x - x_{j-1}}{h} \right)^2 & \text{if } x \in (x_{j-1}, x_j) \\ 1 - 3 \left( \frac{x - x_j}{h} \right)^2 + 2 \left( \frac{x - x_j}{h} \right)^3 & \text{if } x \in (x_j, x_{j+1}) \\ 0 & \text{else} \end{cases}$$

$$\psi_j(x) = \begin{cases} h \left( \frac{x - x_{j-1}}{h} \right)^3 - h \left( \frac{x - x_{j-1}}{h} \right)^2 & \text{if } x \in (x_{j-1}, x_j) \\ h \left( \frac{x - x_j}{h} \right)^3 - 2h \left( \frac{x - x_j}{h} \right)^2 + h \left( \frac{x - x_j}{h} \right) & \text{if } x \in (x_j, x_{j+1}) \\ 0 & \text{else}. \end{cases}$$

The reason for using such high order spaces is the required regularity on the arguments of the operators $M, N$ according to (64), (65). After using these as ansatz and test functions in the variational formulation for space discretisation, we solve the resulting ODE System with an implicit Euler method with step size $h = 0.6$. The interesting cases are those with partial observations and noisy data.

To investigate the case of partial observations we restrict the data to the subinterval $\omega = (0.3, 0.87)$ of $\Omega$. The results for this case are shown in figure 1. On the left the estimated parameter $\hat{q}(t, \cdot)$ is plotted for several time instances between 0 and 60, starting with $\hat{q}(0, \cdot) = 0$. The estimated parameters are the lines with markers, whereas the straight line indicates the exact parameter. On the right, the estimated (lines with markers) and exact (straight lines) states for different time steps are displayed. Although we have just partial observations, the state is estimated quite well also in the ‘unobserved’ region $\Omega \setminus \omega$. One can see that also the estimated parameter gets close to the exact one, but it is shifted to the right, which is due to the fact, that data are just given on the nonsymmetric interval $\omega$. Note that we do not know whether the persistence of excitation condition is satisfied here, which is in fact hard to verify in general.
For the noisy data case we assumed to have data with Gaussian noise with different noise levels $1\%, 2.5\%, 5\%$. In this case of irregular noise, according to section 3.2, parameter and state convergence cannot be proven if $T^* < \infty$, so we expect closeness only for times satisfying condition (46). This can be seen in the numerical results as well, because the error is increasing from a certain time instance on, which corresponds to the semiconvergence phenomenon in regularisation. As one might expect, the time where the error starts to grow

Figure 1. The parameter estimate $\hat{q}(t, x)$ (left) and the state estimate $\hat{u}(t, x)$ (right) at times $t = 0, 6, 15, 30, 45, 60$.

Figure 2. Estimated parameter $\hat{q}(t, x)$ at different times and the parameter error $(\|e\|_Q^2)$ of proposition 3.14 for different noise levels $\delta = 1, 2.5, 5\%$. 

For the noisy data case we assumed to have data with Gaussian noise with different noise levels $\delta = 1\%, 2.5\%, 5\%$. In this case of irregular noise, according to section 3.2, parameter and state convergence cannot be proven if $T^* < \infty$, so we expect closeness only for times satisfying condition (46). This can be seen in the numerical results as well, because the error is increasing from a certain time instance on, which corresponds to the semiconvergence phenomenon in regularisation. As one might expect, the time where the error starts to grow
again gets smaller as the noise level increases. In figure 2 the top row shows the estimated parameter for the three different noise levels for different times up to $t = 60$. In figure 2 the bottom row displays the errors of the estimated parameter ($\|e\|_Q$) as in proposition 3.14 developing over time. For small noise $\delta = 1\%$ the estimated parameter gets close to the exact one, and also the error decreases, as time proceeds, similar to the exact data case, whereas for larger noise $\delta = 2.5\%, 5\%$ the estimated parameter first gets close to the exact parameter up to a certain time instance and then it drifts away again, hence the error increases.

5. Conclusion and outlook

In this paper we have developed and analysed an online parameter identification method for time dependent problems. The main idea was to formulate an alternative dynamic update law for the state and an additional one for the parameter estimate. We showed that the solution of this alternative system of differential equations is well defined and that it converges to the exact parameter and state. The proofs were done for the case of exact data as well as for the case of noisy data and smooth noisy data. The main advantages of this method are, that it imposes less restrictions on the underlying model compared to existing methods and that it is also applicable in case of partial observations. In a numerical example we showed the performance of our online parameter identification method.

The results should be extended in several directions. First of all, we intend to consider more general nonlinearities $C$ in the model and general Banach spaces $X$ and $Q$. Another future goal is to consider time dependent parameters. This would mean that the model itself contains a dynamical update law for the parameter and therefore the online parameter identification method has to be adapted, so that this is taken into account.

Acknowledgments

We thank the reviewers for their valuable comments leading to an improved version of the manuscript. This work was supported by the Karl Popper Kolleg ‘Modeling-Simulation-Optimization’ funded by the Alpen-Adria-Universität Klagenfurt and by the Carinthian Economic Promotion Fund (KWF).

References

[1] Baumeister, J., Scondo, W., Demetriou, M. A. and Rosen, I. G. 1997 On-line parameter estimation for infinite-dimensional dynamical systems SIAM J. Control Optim., 35 678–713
[2] Dihlmann, M. and Haasdonk, B. 2014 A reduced basis Kalman filter for parametrized partial differential equations ESAIM: Control, Optimisation and Calculus of Variations at press
[3] Engl, H., Hanke, M. and Neubauer, A. 1996 Regularization of inverse problems Mathematics and its Applications (Dordrecht: Kluwer Academic)
[4] Evans, L. C. 1998 Partial Differential Equations (Graduate Studies in Mathematics vol 19) (Providence, RI: American Mathematical Society)
[5] Hazewinkel, M., Kaashoek, J. and Leynse, B. 1986 Pattern formation for a one-dimensional evolution equation based on Thoms river basin model Disequilibrium and Self-Organisation (Mathematics and its Applications vol 30) ed C W Kilmister (Berlin: Springer)
[6] Ioannou, P. and Sun, J. 1996 Robust Adaptive Control (Englewood Cliffs, NJ: Prentice-Hall)
[7] Jategaonkar, R. 2006 Flight Vehicle System Identification: A Time Domain Methodology (Progress in Astronautics and Aeronautics vol 216) (Reston, VA: American Institute of Aeronautics and Astronautics)
[8] Keung Y and Zou J 1998 Numerical identifications of parameters in parabolic systems Inverse Problems 14 83–100

[9] Kügler P 2010 Online parameter identification without ricatti-type equations in a class of time-dependent partial differential equations: an extended state approach with potential to partial observations Inverse Problems 26 035004

[10] Maasoumy M, Moridian B, Razmara M, Shabakhti M and Sangiovanni Vincentekku A 2014 Online simultaneous date estimation and parameter adaptation for building predictive control Center for Research in Energy Systems Transformation (CREST) Retrieved from: https://escholarship.org/uc/item/2pg8s65g.

[11] Narendra K and Annaswamy A 2005 Stable Adaptive Systems (Mineola, New York: Dover Publications)

[12] Rahimi-Eichi H, Baronti F and Chow M 2014 Online adaptive parameter identification and state-of-charge coestimation for lithium-polymer battery cells IEEE Trans. Ind. Electron. 61 2053–61

[13] Sastry S and Bodson M 2011 Adaptive Control: Stability, Convergence and Robustness (Mineola, New York: Dover Publications)

[14] Yamamoto M and Zou J 2001 Simultaneous reconstruction of the initial temperature and heat radiative coefficient Inverse Problems 17 1181–202