Optimal control and model-order reduction of an abstract parabolic system containing a controlled bilinear form

Applied to the example of a controlled advection term in an advection-diffusion equation

Stefan Banholzer, Dennis Beermann

Department of Mathematics and Statistics, University of Konstanz, 78457 Konstanz, Germany
(E-Mail: Dennis.Beermann@uni-konstanz.de)

Abstract

In the present paper, a linear parabolic evolution equation is considered whose bilinear form is controlled from a general Banach space. The control-to-state operator and some important properties thereof are presented. For a quadratic objective function, the gradient in the control space is derived. A-posteriori error estimators are presented for a given reduced-order model (ROM) with respect to both the cost function and the gradient.

Keywords: Partial differential equations, optimal control, reduced-order modelling, a-posteriori error analysis

1 Introduction

Many real-world, time-dependent systems can be modeled by parabolic evolution equations. The resulting systems usually contain one or several algebraic parameters that have to be specified in an optimal way to suit the demands of the application. Two of the most prominent examples are:

1. Parameter identification: Identify the set of parameters such that the solution to the model best approximates the measured data from the real-life system. Usually, this leads to inverse problems, see Isakov (2006) or Vogel (1999)
2. Optimal control: Use external influences to control the system in a way that the solution of the model approximates a predefined desired state, see Tröltzsch (2010) or Hinze et al. (2009).

A parabolic system is usually defined by an operator and an inhomogeneity. In the standard literature of optimal control, the external influence on the system tends to take the form of an inhomogeneity, meaning that it is additively separated by the solution variable. In this paper, we consider the case where the operator *itself* is controlled along with the inhomgeneity, thereby leading the way to more complex optimization models. We will only treat linear parabolic equations in this paper. However, even if the operator depends linearily on the control variable, the control-to-state operator itself will be nonlinear. As an example application, we consider a source-free advection-diffusion equation where the advective term may be controlled in order to steer the system. Due to the fact that in an optimal control run, large systems have to be solved repeatedly, it is often advisable to employ model-order reduction in order to save computation time. There is extensive literature to be found for a general overview. For Reduced Basis (RB) techniques, we refer to Patera and Rozza (2006), Hesthaven et al. (2016) and Quarteroni et al. (2016). As far as Proper Orthogonal Decomposition (POD) is concerned, Holmes et al. (2012) and Gubisch and Volkwein (2016) offer an excellent introduction. For model order reduction techniques to work properly, error estimators are required to measure the quality of reduced properties during the optimization programs. We present a general error estimator for a quadratic function and also a specific one for the gradients of the two most common cost functions in the case of the advection-diffusion equation.

This paper is organized as follows:
In Section 2, we introduce the general parabolic equation containing a control variable. Under given coercivity assumptions, we show that the system admits a unique solution for every control, allowing us to define the (nonlinear) control-to-state operator and to present an energy estimate for the solution. We then show that under certain requirements as to how the control variable influences the operator and the inhomogeneity, this operator is continuous and Fréchet-differentiable. Gradients of generalized quadratic cost functions are derived in Section 3. Section 4 focuses on Reduced-order modelling and introduces general error estimators for the state solution and a quadratic cost function from Section 3. In Section 5, we consider the concrete example of controlling an advection term in an advection-diffusion equation: For a cost function, we consider tracking both the entire trajectory and the state at the final time. For both cases, we derive ROM-error estimators for the gradients.

2 The general parabolic equation

Throughout these pages, we consider the following linear parabolic evolution system:

$$y_t(t) + A(u)(t)y(t) = f(u)(t) \text{ in } V' \qquad \text{f.a.a. } t \in (0,T) \tag{1a}$$
$$y(0) = y_0 \qquad \text{in } H \tag{1b}$$

where $V \hookrightarrow H = H' \hookrightarrow V'$ is a Gelfand triple and $T > 0$ is the final time. The control space is given by $\mathcal{U} := L^2(0,T;U)$ where U is a Hilbert space. Let $A : \mathcal{U} \to L^\infty(0,T;L(V,V'))$ and $f : \mathcal{U} \to L^2(0,T;V')$ be the control-dependent bilinear form and inhomogeneity of equation (1a). Lastly, $y_0 \in H$ is an initial time.

If we fix a control $u \in \mathcal{U}$, we may define $B := A(u)$ and $g := f(u)$ so that (1) reads:

$$y_t(t) + B(t)y(t) = g(t) \text{ in } V' \qquad \text{f.a.a. } t \in (0,T) \tag{2a}$$
$$y(0) = y_0 \qquad \text{in } H \tag{2b}$$

Observe that $\|B(t)\|_{L(V,V')} \leq C$ f.a.a. $t \in (0,T)$ with $C = \|A(u)\|_{L^\infty(0,T;L(V,V'))}$.

We start by giving a solvability result on (2):

2.1 Theorem. *Assume that there exist constants $\alpha > 0, \beta \geq 0$ such that $B \neq 0$ is uniformly coercive:*

$$\langle B(t)\varphi, \varphi\rangle_{V'\times V} \;\geq\; \alpha \|\varphi\|_V^2 - \beta \|\varphi\|_H^2 \quad \textit{f.a.a. } t \in (0,T), \textit{ for all } \varphi \in V \tag{3}$$

Then there exists a unique solution $y \in W(0,T) := L^2(0,T;V) \cap H^1(0,T;V')$ of (2) that satisfies

$$\|y(T)\|_H^2 + \|y\|_{L^2(0,T;V)}^2 + \|y_t\|_{L^2(0,T;V')}^2 \;\leq\; C\left(\|y_0\|_H^2 + \|g\|_{L^2(0,T;V')}^2\right) \tag{4}$$

where the constant $C > 0$ depends continuously on $\|B\|_{L^\infty(0,T;L(V,V'))}$ and α, β in (3). In particular, it holds

$$\|y(T)\|_H^2 + \|y\|_{L^2(0,T;V)}^2 \; dt \leq \frac{e^{2\beta T}}{\alpha}\left(\|y_0\|_H^2 + \frac{1}{\alpha}\|g\|_{L^2(0,T;V')}^2\right) \tag{5}$$

Furthermore, the mapping $(g, y_0) \mapsto y$ is linear from $L^2(0,T;V') \times H$ to $W(0,T)$.

Proof. We define the time-dependent bilinear form

$$b : V \times V \times (0,T) \to \mathbb{R} : \quad b(\varphi,\psi;t) := \langle B(t)\varphi, \psi\rangle_{V'\times V}$$

and observe that

$$|b(\varphi,\psi,t)| \leq \|B(t)\|_{L(V,V')} \cdot \|\varphi\|_V \cdot \|\psi\|_V \leq \underbrace{\|B\|_{L^\infty(0,T;L(V,V'))}}_{\neq 0 \text{ since } B \neq 0} \cdot \|\varphi\|_V \cdot \|\psi\|_V$$

So the bilinear form b is t-uniformly continuous and, because of (3), t-uniformly coercive. Equation (2a) is now equivalent to

$$\langle y_t(t), \varphi \rangle_{V' \times V} + b(y(t), \varphi; t) = \langle g(t), \varphi \rangle_{V' \times V} \quad \text{for all } \varphi \in V, \text{ f.a.a. } t \in (0,T)$$

As it was shown in (Hinze et al., 2009, Theorem 1.37), a unique solution $y \in W(0,T)$ of this problem exists. The fact that y is linear in (g, y_0) can be easily verified by hand. For the energy estimates (4) and (5), we start by utilizing (2) along with Young's inequality:

$$\begin{aligned}
\frac{d}{dt} \|y(t)\|_H^2 &= 2\langle y_t(t), y(t) \rangle_{V' \times V} = 2\Big[\langle g(t), y(t) \rangle_{V' \times V} - \langle B(t)y(t), y(t) \rangle_{V' \times V}\Big] \\
&\leq 2\left[\frac{1}{2\alpha} \|g(t)\|_{V'}^2 + \frac{\alpha}{2} \|y(t)\|_V^2 - \alpha \|y(t)\|_V^2 + \beta \|y(t)\|_H^2\right] \\
&= \frac{1}{\alpha} \|g(t)\|_{V'}^2 - \alpha \|y(t)\|_V^2 + 2\beta \|y(t)\|_H^2
\end{aligned} \tag{6}$$

This especially implies

$$\frac{d}{dt} \|y(t)\|_H^2 \leq \frac{1}{\alpha} \|g(t)\|_{V'}^2 + 2\beta \|y(t)\|_H^2$$

Utilizing Gronwall's Lemma, we obtain

$$\|y(t)\|_H^2 \leq e^{2\beta t} \left(\|y_0\|_H^2 + \frac{1}{\alpha} \int_0^t \|g(s)\|_{V'}^2 \ ds \right) \tag{7}$$

We now return to (6) and integrate over $(0,T)$:

$$\|y(T)\|_H^2 - \|y(0)\|_H^2 + \alpha \int_0^T \|y(t)\|_V^2 \ dt \leq \frac{1}{\alpha} \int_0^T \|g(t)\|_{V'}^2 + 2\beta \int_0^T \|y(t)\|_H^2 \ dt$$

By utilizing (2b) and (7), this leads to

$$\begin{aligned}
\|y(T)\|_H^2 + \alpha \int_0^T \|y(t)\|_V^2 \ dt &\leq \|y_0\|_H^2 + \frac{1}{\alpha} \|g\|_{L^2(0,T;V')}^2 + 2\beta \int_0^T e^{2\beta t} \left(\|y_0\|_H^2 + \frac{1}{\alpha} \int_0^t \|g(s)\|_{V'}^2 \ ds \right) \ dt \\
&\leq \left(1 + 2\beta \int_0^T e^{2\beta t} \ dt \right) \|y_0\|_H^2 + \frac{1}{\alpha} \|g\|_{L^2(0,T;V')}^2 + \frac{2\beta}{\alpha} \int_0^T e^{2\beta t} \int_0^T \|g(s)\|_{V'}^2 \ ds \ dt \\
&= e^{2\beta T} \left(\|y_0\|_H^2 + \frac{1}{\alpha} \|g\|_{L^2(0,T;V')}^2 \right)
\end{aligned} \tag{8}$$

In particular, this proves (5).

To obtain the estimate for y_t, let $v \in V$ be given arbitrarily with $\|v\|_V = 1$. We observe using (2a):

$$|\langle y_t(t), v \rangle_{V' \times V}| \leq \|g(t) - B(t)y(t)\|_{V'} \leq \Big(1 + \|B\|_{L^\infty(0,T;L(V,V'))} \Big) \left(\|g(t)\|_{V'} + \|y(t)\|_V \right)$$

This implies:

$$\|y_t\|_{L^2(0,T;V')}^2 \leq \underbrace{2 \Big(1 + \|B\|_{L^\infty(0,T;L(V,V'))} \Big)^2}_{=:C_t} \Big(\|g\|_{L^2(0,T,V')}^2 + \|y\|_{L^2(0,T;V)}^2 \Big)$$

and if we utilize (5), we can further deduce:

$$\begin{aligned}
\|y_t\|_{L^2(0,T;V')}^2 &\leq C_t \left(\|g\|_{L^2(0,T;V')}^2 + \frac{e^{2\beta T}}{\alpha} \left(\|y_0\|_H^2 + \frac{1}{\alpha} \|g\|_{L^2(0,T;V')}^2 \right) \right) \\
&= C_t \frac{e^{2\beta T}}{\alpha} \|y_0\|_H^2 + C_t \left(1 + \frac{e^{2\beta T}}{\alpha^2} \right) \|g\|_{L^2(0,T;V')}^2 \\
&\leq \underbrace{C_t \left(1 + \frac{\alpha + 1}{\alpha^2} e^{2\beta T} \right)}_{=:\tilde{C}} \Big(\|y_0\|_H^2 + \|g\|_{L^2(0,T;V')}^2 \Big)
\end{aligned} \tag{9}$$

We can easily verify that $\frac{e^{2\beta T}}{\alpha} \leq \tilde{C}$. Therefore, by setting $C := 2\tilde{C}$, (4) holds true and C depends continuously on α, β and $\|B\|$. □

To derive from Theorem 2.1 the solvability of equation (1) for any given control $u \in \mathcal{U}$, we have to make some assumptions on A:

2.2 Corollary. *Assume that for every control $u \in \mathcal{U}$, there are $\alpha_u > 0$ and $\beta_u \geq 0$ such that*

$$\langle A(u)(t)\varphi, \varphi\rangle_{V'\times V} \;\geq\; \alpha_u \|\varphi\|_V^2 - \beta_u \|\varphi\|_H^2 \quad \text{f.a.a. } t \in (0,T), \text{ for all } \varphi \in V \tag{10}$$

Then for every control $u \in \mathcal{U}$, there exists a unique solution $y \in W(0,T)$ of (1) which satisfies

$$\|y(T)\|_H^2 + \|y\|_{L^2(0,T;V)}^2 + \|y_t\|_{L^2(0,T;V')}^2 \;\leq\; C_u \left(\|y_0\|_H^2 + \|f(u)\|_{L^2(0,T;V')}^2 \right) \tag{11}$$

where the constant C_u depends continuously on $\|A(u)\|_{L^\infty(0,T;L(V,V'))}$ as well as α_u, β_u. We write $y = G(u)$ and have a solution operator $G : \mathcal{U} \to W(0,T)$.

Next we are interested in properties of the solution operator. For this we first need a result about the coercivity constants α_u, β_u in (10).

2.3 Lemma. *Assume that (10) is satisfied for a control $\bar{u} \in \mathcal{U}$ and that A is continuous in $\bar{u}$. Then for every $\varepsilon > 0$, there is $\delta > 0$ such that*

$$\langle A(u)(t)\varphi, \varphi\rangle_{V'\times V} \geq (\alpha_{\bar{u}} - \varepsilon) \|\varphi\|_V^2 - \beta_{\bar{u}} \|\varphi\|_H^2 \quad \text{for all } \varphi \in V \tag{12}$$

holds for all $u \in \mathcal{U}$ with $\|u - \bar{u}\|_{\mathcal{U}} < \delta$. In this sense, the coercivity constants α, β from (10) are continuous with respect to u.

Proof. We have for all $\varphi \in V$:

$$\begin{aligned}
\langle A(u)(t)\varphi, \varphi\rangle_{V'\times V} &= \langle A(\bar{u})(t)\varphi, \varphi\rangle_{V'\times V} + \langle [A(u) - A(\bar{u})](t)\varphi, \varphi\rangle_{V'\times V} \\
&\geq \alpha_{\bar{u}} \|\varphi\|_V^2 - \beta_{\bar{u}} \|\varphi\|_H^2 - \|[A(u) - A(\bar{u})](t)\|_{L(V,V')} \cdot \|\varphi\|_V^2 \\
&\geq \left(\alpha_{\bar{u}} - \|A(u) - A(\bar{u})\|_{L^\infty(0,T;L(V,V'))} \right) \|\varphi\|_V^2 - \beta_{\bar{u}} \|\varphi\|_H^2
\end{aligned}$$

Now, since A is continuous in $\bar{u}$, there is $\delta > 0$ such that for $\|u - \bar{u}\|_{\mathcal{U}} < \delta$, it is $\|A(u) - A(\bar{u})\| < \varepsilon$. This proves the lemma. □

2.4 Lemma. *Let (10) be satisfied and A and f be continuous mappings on $\mathcal{U}$. Then the solution operator G is continuous. If A and f are locally Lipschitz continuous, then G is locally Lipschitz continuous.*

Proof. Consider two controls $u_1, u_2 \in \mathcal{U}$ with $\|u_1 - u_2\| < \varepsilon$ and their according solutions $y_1, y_2 \in W(0,T)$, i.e. $y_i = G(u_i)$ $(i = 1, 2)$. Then the difference $y := y_1 - y_2$ solves the following differential equation for almost all $t \in (0,T)$ in V':

$$\begin{aligned}
& y_t(t) + A(u_1)(t)y_1(t) - A(u_2)(t)y_2(t) = f(u_1)(t) - f(u_2)(t) \\
\Leftrightarrow \quad & y_t(t) + A(u_1)(t)y(t) = \underbrace{[f(u_1) - f(u_2)]\,(t)}_{=:g_1(t)} - \underbrace{[A(u_1) - A(u_2)]\,(t)y_2(t)}_{=:g_2(t)}
\end{aligned}$$

along with $y(0) = y_0 - y_0 = 0$. We know that $g_1 \in L^2(0,T;V')$. Furthermore,

$$\|g_2(t)\|_{V'} \leq \|[A(u_1) - A(u_2)](t)\|_{L(V,V')} \cdot \|y_2(t)\|_V \leq \|A(u_1) - A(u_2)\|_{L^\infty(0,T;L(V,V'))} \cdot \|y_2(t)\|_V$$

which yields $g_2 \in L^2(0,T;V')$ with

$$\|g_2\|_{L^2(0,T;V')} \leq \|A(u_1) - A(u_2)\|_{L^\infty(0,T;L(V,V'))} \cdot \|y_2\|_{L^2(0,T;V)}$$

We can therefore apply Theorem 2.1 and obtain the estimate

$$\begin{aligned}\|y\|_{W(0,T)}^2 &\le C_{A(u_1)} \Big(\|f(u_1) - f(u_2)\|_{L^2(0,T;V')} + \|A(u_1) - A(u_2)\|_{L^\infty(0,T;L(V,V'))} \cdot \|y_2\|_{L^2(0,T;V)} \Big)^2 \\ &\le 2C_{A(u_2)} \Big(\|f(u_1) - f(u_2)\|_{L^2(0,T;V')} + \|A(u_1) - A(u_2)\|_{L^\infty(0,T;L(V,V'))} \cdot \|y_2\|_{L^2(0,T;V)} \Big)^2\end{aligned}$$

The last inequality holds if ε is small enough. This is due to the fact that by Theorem 2.1, $C_{A(u)}$ depends continuously on the coercivity constants α_u and β_u as well as $\|A(u)\|$. It was shown in Lemma 2.3 that these parameters can be seen as depending continuously on u which makes the constant $C_{A(u)}$ depend continuously on u. By the continuity of f and A from $\mathcal{U}$, we can see that $\|y\|_{W(0,T)} \to 0$ as $u_1 \to u_2$, implying that G is continuous in u_2. If, in addition, f and A are locally Lipschitz continuous, we obtain for small enough ε:

$$\|y\|_{W(0,T)}^2 \le 2C_{A(u_2)} \Big(L_f(u_2) + L_A(u_2) \|y_2\|_{L^2(0,T;V)} \Big)^2 \cdot \|u_1 - u_2\|_{\mathcal{U}}^2$$

where $L_f(u_2)$ and $L_A(u_2)$ are the local Lipschitz constants in u_2. Therefore, G is locally Lipschitz-continuous in u_2. □

2.5 Lemma. *Let (10) be satisfied and the mappings A and f be Fréchet differentiable (in u). Then the solution operator G is Fréchet differentiable and its derivative is given by $G'(u)h = y^h$ for $u, h \in \mathcal{U}$, where $y^h \in W(0,T)$ satisfies the system*

$$y_t^h(t) + A(u)(t) y^h(t) = (f'(u)h)(t) - (A'(u)h)(t)\bar{y}(t) \text{ in } V' \qquad \text{f.a.a. } t \in (0,T) \qquad (13a)$$

$$y^h(0) = 0 \qquad \text{in } H \qquad (13b)$$

and $\bar{y}$ is the solution for the control u, i.e. $\bar{y} = G(u)$.

Proof. Consider a control $u \in \mathcal{U}$ and a direction $h \in \mathcal{U}$. We define $y := G(u+h) - G(u)$. We proceed similarly to the proof of Lemma 2.4 and observe that y satisfies the system

$$y_t(t) + A(u)(t)y(t) = [f(u+h) - f(u)](t) - [A(u+h) - A(u)](t)\,[G(u+h)(t)] \text{ in } V' \quad \text{f.a.a. } t \in (0,T) \qquad (14a)$$

$$y(0) = 0 \qquad \text{in } H \qquad (14b)$$

Now, since f and A are Fréchet differentiable, this means

$$f(u+h) - f(u) = f'(u)h + r_f(u,h), \quad A(u+h) - A(u) = A'(u)h + r_A(u,h)$$

with $r_f(u,h) \in L^2(0,T;V')$, $r_A(u,h) \in L^\infty(0,T;L(V,V'))$ such that

$$\frac{\|r_f(u,h)\|_{L^2(0,T;V')}}{\|h\|_{\mathcal{U}}} \to 0, \quad \frac{\|r_A(u,h)\|_{L^\infty(0,T;L(V,V'))}}{\|h\|_{\mathcal{U}}} \to 0 \quad \text{as } \|h\|_{\mathcal{U}} \to 0 \qquad (15)$$

Inserting this into the right-hand side of (14a) yields

$$\begin{aligned} & [f(u+h) - f(u)](t) - [A(u+h) - A(u)](t)\,[G(u+h)(t)] \\ =\; & [f'(u)h + r_f(u,h)](t) - [A'(u)h + r_A(u,h)](t)[G(u+h)(t)] \\ =\; & [(f'(u)h)(t) - (A'(u)h)(t)G(u)(t)] \\ & + [r_f(u,h)(t) - r_A(u,h)(t)G(u+h)(t) - (A'(u)h)(t)(G(u+h) - G(u))(t)] \end{aligned}$$

So now y solves the system

$$y_t(t) + A(u)(t)y(t) = g_1(t) + g_2(t) \text{ in } V' \qquad \text{f.a.a. } t \in (0,T) \qquad (16a)$$

$$y(0) = 0 \qquad \text{in } H \qquad (16b)$$

with

$$\begin{aligned} g_1(t) &:= (f'(u)h)(t) - (A'(u)h)(t)G(u)(t) \\ g_2(t) &:= r_f(u,h)(t) - r_A(u,h)(t)G(u+h)(t) - (A'(u)h)(t)(G(u+h) - G(u))(t) \end{aligned}$$

Note that the term $y = G(u+h) - G(u)$ still appears on the right-hand side of (16a) within g_2 yet is treated as an inhomogeneity of (1a). This is possible since we already know that y exists and is a solution to (16). It is straightforward to show that $g_1, g_2 \in L^2(0,T;V')$, so Theorem 2.1 is applicable. Therefore, the solution of this equation depends linearly on the right-hand side of (16a), meaning that y allows for the decomposition $y = y^h + y^\delta$ where $y^h, y^\delta \in W(0,T)$ satisfy the systems

$$\begin{aligned} y_t^h(t) + A(u)(t)y^h(t) &= g_1(t) \text{ in } V' && \text{f.a.a. } t \in (0,T) \\ y(0) &= 0 && \text{in } H \end{aligned}$$

as well as

$$\begin{aligned} y_t^\delta(t) + A(u)(t)y^\delta(t) &= g_2(t) \text{ in } V' && \text{f.a.a. } t \in (0,T) \\ y(0) &= 0 && \text{in } H \end{aligned}$$

We want to show that $y^h = G'(u)h$. For this, we have to show that y^h is linear and continuous in h and that ${\|y^\delta\|_{W(0,T)}}/{\|h\|_{\mathcal{U}}} \to 0$ as $\|h\|_{\mathcal{U}} \to 0$.
We start with y^h. The linearity follows directly since g_1 depends linearly on h and y^h depends linearly on g_1 as it was shown in Theorem 2.1. For the continuity, we infer from (4) that

$$\begin{aligned} \left\|y^h\right\|_{W(0,T)}^2 &\leq C_{A(u)} \left\|f'(u)h - (A'(u)h)(\cdot)G(u)(\cdot)\right\|_{L^2(0,T;V')}^2 \\ &\leq C_{A(u)} \Big(\|f'(u)\|_{L(\mathcal{U},L^2(0,T;V'))} + \|A'(u)\|_{L(\mathcal{U},L^\infty(0,T;L(V,V')))} \cdot \|G(u)\|_{L^2(0,T;V)} \Big)^2 \cdot \|h\|_{\mathcal{U}}^2 \end{aligned} \tag{17}$$

The above estimate shows that y^h is continuous in $h = 0$ which, along with the linearity in h, implies that y^h is continuous in h everywhere. Altogether, we have shown that

$$(h \mapsto y^h) \in L(\mathcal{U}, W(0,T))$$

We now turn to y^δ. Again from the energy estimate (4), we infer that

$$\begin{aligned} \left\|y^\delta\right\|_{W(0,T)}^2 &\leq C_{A(u)} \left\|r_f(u,h) - r_A(u,h)(\cdot)G(u+h)(\cdot) - (A'(u)h)(\cdot)(G(u+h) - G(u))(\cdot)\right\|_{L^2(0,T;V')}^2 \\ &\leq C_{A(u)} \Big(\|r_f(u,h)\|_{L^2(0,T;V')} + \|r_A(u,h)\|_{L^\infty(0,T;L(V,V'))} \cdot \|G(u+h)\|_{L^2(0,T;V)} \\ &\quad + \|A'(u)\|_{L(\mathcal{U},L^\infty(0,T;L(V,V')))} \cdot \|h\|_{\mathcal{U}} \cdot \|G(u+h) - G(u)\|_{L^2(0,T;V')} \Big)^2 \end{aligned}$$

This yields

$$\begin{aligned} \left(\frac{\|y^\delta\|_{W(0,T)}}{\|h\|_{\mathcal{U}}} \right)^2 &\leq C_{A(u)} \Bigg(\frac{\|r_f(u,h)\|_{L^2(0,T;V')}}{\|h\|_{\mathcal{U}}} + \frac{\|r_A(u,h)\|_{L^\infty(0,T;L(V,V'))}}{\|h\|_{\mathcal{U}}} \cdot \|G(u+h)\|_{L^2(0,T;V)} \\ &\quad + \|A'(u)\|_{L(\mathcal{U},L^\infty(0,T;L(V,V')))} \cdot \|G(u+h) - G(u)\|_{L^2(0,T;V')} \Bigg)^2 \end{aligned}$$

Again with the fact that G is continuous in u and (15), we obtain ${\|y^\delta\|_{W(0,T)}}/{\|h\|_{\mathcal{U}}} \to 0$ as $\|h\|_{\mathcal{U}} \to 0$. This shows that the derivative of G at u in the direction h is indeed given by y^h and we can write $G'(u)h = y^h$.

□

3 Quadratic cost functions

In this section, let us assume that we are given a cost function of the form

$$\hat{J} : \mathcal{U} \to \mathbb{R}, \quad \hat{J}(u) := \tfrac{1}{2} \left\|\Phi G(u) - y_d\right\|_X^2 \tag{18}$$

where $G : \mathcal{U} \to W(0,T)$ is the solution operator, X is a Hilbert space and $\Phi \in L(W(0,T), X)$ is an observation operator. The vector $y_d \in X$ is called a desired state. We would like to derive a gradient representation of J. First of all, we observe that we can use the decomposition $\hat{J} = J \circ G$ with the non-reduced cost function

$$J : W(0,T) \to \mathbb{R}, \quad J(y) := \tfrac{1}{2} \|\Phi y - y_d\|_X^2$$

The derivative of this function is easy to compute:

3.1 Lemma. *The function J is Fréchet differentiable and the derivative is given by*

$$J'(y)h_y = \langle \Phi y - y_d, \Phi h_y \rangle_X \quad \textit{for all } y, h_y \in W(0,T) \tag{19}$$

Proof. Let $y, h_y \in W(0,T)$ be arbitrarily given. Then we observe

$$\begin{aligned} J(y + h_y) - J(y) &= \tfrac{1}{2} \left(\|\Phi y + \Phi h_y - y_d\|_X^2 - \|\Phi y - y_d\|_X^2 \right) \\ &= \tfrac{1}{2} \left(\|\Phi y - y_d\|_X^2 + 2\langle \Phi y - y_d, \Phi h_y \rangle_X + \|\Phi h_y\|_X^2 - \|\Phi y - y_d\|_X^2 \right) \\ &= \langle \Phi y - y_d, \Phi h_y \rangle_X + \tfrac{1}{2} \|\Phi h_y\|_X^2 \end{aligned}$$

and clearly, $h_y \mapsto \langle \Phi y - y_d, \Phi h_y \rangle_X$ is a linear and continuous mapping from $W(0,T)$ to $\mathbb{R}$. Furthermore,

$$\frac{\frac{1}{2} \|\Phi h_y\|_X^2}{\|h_y\|_{W(0,T)}} \leq \tfrac{1}{2} \|\Phi\|_{L(W(0,T),X)}^2 \cdot \|h_y\|_{W(0,T)} \;\to\; 0 \quad \text{as } \|h_y\|_{W(0,T)} \to 0$$

This implies that J is differentiable in y with the proposed derivative. □

Using Lemma 3.1, we can immediately see that $\hat{J}$ is differentiable:

3.2 Corollary. *Let (10) be satisfied and the mappings A and f be Fréchet differentiable from $\mathcal{U}$. Then the cost function $\hat{J}$ is Fréchet differentiable from $\mathcal{U}$ to $\mathbb{R}$ and the derivative is given by*

$$\hat{J}'(u)h = \langle \Phi \bar{y} - y_d, \Phi y^h \rangle_X \tag{20}$$

where $\bar{y} := G(u) \in W(0,T)$ is the state solution and $y^h := G'(u)h$ the solution to (13).

Proof. Seeing as $\hat{J} = J \circ G$, and that J and G themselves are differentiable as it was shown in the Lemmata 2.5 and 3.1, $\hat{J}$ is differentiable by the chain rule and the derivative is given by

$$\hat{J}'(u)h = J'(G(u))\,[G'(u)h] \quad \text{for all } u, h \in \mathcal{U}$$

It was shown in Lemma 2.5 that $y^h := G'(u)h$ satisfies the system (13). Inserting this into the representation (19), we end up with

$$\hat{J}'(u)h = \langle \Phi G(u) - y_d, \Phi y^h \rangle_X \tag{21}$$

□

We have given a representation for the derivative of the reduced cost function. For the purposes of optimization, the notion of a gradient is additionally required:

3.3 Corollary. *Let (10) be satisfied and the mappings A and f be Fréchet differentiable. Then the cost function $\hat{J}$ has a gradient which is given by*

$$\nabla \hat{J} : \mathcal{U} \to \mathcal{U} : \quad \nabla \hat{J}(u) = (\Phi G'(u))^* \,(\Phi G(u) - y_d) \tag{22}$$

where $(\Phi G'(u))^* \in L(X;\mathcal{U})$ is the Hilbert adjoint of the operator $\Phi G'(u) \in L(\mathcal{U}, X)$.

Proof. Follows directly from the representation (21) by definition of the Hilbert adjoint. □

4 Reduced Order Modelling

In this chapter, we assume that we are given a finite-dimensional subspace $V^\ell \subset V$ which is spanned by a V-orthonormal basis $\{\varphi_1, ..., \varphi_\ell\}$. We employ a Galerkin projection of (1) onto V^ℓ and look for a function y^ℓ satisfying:

$$\langle y_t^\ell(t), \varphi\rangle_{V'\times V} + \langle A(u)(t)y^\ell(t), \varphi\rangle_{V'\times V} = \langle f(u)(t), \varphi\rangle_{V'\times V} \quad \text{f.a.a. } t \in (0,T), \text{ for all } \varphi \in V^\ell \tag{23a}$$

$$y^\ell(0) = \mathcal{P}^\ell y_0 \text{ in } H \tag{23b}$$

where $\mathcal{P}^\ell \in L(V)$ is a projection operator onto V^ℓ. We express y^ℓ through the basis of V^ℓ:

4.1 Lemma. *Assume that we are given a coefficient function $a^\ell \in H^1(0,T;\mathbb{R}^\ell)$ such that*

$$y^\ell : (0,T) \to V, \quad y^\ell(t) := \sum_{i=1}^{\ell} a_i^\ell(t)\varphi_i \quad \text{f.a.a. } t \in (0,T) \tag{24}$$

Then it is $y^\ell \in H^1(0,T;V) \subset W(0,T)$ with derivative $y_t^\ell(t) = \sum_{i=1}^\ell \dot{a}_i^\ell(t)\varphi_i$ for almost all $t \in (0,T)$.

Proof. Since $W(0,T) = L^2(0,T;V) \cap H^1(0,T;V')$, we proceed in two steps:

(i) For almost all $t \in (0,T)$, we have

$$\left\|y^\ell(t)\right\|_V \leq \sum_{i=1}^{\ell} \left|a_i^\ell(t)\right| \cdot \|\varphi_i\|_V \leq \left(\sum_{i=1}^{\ell} \left|a_i^\ell(t)\right|^2\right)^{1/2} \cdot \underbrace{\left(\sum_{i=1}^{\ell} \|\varphi_i\|_V^2\right)^{1/2}}_{=:C_V} = C_V \left|a^\ell(t)\right|_{\mathbb{R}^\ell}$$

and, since $a^\ell \in L^2(0,T;\mathbb{R}^\ell)$, this yields $y^\ell \in L^2(0,T;V)$.

(ii) For almost all $t \in (0,T)$, it holds that

$$\begin{aligned}\left\|y^\ell(t+h) - y^\ell(t) - h\sum_{i=1}^{\ell} \dot{a}_i^\ell(t)\varphi_i\right\|_V &\leq \sum_{i=1}^{\ell} \left|a_i^\ell(t+h) - a_i^\ell(t) - \dot{a}_i^\ell(t)h\right| \cdot \|\varphi_i\|_V \\ &\leq C_V \left|a^\ell(t+h) - a^\ell(t) - \dot{a}^\ell(t)h\right|_{\mathbb{R}^\ell}\end{aligned}$$

which implies

$$\frac{\left\|y^\ell(t+h) - y^\ell(t) - h\sum_{i=1}^{\ell} \dot{a}_i^\ell(t)\varphi_i\right\|_V}{|h|} \leq C_V \frac{\left|a^\ell(t+h) - a^\ell(t) - \dot{a}^\ell(t)h\right|_{\mathbb{R}^\ell}}{|h|} \to 0 \quad \text{as } |h| \to 0$$

Furthermore, the mapping $h \mapsto h\sum_{i=1}^{\ell} \dot{a}_i(t)\varphi_i$ is obviously linear and continuous from $\mathbb{R}$ to $L^2(0,T;V)$. So indeed $y^\ell \in H^1(0,T;V)$ with the proposed derivative.

□

Inserting y^ℓ from (24) into (23) yields a system for a^ℓ:

$$\dot{a}^\ell(t) + A^\ell(u)(t)a^\ell(t) = f^\ell(u)(t) \text{ in } \mathbb{R}^\ell \qquad \text{f.a.a. } t \in (0,T) \tag{25a}$$

$$a(0) = a^0 \text{ in } \mathbb{R}^\ell \tag{25b}$$

where $a^0 \in \mathbb{R}^\ell$ is the basis representation of $\mathcal{P}^\ell y_0$ in the basis of V^ℓ. Furthermore, $A^\ell(u) : (0,T) \to \mathbb{R}^{\ell\times\ell}$ and $f^\ell(u) : (0,T) \to \mathbb{R}^\ell$ for all $u \in \mathcal{U}$ with

$$A_{ij}^\ell(u)(t) = \langle A(u)(t)\varphi_j, \varphi_i\rangle_{V'\times V} \qquad \text{for all } u \in \mathcal{U}, \text{ f.a.a. } t \in (0,T)$$

$$f_i^\ell(u)(t) = \langle f(u)(t), \varphi_i\rangle_{V'\times V} \qquad \text{for all } u \in \mathcal{U}, \text{ f.a.a. } t \in (0,T)$$

4.2 Lemma. *The system (25) admits a unique solution $a^\ell \in H^1(0,T;\mathbb{R}^\ell)$. The function y^ℓ from (24) then solves the system (23) which is its unique solution. We define the solution operator $G_\ell : \mathcal{U} \to W(0,T)$, $u \mapsto y^\ell$.*

Proof. The evolution equation (25a) can equivalently be understood in $(\mathbb{R}^\ell)'$ since $\mathbb{R}^\ell \cong (\mathbb{R}^\ell)'$. We will therefore utilize the trivial Gelfand triple $W(0,T;\mathbb{R}^\ell;\mathbb{R}^\ell)$, i.e. $W(0,T;H;V)$ where $H = V = \mathbb{R}^\ell$. Furthermore, it is:

$$\left\|A^\ell(u)(t)\right\|_{L(\mathbb{R}^\ell,\mathbb{R}^\ell)} \leq C \max_{i,j=1,\dots,\ell} \left|A^\ell_{ij}(u)(t)\right| \leq C \left\|A(u)(t)\right\|_{L(V,V')} \leq C \left\|A(u)\right\|_{L^\infty(0,T;L(V,V'))}$$

where we have used the fact that in the finite-dimensional space $L(\mathbb{R}^\ell,\mathbb{R}^\ell)$, all norms are equivalent and that the system $\{\varphi_1,\dots,\varphi_\ell\}$ is V-orthonormal. Therefore, we have shown that $A^\ell(u) \in L^\infty(0,T;L(\mathbb{R}^\ell,(\mathbb{R}^\ell)'))$ for every $u \in \mathcal{U}$. Furthermore, it is easy to show that $A^\ell(u)$ satisfies (10) due to the fact that $\mathbb{R}^\ell$ is finite-dimensional and for all $u \in \mathcal{U}$, we have

$$\left|A^\ell_{ij}(u)(t)\right| \leq \left\|A(u)(t)\right\|_{L(V,V')} \leq \left\|A(u)\right\|_{L^\infty(0,T;L(V,V'))} \quad \text{for all } i,j = 1,\dots,\ell, \text{ f.a.a. } t \in (0,T)$$

At last, we have

$$\left\|f^\ell(u)(t)\right\|^2_{(\mathbb{R}^\ell)'} = \left|f^\ell(u)(t)\right|^2_{\mathbb{R}^\ell} = \sum_{i=1}^{p} \left|\langle f(u)(t), v^i\rangle_{V'\times V}\right|^2 \leq C_V^2 \left\|f(u)(t)\right\|^2_{V'}$$

with the constant C_V from Lemma 4.1. Therefore, $f^\ell(u) \in L^\infty(0,T;(\mathbb{R}^\ell)')$. Utilizing Corollary 2.2, this means that (25) admits a unique solution $a^\ell \in W(0,T;\mathbb{R}^\ell;(\mathbb{R}^\ell)')$. Due to the fact that $(\mathbb{R}^\ell)' \cong \mathbb{R}^\ell$, this also means that $a^\ell \in H^1(0,T;\mathbb{R}^\ell)$. By Lemma 4.1, this implies $y^\ell \in H^1(0,T;V)$ and it satisfies (23).
The fact that (23) has a unique solution can be proven in the very same way as in Section 2, along with every other result in that section, by replacing the space V with V^ℓ. ☐

4.3 Theorem. *Assume that (10) holds for every control $u \in \mathcal{U}$ and let $y \in W(0,T)$ satisfy the full system (1) and $y^\ell \in W(0,T)$ the reduced system (23). Then the following a-posteriori error estimate holds true:*

$$\left\|y(T) - y^\ell(T)\right\|^2_H + \left\|y - y^\ell\right\|^2_{L^2(0,T;V)} \leq \frac{e^{2\beta_u T}}{\alpha_u}\left(\left\|(1-\mathcal{P}^\ell)y_0\right\|^2_H + \frac{1}{\alpha_u}\left\|R^\ell\right\|^2_{L^2(0,T;V')}\right) \tag{26}$$

where α_u, β_u are the coercivity constants of the operator $A(u)$ from (10) and the residual $R^\ell \in L^2(0,T;V')$ is given by

$$R^\ell(t) = y^\ell_t(t) + A(u)(t)y^\ell(t) - f(u)(t) \ \in V' \quad \text{f.a.a. } t \in (0,T) \tag{27}$$

Proof. We define the error $e := y - y^\ell \in W(0,T)$ and observe that for every $\varphi \in V$, it holds:

$$\begin{aligned}\langle e_t(t),\varphi\rangle_{V'\times V} + \langle A(u)(t)e(t),\varphi\rangle_{V'\times V} = f(u)(t) - \left(\langle y^\ell_t(t),\varphi\rangle_{V'\times V} + \langle A(u)(t)y^\ell(t),\varphi\rangle\right) = -\langle R^\ell(t),\varphi\rangle_{V'\times V}\\ e(0) = (1-\mathcal{P}^\ell)y_0 \text{ in } H\end{aligned}$$

It follows from (5) applied to e that

$$\|e(T)\|^2_V + \|e\|^2_{L^2(0,T;V)} \leq \frac{e^{2\beta_u T}}{\alpha_u}\left(\left\|(1-\mathcal{P}^\ell)y_0\right\|^2_H + \frac{1}{\alpha_u}\left\|R^\ell\right\|^2_{L^2(0,T;V')}\right)$$

☐

In addition to a quadratic cost function $\hat{J}$ as defined in (18), we define the according reduced-order cost function

$$\hat{J}^\ell : \mathcal{U} \to \mathbb{R}, \quad \hat{J}^\ell(u) := J(G_\ell(u)) = \tfrac{1}{2}\left\|\Phi G_\ell(u) - y_d\right\|^2_X$$

We will use the inequality (26) to estimate the error made in the cost function between the full and reduced-order model:

4.4 Corollary. *Assume that (10) holds for every control $u \in \mathcal{U}$ and consider a quadratic cost function $\hat{J}$ as defined in (18). Then the following estimate holds for the cost function:*

$$\left|\hat{J}(u) - \hat{J}^\ell(u)\right| \leq \frac{1}{2}\left\|\Phi(y-y^\ell)\right\|_X^2 + \left\|\Phi(y-y^\ell)\right\|_X \cdot \sqrt{2\hat{J}^\ell(u)} \tag{28}$$

If either $\Phi y = y \in L^2(0,T;L^2(\Omega))$ or $\Phi y = y(T) \in L^2(\Omega)$, we can use estimate (26) for $\left\|\Phi(y-y^\ell)\right\|_X$.

Proof. We begin by utilizing the third Pythagorean Theorem:

$$\begin{aligned}
\left|\hat{J}(u) - \hat{J}^\ell(u)\right| &= \frac{1}{2}\left(\|\Phi G(u) - y_d\|_X^2 - \|\Phi G_\ell(u) - y_d\|_X^2\right) \\
&\leq \frac{1}{2}\left\|\Phi y - y_d - (\Phi y^\ell - y_d)\right\|_X \cdot \left(\|\Phi y - y_d\|_X + \left\|\Phi y^\ell - y_d\right\|_X\right) \\
&\leq \frac{1}{2}\left\|\Phi(y-y^\ell)\right\|_X \cdot \left(\left\|\Phi(y-y^\ell)\right\|_X + 2\left\|\Phi y^\ell - y_d\right\|_X\right) \\
&= \frac{1}{2}\left\|\Phi(y-y^\ell)\right\|_X^2 + \left\|\Phi(y-y^\ell(u))\right\|_X \cdot \sqrt{2\hat{J}^\ell(u)}
\end{aligned}$$

□

In addition to an estimate for the cost function, we require one for the gradient. However, this will depend strongly on the concrete parabolic system and the cost function, which we will cover later.

5 Controlling a convection term

Consider the equation

$$\begin{aligned}
y_t(t,x) - \kappa\Delta y(t,x) + v(u)(t,x)\cdot\nabla y(t,x) &= 0 && \text{in } Q := (0,T)\times\Omega && \text{(29a)} \\
\tfrac{\partial y}{\partial n}(t,x) &= 0 && \text{on } \Sigma := (0,T)\times\Gamma && \text{(29b)} \\
y(0) &= y_0 && \text{in } \Omega && \text{(29c)}
\end{aligned}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain with boundary Γ. We consider the Gelfand triple $V := H^1(\Omega)$, $H := L^2(\Omega)$. For the controlled convection term v, we demand that $v : \mathcal{U} \to L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d))$. As an application to (29), one can for example think of two different fluids that are supposed to be mixed by steering the rotation velocity of mixers in the domain.
To derive the weak formulation of (29), we assume that $y(t) \in H^1(\Omega)$ and $y_t(t) \in (H^1(\Omega))'$ and then 'test' (29a) with a function $\psi \in H^1(\Omega)$ which yields, using Green's identity:

$$\langle y_t(t),\psi\rangle_{H^1(\Omega)'\times H^1(\Omega)} + \kappa\int_\Omega \nabla y(t,x)\cdot\nabla\psi(x)\ dx + \int_\Omega (v(u)(t,x)\cdot\nabla y(t,x))\psi(x)\ dx = 0$$

This motivates us to define, for $u \in \mathcal{U}$ and $t \in (0,T)$:

$$\begin{aligned}
&A(u)(t) \in L(H^1(\Omega), H^1(\Omega)') \\
&\langle A(u)(t)\varphi,\psi\rangle_{V'\times V} := \kappa\int_\Omega \nabla\varphi(x)\cdot\nabla\psi(x)\ dx + \int_\Omega (v(u)(t,x)\cdot\nabla\varphi(x))\psi(x)\ dx
\end{aligned} \tag{30}$$

Of course, we still have to show that this is indeed an element of $L(V,V')$. Basic estimates reveal that

$$\left|\langle A(u)(t)\varphi,\psi\rangle_{H^1(\Omega)'\times H^1(\Omega)}\right| \leq \kappa\|\varphi\|_{H^1(\Omega)}\cdot\|\psi\|_{H^1(\Omega)} + \|v(u)(t,\cdot)\|_{L^\infty(Q)}\cdot\|\varphi\|_{H^1(\Omega)}\cdot\|\psi\|_{H^1(\Omega)}$$

and therefore

$$\|A(u)(t)\|_{L(V,V')} \leq \kappa + \|v(u)(t,\cdot)\|_{L^\infty(Q)} \tag{31}$$

We can therefore write (29) in the form of (1):

$$\begin{aligned}
y_t(t) + A(u)(t)y(t) &= 0 \ \text{ in } H^1(\Omega)' && \text{f.a.a. } t \in (0,T) \\
y(0) &= y_0 && \text{in } L^2(\Omega)
\end{aligned} \tag{32}$$

5.1 Properties of the bilinear form

We start by showing some elementary properties of the bilinear form A:

5.1 Lemma. *Consider the operator A as defined by (30). Then it is*

$$A(u) \in L^\infty(0,T;L(V,V')) \quad \text{for all } u \in \mathcal{U} \tag{33}$$

and condition (10) holds for A with the coercivity constants

$$\alpha_u = \frac{\kappa}{2}, \quad \beta_u = \frac{\kappa}{2} + \frac{\|v(u)\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d))}}{2\kappa}^2 \tag{34}$$

Proof. The fact that (33) holds can be seen from (31) which implies

$$\|A(u)\|_{L^\infty(0,T;L(V,V'))} \leq \kappa + \|v(u)\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d))} \quad \text{for all } u \in \mathcal{U}$$

To prove that (10) holds, let $u \in \mathcal{U}$ be arbitrary but fixed. We then observe

$$\langle A(u)(t)\varphi,\varphi\rangle_{V'\times V} = \underbrace{\kappa\int_\Omega \nabla\varphi(x)\cdot\nabla\varphi(x)\;dx}_{=:(I)} + \underbrace{\int_\Omega (v(u)(t,x)\cdot\nabla\varphi(x))\varphi(x)\;dx}_{=:(II)}$$

For the first term, we simply obtain

$$(I) = \kappa\left(\|\varphi\|_{H^1(\Omega)}^2 - \|\varphi\|_{L^2(\Omega)}^2\right)$$

The second term can be estimated using Young's inequality as follows:

$$\begin{aligned}
(II) &\geq -\|v(u)\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d))}\cdot\|\nabla\varphi\|_{L^2(\Omega;\mathbb{R}^d)}\cdot\|\varphi\|_{L^2(\Omega)}\\
&\geq -\|v(u)\|\left(\tfrac{\kappa}{2\|v(u)\|}\|\nabla\varphi\|_{L^2(\Omega;\mathbb{R}^d)}^2 + \tfrac{\|v(u)\|}{2\kappa}\|\varphi\|_{L^2(\Omega)}^2\right)\\
&= -\tfrac{\kappa}{2}\left(\|\varphi\|_{H^1(\Omega)}^2 - \|\varphi\|_{L^2(\Omega)}^2\right) - \tfrac{\|v(u)\|^2}{2\kappa}\|\varphi\|_{L^2(\Omega)}^2
\end{aligned}$$

Adding (I) and (II) again, we end up with

$$\langle A(u)(t)\varphi,\varphi\rangle_{V'\times V} \geq \tfrac{\kappa}{2}\|\varphi\|_{H^1(\Omega)}^2 - \left(\tfrac{\|v(u)\|^2}{2\kappa} + \tfrac{\kappa}{2}\right)\|\varphi\|_{L^2(\Omega)}^2$$

so condition (10) is satisfied with the coercivity constants proposed in (34). □

5.2 Lemma. *Let the mapping v be Fréchet differentiable from $\mathcal{U}$ to $L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d))$. Then the operator A is Fréchet differentiable from $\mathcal{U}$ to $L^\infty(0,T;L(V,V'))$ with derivative*

$$\langle(A'(u)h)(t)\varphi,\psi\rangle_{V'\times V} = \int_\Omega [(v'(u)h)(t,x)\cdot\nabla\varphi(x)]\,\psi(x)\;dx \quad \text{for all } \varphi,\psi \in V \tag{35}$$

Proof. The proof is straightforward and will not be carried out here. □

5.3 Example. *Assume that $U = \mathbb{R}^p$ and that there exists continuously differentiable coefficients $\eta_i : \mathbb{R}\to\mathbb{R}$ and shape functions $v^i \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d))$ such that $v(u) := \sum_{i=1}^p \eta_i(u_i(t))v^i$ for $u \in \mathcal{U}$. Then it can be shown that v is differentiable with derivative*

$$(v'(u)h)(t,x) = \sum_{i=1}^p \eta_i'(u_i(t))h_i(t)v^i(t,x) \quad \text{for all } u,h \in \mathcal{U}, t \in (0,T), x \in \Omega \tag{36}$$

Next, we consider error estimation and start with Corollary 4.4: The estimator depends on the variables $\alpha(u)$, $\beta(u)$ and R^ℓ which we have to clarify for the current situation:

5.4 Corollary. *Let $y = G(u)$ and $y^\ell = G_\ell(u)$ be full and reduced solutions to (32). Then we obtain the error estimate*

$$\left\|y(T) - y^\ell(T)\right\|^2_{L^2(\Omega)} + \left\|y - y^\ell\right\|^2_{L^2(0,T;V)} \le \frac{2}{\kappa} \exp\left(\left(\kappa + \tfrac{\|v(u)\|_\infty}{\kappa}\right) T\right) \left(\left\|(1 - \mathcal{P}^\ell)y_0\right\|^2_H + \frac{2}{\kappa}\left\|R^\ell\right\|^2_{L^2(0,T;V')}\right) \tag{37}$$

where the residual R^ℓ is given by (27).

Proof. This is a direct consequence of (26) with the coercivity constants α_u, β_u from (34). □

For error estimation of derivatives, we have to turn to analyzing the concrete cost functions:

5.2 Gradients in $L^2(Q)$

Following Section 3, we would like to analyze a specific cost functions for the system (29):

$$\hat{J}_1(u) := J_1(G(u)) := \frac{1}{2}\int_0^T \int_\Omega (y(t,x) - y_Q(t,x))^2 \, dx \, dt$$

with a desired function $y_Q \in L^2(0,T;L^2(\Omega))$. In order to accurately describe $\hat{J}_1$ as in (18), we introduce the operator

$$\Phi_Q : W(0,T) \to L^2(0,T;L^2(\Omega)), \quad (\Phi_Q y)(t,x) := y(t,x) \quad \text{f.a.a. } t \in (0,T), x \in \Omega$$

which is trivially continuous. Thereby, we can write

$$\hat{J}_1(u) = \tfrac{1}{2}\left\|\Phi_Q G(u) - y_Q\right\|^2_{L^2(0,T;L^2(\Omega))} \quad \text{for all } u \in \mathcal{U}$$

We can therefore accurately define a gradient $\nabla\hat{J}_1 : \mathcal{U} \to \mathcal{U}$ by use of Corollary 3.3 if we know how the adjoint operator $(\Phi_Q G'(u))^*$ looks like. For this, let us define the adjoint equation for an arbitrary $z \in L^2(0,T;L^2(\Omega))$ and $u \in \mathcal{U}$:

$$\begin{aligned} -\langle p_t(t), \varphi\rangle_{V'\times V} + \langle A(u)(t)\varphi, p(t)\rangle_{V'\times V} &= \langle z(t), \varphi\rangle_H \quad \text{for all } \varphi \in V, \text{ f.a.a. } t \in (0,T) \\ p(T) &= 0 \quad \text{in } H \end{aligned} \tag{38}$$

Because of the negative sign in front of the time derivative in (38), we also call this a backwards equation. With identical arguments as for the forward equation, (38) admits a unique solution $p \in W(0,T)$.

5.5 Lemma. *Let v be Fréchet differentiable. Then for arbitrary $u \in \mathcal{U}$, the adjoint operator is given by*

$$(\Phi_Q G'(u))^* \in L(L^2(0,T;L^2(\Omega)), \mathcal{U}), \quad [(\Phi_Q G'(u))^* z](t) = [\mathcal{B}(u)^* \Phi_Q p](t) \text{ in } U, \quad \text{f.a.a. } t \in (0,T) \tag{39}$$

where $p \in W(0,T)$ is the solution to the adjoint equation (38) and $\mathcal{B}$ is given by

$$\mathcal{B} : \mathcal{U} \to L(\mathcal{U}, L^2(0,T;L^2(\Omega))) \quad (\mathcal{B}(u)h)(t,x) := -(v'(u)h)(t,x) \cdot (\nabla G(u))(t,x) \quad \text{f.a.a. } t \in (0,T), x \in \Omega$$

Proof. Let $u, h \in \mathcal{U}$ and $z \in L^2(0,T;L^2(\Omega))$. Then it is

$$\langle (\Phi_Q G'(u))h, z\rangle_{L^2(0,T;L^2(\Omega))} = \int_0^T \int_\Omega (G'(u)h)(t,x) z(t,x) \, dx \, dt$$

We write $y^h := G'(u)h$ and choose $y^h(t) \in V$ as a test function φ in (38):

$$\begin{aligned} ... &= \int_0^T \left[-\langle p_t(t), y^h(t)\rangle_{V'\times V} + \langle A(u)(t) y^h(t), p(t)\rangle_{V'\times V}\right] dt \\ &= \langle p(0), y^h(0)\rangle_H - \langle p(T), y^h(T)\rangle_H + \int_0^T \left[\langle y^h_t(t), p(t)\rangle_{V'\times V} + \langle A(u)(t) y^h(t), p(t)\rangle_{V'\times V}\right] dt \end{aligned}$$

where we have utilized the formula of partial integration for functions in $W(0,T)$, compare Zeidler (1990). Next, using $y^h(0) = p(T) = 0$ along with using $p(t)$ as a test function φ in (13), we obtain by writing $\bar{y} := G(u)$:

$$\begin{aligned} ... &= -\int_0^T \langle (A'(u)h)(t)\bar{y}(t), p(t)\rangle_{V'\times V}\ dt \overset{(35)}{=} -\int_0^T \langle (v'(u)h)(t)\cdot\nabla\bar{y}(t), p(t)\rangle_H\ dt \\ &= \langle (\mathcal{B}(u)h), \Phi_Q p\rangle_{L^2(0,T;L^2(\Omega))} = \langle h, \mathcal{B}(u)^*\Phi_Q p\rangle_{\mathcal{U}} \end{aligned}$$

Lastly, it is obvious that $\mathcal{B}(u) \in L(\mathcal{U}, L^2(0,T;L^2(\Omega)))$ and so the Hilbert adjoint $\mathcal{B}(u)^*$ is well-defined. □

5.6 Example. *Assume that we are in the situation of Example 5.3. Then the operator from Lemma 5.5 has, for every $u \in \mathcal{U}$, the adjoint*

$$(\mathcal{B}(u)^* z)_i(t) = -\eta_i'(u_i(t)) \int_\Omega v^i(t,x)\cdot(\nabla G(u))(t,x)\ z(t,x)\ dx \quad \textit{for all } i = 1,...,p,\ \textit{f.a.a. } t \in (0,T)$$

Therefore, the functional $\hat{J}_1$ has the gradient

$$(\nabla\hat{J}_1(u))_i(t) = -\eta_i'(u_i(t)) \int_\Omega v^i(t,x)\cdot(\nabla G(u))(t,x)\ p(t,x)\ dx \quad \textit{for all } i = 1,...,p,\ \textit{f.a.a. } t \in (0,T) \tag{40}$$

where $p \in W(0,T)$ is the solution to the adjoint equation (38) with $z = \Phi_Q G(u) - y_Q$.

5.7 Remark. *Assume that the control space was given by U rather than $\mathcal{U} = L^\infty(0,T;U)$, meaning the controls would be constant in time. In this case we define the mapping*

$$\Psi : U \to \mathcal{U}, \quad (\Psi u)(t) := u \quad \textit{f.a.a. } t \in (0,T)$$

It is quite obvious that $\Psi \in L(U,\mathcal{U})$. Let us now consider a cost function defined on U:

$$\hat{\mathcal{K}}_1 : U \to \mathbb{R}, \quad \hat{\mathcal{K}}_1(u) := J_1(\Psi u) = \tfrac{1}{2}\left\|\Phi_Q G(\Psi u) - y_Q\right\|^2_{L^2(0,T;L^2(\Omega))}$$

By the chain rule, $\hat{\mathcal{K}}_1$ is Fréchet differentiable if J_1 is Fréchet differentiable which in turn is the case if v is differentiable. The derivative is then given by $\hat{\mathcal{K}}_1'(u)h = J_1'(\Psi u)(\Psi h)$. This implies the existence of a gradient $\nabla\hat{\mathcal{K}}_1(u) = \Psi^*\nabla J_1(\Psi u)$ for all $u \in U$, where Ψ^* is the Hilbert adjoint of Ψ. In order to compute this, let $u \in U$ and $v \in \mathcal{U}$:

$$\langle \Psi u, v\rangle_{\mathcal{U}} = \int_0^T \langle u, v(t)\rangle_U\ dt = \langle u, \int_0^T v(t)\ dt\rangle_U = \langle u, \Psi^* v\rangle_U$$

Therefore, the gradient is given by $\nabla\hat{\mathcal{K}}_1(u) = \int_0^T \nabla J_1(\Psi u)(t)\ dt$ for all $u \in U$. Inserting this into (40) in the situation of Example 5.3, we obtain the representation:

$$(\nabla\hat{\mathcal{K}}_1(u))_i = -\eta_i'(u_i)\int_0^T \int_\Omega v^i(t,x)\cdot(\nabla G(u))(t,x)\ p(t,x)\ dx\ dt \quad \textit{for all } i = 1,...,p$$

We return to the task of error estimation and start by introducing three preliminary estimators for the following terms:

(i) $\|(\Phi_Q G_\ell'(u))^*\|_{L(L^2(0,T;L^2(\Omega)),\mathcal{U})}$ (Lemma 5.8).

(ii) $\left\|p - p^\ell\right\|_{L^2(0,T;V)}$ where $p, p^\ell \in W(0,T)$ are the full and reduced solutions to the adjoint equation (38) to a right-hand side $z \in L^2(0,T;L^2(\Omega))$ (Lemma 5.9).

(iii) $\|(\Phi_Q(G'(u) - G_\ell'(u)))^* z\|_{\mathcal{U}}$ for an element $z \in L^2(0,T;L^2(\Omega))$ (Lemma 5.10).

Using these three estimates, we will be able to present an error estimator for the reduced gradient (Theorem 5.11).

5.8 Lemma. *Let v be Fréchet differentiable. Then the following estimate holds:*

$$\|(\Phi_Q G'(u))^*\|_{L(L^2(0,T;L^2(\Omega)),\mathcal{U})} \leq \sqrt{C_u} \cdot \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \left(\left\|y^\ell\right\|_{L^2(0,T;V)} + \left\|y - y^\ell\right\|_{L^2(0,T;V)}\right) \tag{41}$$

where the constant is given by

$$C_u = \frac{4}{\kappa^2} \exp\left(\left(\kappa + \tfrac{\|v(u)\|_\infty}{\kappa}\right) T\right) \tag{42}$$

Note that (41) becomes an a-posteriori error estimator if $\left\|y - y^\ell\right\|$ if further estimated by the right-hand side of (37).

Proof. For any $z \in L^2(0,T;L^2(\Omega))$, we have

$$\|(\Phi_Q G'(u))^* z\|_{\mathcal{U}} \overset{(39)}{=} \|B(u)^* \Phi_Q p\|_{\mathcal{U}} \leq \|B(u)\|_{L(\mathcal{U},L^2(0,T;L^2(\Omega)))} \cdot \|\Phi_Q p\|_{L^2(0,T;L^2(\Omega))}$$

We continue individually. First of all, we have for any $h \in \mathcal{U}$:

$$\begin{aligned}\|B(u)h\|^2_{L^2(0,T;L^2(\Omega))} &= \|(v'(u)h) \cdot (\nabla \bar{y})\|^2_{L^2(0,T;L^2(\Omega))} \leq \|v'(u)h\|^2_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d))} \cdot \|\nabla \bar{y}\|^2_{L^2(0,T;L^2(\Omega;\mathbb{R}^d))} \\ &\leq \|v'(u)\|^2_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \|\bar{y}\|^2_{L^2(0,T;V)} \cdot \|h\|^2_{\mathcal{U}}\end{aligned} \tag{43}$$

and therefore

$$\|B(u)\|_{L(\mathcal{U},L^2(0,T;L^2(\Omega)))} \leq \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \left(\left\|y^\ell\right\|_{L^2(0,T;V)} + \left\|y - y^\ell\right\|_{L^2(0,T;V)}\right)$$

Second, we obtain

$$\begin{aligned}\|\Phi_Q p\|^2_{L^2(0,T;L^2(\Omega))} \leq \|p\|^2_{L^2(0,T;V)} &\overset{(8)}{\leq} \frac{e^{2\beta_u T}}{\alpha_u}\Big(\underbrace{\|p(T)\|^2_H}_{=0} + \tfrac{1}{\alpha_u}\|z\|^2_{L^2(0,T;H)}\Big) \\ &\overset{(34)}{=} \frac{4}{\kappa^2}\exp\left(\left(\kappa + \tfrac{\|v(u)\|^2_\infty}{\kappa}\right)T\right) \cdot \|z\|^2_{L^2(0,T;H)} \\ &= C_u \cdot \|z\|^2_{L^2(0,T;H)}\end{aligned} \tag{44}$$

All in all, this yields (41). □

Next, we first require a result similar to (37) for the adjoint state p:

5.9 Lemma. *Let v be differentiable and $p, p^\ell \in W(0,T)$ be the full and reduced solutions to the adjoint equation (38). Then the following estimate holds true:*

$$\left\|p(0) - p^\ell(0)\right\|^2_{L^2(\Omega)} + \left\|p - p^\ell\right\|^2_{L^2(0,T;V)} \leq C_u^2 \left\|\mathcal{S}^\ell\right\|^2_{L^2(0,T;V')} \tag{45}$$

with the constant C_u from (42) and $\mathcal{S}^\ell \in L^2(0,T;V')$ given by

$$\langle \mathcal{S}^\ell(t), \varphi \rangle_{V' \times V} = -\langle p^\ell_t(t), \varphi \rangle_H + \langle A(u)(t)\varphi, p^\ell(t) \rangle_{V' \times V} - z(t) \quad \text{for all } \varphi \in V$$

Proof. We define $e := p - p^\ell \in W(0,T)$ and observe that for all $\varphi \in V$ and almost all $t \in (0,T)$, it holds

$$\begin{aligned}-\langle e_t(t), \varphi \rangle_{V' \times V} + \langle A(u)(t)\varphi, p(t) \rangle_{V' \times V} &= \langle z(t), \varphi \rangle_H + \langle p^\ell_t(t), \varphi \rangle_{V' \times V} - \langle A(u)(t)\varphi, p^\ell(t) \rangle_{V' \times V} \\ &= -\langle \mathcal{S}^\ell(t), \varphi \rangle_{V' \times V}\end{aligned}$$

as well as $e(T) = 0$. Applying (8), this directly results in (45). □

5.10 Lemma. *Let v be Fréchet differentiable. Then the following a-posteriori estimate holds for all $z \in L^2(0,T;L^2(\Omega))$:*

$$\|(\Phi_Q G'(u) - \Phi_Q G'_\ell(u))^* z\|_{\mathcal{U}} \le \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \Big(\sqrt{C_u}\,\big\|y - y^\ell\big\|_{L^2(0,T;V)} \cdot \|z\|_{L^2(0,T;L^2(\Omega))} + \big\|y^\ell\big\|_{L^2(0,T;V)} \cdot \big\|p - p^\ell\big\|_{L^2(0,T;V)}\Big) \tag{46}$$

where the constant $C_u > 0$ is given by (42) and p, p^ℓ are full and reduced solutions to the adjoint equation (38).

Proof. We utilize (39) and obtain:

$$\begin{aligned}
\big\|(\Phi_Q G'(u) - \Phi_Q G'_\ell(u))^* z\big\|_{\mathcal{U}} &= \big\|B(u)^*\Phi_Q p - B^\ell(u)^*\Phi_Q p^\ell\big\|_{\mathcal{U}} \\
&\le \big\|B(u) - B^\ell(u)\big\|_{L(\mathcal{U},L^2(0,T;L^2(\Omega)))} \cdot \|\Phi_Q p\|_{L^2(0,T;L^2(\Omega))} \\
&\quad + \big\|B^\ell(u)\big\|_{L(\mathcal{U},L^2(0,T;L^2(\Omega)))} \cdot \big\|\Phi_Q(p - p^\ell)\big\|_{L^2(0,T;L^2(\Omega))} \\
&\le \big\|B(u) - B^\ell(u)\big\|_{L(\mathcal{U},L^2(0,T;L^2(\Omega)))} \cdot \|p\|_{L^2(0,T;V)} \\
&\quad + \big\|B^\ell(u)\big\|_{L(\mathcal{U},L^2(0,T;L^2(\Omega)))} \cdot \big\|p - p^\ell\big\|_{L^2(0,T;V)}
\end{aligned}$$

We continue to estimate the occuring terms:

(1) For any $h \in \mathcal{U}$, we obtain

$$\begin{aligned}
\big\|B(u)h - B^\ell(u)h\big\|_{L^2(0,T;L^2(\Omega))} &= \big\|(v'(u)h) \cdot (\nabla y - \nabla y^\ell)\big\|_{L^2(0,T;L^2(\Omega))} \\
&\le \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \|h\|_{\mathcal{U}} \cdot \big\|y - y^\ell\big\|_{L^2(0,T;V)}
\end{aligned}$$

and $\big\|y - y^\ell\big\|$ can be further estimated by (26).

(2) We have by the use of (5) for the adjoint equation (38):

$$\|p\|^2_{L^2(0,T;V)} \le \frac{4}{\kappa^2}\exp\left(\left(\kappa + \tfrac{\|v(u)\|_\infty^2}{\kappa}\right)T\right)\|z\|^2_{L^2(0,T;L^2(\Omega))} = C_u\,\|z\|^2_{L^2(0,T;L^2(\Omega))}$$

(3) Exactly as in (1), it follows that

$$\big\|B^\ell(u)\big\|_{L(\mathcal{U},L^2(0,T;L^2(\Omega)))} \le \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \big\|y^\ell\big\|_{L^2(0,T;V)}$$

Applying all these estimates results in (46). □

Finally, we are ready to present an estimator for the gradient:

5.11 Theorem. *Let v be Fréchet differentiable. Then for every $u \in \mathcal{U}$, it holds that*

$$\begin{aligned}
\Big\|\nabla\hat{J}(u) - \nabla\hat{J}^\ell(u)\Big\|_{\mathcal{U}} &\le \sqrt{C_u} \cdot \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \big\|y - y^\ell\big\|_{L^2(0,T;V)} \cdot \\
&\quad \left(\sqrt{2\hat{J}^\ell(u)} + \big\|y^\ell\big\|_{L^2(0,T;V)} + \big\|y - y^\ell\big\|_{L^2(0,T;V)}\right) \\
&\quad + \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \big\|y^\ell\big\|_{L^2(0,T;V)} \cdot \big\|p - p^\ell\big\|_{L^2(0,T;V)}
\end{aligned} \tag{47}$$

Proof. We start by estimating:

$$\begin{aligned}
\Big\|\nabla\hat{J}(u) - \nabla\hat{J}^\ell(u)\Big\|_{\mathcal{U}} &\overset{(22)}{=} \|(\Phi_Q G'(u))^*(\Phi_Q G(u) - y_Q) - (\Phi_Q G'_\ell(u))^*(\Phi_Q G_\ell(u) - y_Q)\|_{\mathcal{U}} \\
&= \Big\|\Big[\Phi_Q G'(u) - \Phi_Q G'_\ell(u)\Big]^*(\Phi_Q G_\ell(u) - y_Q) + \big[\Phi_Q G'(u)\big]^*(\Phi_Q G(u) - \Phi_Q G_\ell(u))\Big\|_{\mathcal{U}} \\
&\le \Big\|\Big[\Phi_Q G'(u) - \Phi_Q G'_\ell(u)\Big]^*(\Phi_Q G_\ell(u) - y_Q)\Big\|_{\mathcal{U}} \\
&\quad + \|(\Phi_Q G'(u))^*\|_{L(L^2(0,T;L^2(\Omega)),\mathcal{U})} \cdot \big\|\Phi_Q y - \Phi_Q y^\ell\big\|_{L^2(0,T;L^2(\Omega))}
\end{aligned}$$

We apply the estimates (46) for $z = \Phi_Q G(u) - y_Q$ and (41) to obtain

$$
\begin{aligned}
... \leq & \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \left(\sqrt{C_u} \cdot \left\|y - y^\ell\right\|_{L^2(0,T;V)} \cdot \left\|y^\ell - y_Q\right\|_{L^2(0,T;L^2(\Omega))}\right. \\
& \left. + \left\|y^\ell\right\|_{L^2(0,T;V)} \cdot \left\|p - p^\ell\right\|_{L^2(0,T;V)}\right) \\
& + \sqrt{C_u} \cdot \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \left(\left\|y^\ell\right\|_{L^2(0,T;V)} + \left\|y - y^\ell\right\|_{L^2(0,T;V)}\right) \cdot \left\|y - y^\ell\right\|_{L^2(0,T;V)}
\end{aligned}
$$

which is identical to (47). □

5.3 Gradients in $L^2(\Omega)$ at T

The second specific cost function will be given by

$$\hat{J}_2(u) := J_2(G(u)) := \frac{1}{2}\int_\Omega (y(T,x) - y_T(x))^2 \ dx$$

with a desired function $y_T \in L^2(\Omega)$. Again, we introduce an operator

$$\Phi_T : W(0,T) \to L^2(\Omega), \quad (\Phi_T y)(x) := y(T,x) \quad \text{f.a.a. } x \in \Omega$$

Obviously, Φ_T is linear. It is furthermore continuous and well-defined since $W(0,T) \hookrightarrow C([0,T];H)$, a property which all Gelfand triples possess, see Evans (2008). Therefore, we have

$$\|\Phi_T y\|_{L^2(\Omega)} = \|y(T)\|_H \leq \max_{t\in[0,T]} \|y(t)\|_H \leq C \|y\|_{W(0,T)}$$

with the constant C from the continuous embedding. We have shown $\Phi_T \in L(W(0,T), L^2(\Omega))$ and can write the cost function as

$$\hat{J}_2(u) = \tfrac{1}{2} \|\Phi_T G(u) - y_T\|^2_{L^2(\Omega)} \quad \text{for all } u \in \mathcal{U}$$

Similarily as in the previous section, we are interested in the gradient $\nabla\hat{J}_2 : \mathcal{U} \to \mathcal{U}$. The essential part will again be the adjoint operator $(\Phi_T G'(u))^*$. We define the alternative adjoint equation for an element $z \in L^2(\Omega)$:

$$
\begin{aligned}
-\langle p_t(t), \varphi\rangle_{V'\times V} + \langle A(u)(t)\varphi, p(t)\rangle_{V'\times V} &= 0 \quad \text{for all } \varphi \in V, \text{ f.a.a. } t \in (0,T) \\
p(T) &= z \quad \text{in } H
\end{aligned}
\tag{48}
$$

5.12 Lemma. *Let v be Fréchet differentiable. Then for arbitrary $u \in \mathcal{U}$, the adjoint operator is given by*

$$(\Phi_T G'(u))^* \in L(L^2(\Omega), \mathcal{U}), \quad [(\Phi_T G'(u))^* z](t) = [\mathcal{B}(u)^* \Phi_Q p](t) \text{ in } U, \text{ f.a.a. } t \in (0,T)$$

Proof. For $u, h \in \mathcal{U}$ and $z \in L^2(\Omega)$, we observe

$$\langle(\Phi_T G'(u))h, z\rangle_{L^2(\Omega)} = \int_\Omega (G'(u)h)(T,x)z(x) \ dx = \int_\Omega y^h(T,x)p(T,x) \ dx = \langle y^h(T), p(T)\rangle_H$$

Again by using the formular of partial integration, this results in

$$
\begin{aligned}
... &= \langle\underbrace{y^h(0)}_{=0}, p(0)\rangle + \int_0^T \langle p_t(t), y^h(t)\rangle_H + \langle y^h_t(t), p(t)\rangle_H \ dt \\
&\overset{(13)\&(48)}{=} \int_0^T \langle A(u)(t)y^h(t), p(t)\rangle_{V'\times V} - \langle A(u)(t)y^h(t), p(t)\rangle_{V'\times V} - \langle(A'(u)h)(t)\bar{y}(t), p(t)\rangle_{V'\times V} \ dt \\
&= -\int_0^T \langle(A'(u)h)(t)\bar{y}(t), p(t)\rangle_{V'\times V} \ dt \overset{(35)}{=} -\int_0^T \langle(v'(u)h)(t) \cdot \nabla\bar{y}(t), p(t)\rangle_H \ dt \\
&= \langle\mathcal{B}(u)h, \Phi_Q p\rangle_{L^2(0,T;L^2(\Omega))} = \langle h, \mathcal{B}(u)^*\Phi_Q p\rangle_{\mathcal{U}}
\end{aligned}
$$

In the second equality, we have used $p(t)$ as a test function φ in (13) and $y^h(t)$ as a test function φ in (48). It was already shown in Lemma 5.5 that $\mathcal{B}(u)$ is linear and continuous. □

5.13 Example. *Assume that we are in the situation of Example 5.3. Then the function $\hat{J}_2$ has the gradient*

$$(\hat{J}_2(u))_i(t) = -\eta_i'(u_i(t)) \int_\Omega v^i(t,x) \cdot (\nabla G(u))(t,x) p(t,x) \; dx \quad \text{for all } i = 1,..,p, \; \text{f.a.a. } t \in (0,T)$$

where $p \in W(0,T)$ is the solution to the adjoint equation (48) with $z = \Phi_T G(u) - y_T$.

Proof. Starting from Lemma 5.12, we have inserted the adjoint representation for $(\Phi_Q G'(u))^*$ from Lemma 5.6. □

5.14 Remark. *Assume again that the control space is given by U instead of $\mathcal{U}$. We define the cost function for time-independent controls:*

$$\hat{\mathcal{K}}_2 : U \to \mathbb{R}, \quad \hat{\mathcal{K}}_2(u) := J_2(\Psi u) = \tfrac{1}{2} \left\| \Phi_T G(\Psi u) - y_T \right\|^2_{L^2(\Omega)}$$

Similarily as in Remark 5.7, if v is continuously differentiable, then $\hat{\mathcal{K}}_2$ has a gradient which is, in case of the Example 5.3, given by

$$(\hat{\mathcal{K}}_2(u))_i = -\eta_i'(u_i) \int_0^T \int_\Omega v^i(t,x) \cdot (\nabla G(u))(t,x) \; p(t,x) \; dx \; dt \quad \text{for all } i = 1, ..., p$$

Again, we return to a-posteriori error estimation. For the objective function $\hat{J}_1$, the Lemmata 5.8, 5.9 and 5.10 where essential to derive an a-posteriori estimator for the gradient $\nabla \hat{J}_1$. Only few things change when considering the objective function $\hat{J}_2$ instead, so we will wrap these up in a single lemma:

5.15 Lemma. *Let v be Fréchet differentiable. Then the following estimates hold true:*

a) For the pendant to (41), we have

$$\|(\Phi_T G'(u))^*\|_{L(L^2(0,T;L^2(\Omega)),\mathcal{U})} \le \sqrt{\widetilde{C}_u} \cdot \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \left(\left\|y^\ell\right\|_{L^2(0,T;V)} + \left\|y - y^\ell\right\|_{L^2(0,T;V)} \right) \tag{49}$$

where the constant $\widetilde{C}_u$ is given by

$$\widetilde{C}_u = \frac{2}{\kappa} \exp\left(\left(\kappa + \tfrac{\|v(u)\|_\infty^2}{\kappa} \right) T \right) \tag{50}$$

b) The estimate (45) changes to

$$\left\|p(0) - p^\ell(0)\right\|_H^2 + \left\|p - p^\ell\right\|^2_{L^2(0,T;V)} \le \widetilde{C}_u \left\|(1 - \mathcal{P}^\ell) z\right\|^2_{L^2(\Omega)} + C_u \left\|\mathcal{S}^\ell\right\|^2_{L^2(0,T;V')} \tag{51}$$

where $S^\ell \in L^2(0,T;V')$ given by

$$\langle \mathcal{S}^\ell(t), \varphi \rangle_{V' \times V} = -\langle p_t^\ell(t), \varphi \rangle_H + \langle A(u)(t)\varphi, p^\ell(t) \rangle_{V' \times V} \quad \text{for all } \varphi \in V \tag{52}$$

c) (46) takes, for every $z \in L^2(\Omega)$, the new form

$$\begin{aligned} \|(\Phi_T G'(u) - \Phi_T G_\ell'(u))^* z\|_{\mathcal{U}} \le \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \Big(\sqrt{\widetilde{C}_u} \cdot \left\|y - y^\ell\right\|_{L^2(0,T;V)} \cdot \|z\|_{L^2(\Omega)} \\ + \left\|y^\ell\right\|_{L^2(0,T;V)} \cdot \left\|p - p^\ell\right\|_{L^2(0,T;V)} \Big) \end{aligned} \tag{53}$$

Proof. a) Considering the original proof from Lemma 5.8, the only thing that changes is that due to the new form of the adjoint equation (48), we get

$$\|p\|^2_{L^2(0,T;V)} \overset{(8)}{\le} \frac{e^{2\beta_u T}}{\alpha_u} \left(\|p(T)\|^2_{L^2(\Omega)} + \frac{1}{\alpha_u} \cdot 0 \right) = \frac{e^{2\beta_u T}}{\alpha_u} \|z\|^2_{L^2(\Omega)} = \widetilde{C}_u \|z\|^2_{L^2(\Omega)}$$

which then results in (49).

b) The proof from Lemma 5.9 differs only in the fact that now, we have $p(T) = z$ and $p^\ell(T) = \mathcal{P}^\ell z$ where $\mathcal{P}^\ell$ is a projection onto V^ℓ. This results in (45).

c) The thing that differs from Lemma 5.10 is again the fact that

$$\|p\|^2_{L^2(0,T;V)} \leq \widetilde{C}_u \|z\|^2_{L^2(\Omega)}$$

which then results in (53).

□

5.16 Theorem. *Let v be Fréchet differentiable. Then for every $u \in \mathcal{U}$, it holds that*

$$\begin{aligned}\left\|\nabla \hat{J}(u) - \nabla \hat{J}^\ell(u)\right\|_{\mathcal{U}} \leq \sqrt{\widetilde{C}_u} \cdot \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \left\|y - y^\ell\right\|_{L^2(0,T;V)} \cdot \\ \left(\sqrt{2\hat{J}^\ell(u)} + \left\|y^\ell\right\|_{L^2(0,T;V)} + \left\|y - y^\ell\right\|_{L^2(0,T;V)}\right) \\ + \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \left\|y^\ell\right\|_{L^2(0,T;V)} \cdot \left\|p - p^\ell\right\|_{L^2(0,T;V)}\end{aligned} \tag{54}$$

Proof. Identically to Theorem 5.11, we obtain

$$\begin{aligned}\left\|\nabla \hat{J}(u) - \nabla \hat{J}^\ell(u)\right\|_{\mathcal{U}} \leq \left\|\left[\Phi_T G'(u) - \Phi_T G'_\ell(u)\right]^* (\Phi_T G_\ell(u) - y_T)\right\|_{\mathcal{U}} \\ + \|(\Phi_T G'(u))^*\|_{L(L^2(0,T;L^2(\Omega)),\mathcal{U})} \cdot \left\|\Phi_T y - \Phi_T y^\ell\right\|_{L^2(0,T;L^2(\Omega))}\end{aligned}$$

We apply the estimates (53) for $z = \Phi_T G(u) - y_T$ and (49) to obtain

$$\begin{aligned}... \leq \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \Big(\sqrt{\widetilde{C}_u} \cdot \left\|y - y^\ell\right\|_{L^2(0,T;V)} \cdot \left\|y^\ell(T) - y_T\right\|_{L^2(0,T;L^2(\Omega))} \\ + \left\|y^\ell\right\|_{L^2(0,T;V)} \cdot \left\|p - p^\ell\right\|_{L^2(0,T;V)}\Big) \\ + \sqrt{\widetilde{C}_u} \cdot \|v'(u)\|_{L(\mathcal{U},L^\infty(0,T;L^\infty(\Omega;\mathbb{R}^d)))} \cdot \left(\left\|y^\ell\right\|_{L^2(0,T;V)} + \left\|y - y^\ell\right\|_{L^2(0,T;V)}\right) \cdot \left\|y - y^\ell\right\|_{L^2(0,T;V)}\end{aligned}$$

which is identical to (54). □

6 Conclusion

In the present paper, we consider general linear, inhomogeneous parabolic system whose bilinear form and inhomogeneity are controlled by a possibly time-dependent term from an arbitrary Hilbert space. We show that under certain assumptions, there is a unique solution for every control, enabling us to define a solution operator. However, allowing a control influence in the bilinear form complicates the analysis compared to the case in which only the inhomogeneity is controlled. Even if the bilinear form depends linearly on the control, the solution of the PDE does not. Therefore, the differentiability of the solution operator as well as the representation of the derivative can only be shown under certain additional assumptions (see Lemma 2.5).

In this context, we introduce the concept of model order reduction and show an a-posteriori estimate that gives us an upper bound to the error between the solution of the full system and the reduced system (see Theorem 4.3).

This analysis allows us to consider these kinds of PDEs in the context of optimal control problems, where the PDE serves as a side condition. We introduce a standard quadratic cost function and present a representation of the gradient. It contains an adjoint operator that is in general not easy to evaluate.

We show that the application of model order reduction can be justified by providing an a-posteriori error estimate for the cost function, giving us an upper bound on the error between the cost function and the reduced cost function (see Corollary 4.4).

In an example, we consider a source-free homogeneous advection-diffusion equation with homogeneous Neumann boundary condition, in which the control function acts on the advection term. We look at two different typical cost functions: The first one measures the L^2-norm on the whole time- and spatial domain of the difference between the solution of the underlying PDE and a given desired function, whereas the second one measures the L^2-norm only in the spatial domain of the difference between the solution of the PDE evaluated at the end time-point and a given function.
For both cost functions and this concrete PDE, we develop a representation of the gradient, which is easier to evaluate by using the adjoint equation. With this representation, we are able to show further a-posteriori estimates on the gradient of both cost functions (see Theorems 5.11 and 5.16).
In total, we can conclude that allowing control terms also in the bilinear form complicates the analysis of the PDE and its solution operator in the context of optimal control. However, we were able to show on a concrete example not only that the analysis is possible, but even that the application of model-order reduction schemes is justified by a-posteriori estimates both on the cost function and its gradient. All these things combined give us enough theoretical backing to tackle such an optimal control problem numerically, which shall be approached in future work.

Bibliography

L. C. Evans. *Partial differential equations.* American Mathematical Society, 2008.

M. Gubisch and S. Volkwein. Proper orthogonal decomposition for linear-quadratic optimal control. *P. Benner, A. Cohen, M. Ohlberger, and K. Willcox (eds.), Model Reduction and Approximation: Theory and Algorithms. SIAM, Philadelphia, PA*, 2016.

J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations.* Springer Briefs in Mathematics, 2016.

M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints.* Springer, 2009.

P. Holmes, J. Lumley, G. Berkooz, and C. Rowley. *Turbulence, coherent Structures, dynamical systems and symmetry.* Cambridge Monographs on Mechanics, Cambridge University Press, second edition, 2012.

V. Isakov. *Inverse problems for partial differential equations.* Springer, 2006.

A. T. Patera and G. Rozza. *Reduced basis approximation and a posteriori error estimation for parametrized partial differential equations.* MIT Pappalardo Graduate Monographs in Mechanical Engineering, 2006.

A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations.* Springer, 2016.

F. Tröltzsch. *Optimal Control of partial differential equations: theory, methods and applications.* American Mathematical Society, 2010.

C. Vogel. Sparse matrix computations arising in distributed parameter identification. *SIAM Journal on matrix analysis and applications, 07/1999, Vol. 20, No.4*, 1999.

E. Zeidler. *Nonlinear Functional Analysis and its Application II/A: Linear Monotone Operators.* Springer, 1990.