Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstraße 6 • 53115 Bonn • Germany
phone +49 228 73-3427 • fax +49 228 73-7527
www.ins.uni-bonn.de

L. Bonifacius, I. Neitzel

Second order optimality conditions for optimal control of quasilinear parabolic equations

INS Preprint No. 1705

March 2017

SECOND ORDER OPTIMALITY CONDITIONS FOR OPTIMAL CONTROL OF QUASILINEAR PARABOLIC EQUATIONS

Lucas Bonifacius
Technische Universität München
Fakultät für Mathematik
Boltzmannstr. 3, 85748 Garching, Germany

Ira Neitzel
Rheinische Friedrich-Wilhelms-Universität Bonn
Institut für Numerische Simulation
Wegelerstr. 6, 53115 Bonn, Germany

Abstract. We discuss an optimal control problem governed by a quasilinear parabolic PDE including mixed boundary conditions and Neumann boundary control, as well as distributed control. Second order necessary and sufficient optimality conditions are derived. The latter leads to a quadratic growth condition without two-norm discrepancy. Furthermore, maximal parabolic regularity of the state equation in Bessel-potential spaces $H_D^{-\zeta,p}$ with uniform bound on the norm of the solution operator is proved and used to derive stability results with respect to perturbations of the nonlinear differential operator.

1. Introduction

This article is concerned with optimal control problems governed by quasilinear parabolic partial differential equations (PDEs). Our goal is twofold. First, we aim at establishing second order optimality conditions with minimal gap between necessary and sufficient conditions. Second, we are interested in the precise regularity of the state equation, which is crucial for, e.g., Lipschitz stability estimates. We use the theory to study perturbations of the problem with respect to the nonlinearity, where we rely both on second order optimality conditions and the improved regularity. The prototypical problem with control q and state u is

(1.1a) $$\text{Minimize } J(u,q) \coloneqq \frac{1}{2}\|u-\hat{u}\|^2_{L^2((0,T)\times\Omega)} + \frac{\lambda}{2}\|q\|^2_{L^2(\Lambda,\varrho)},$$

(1.1b) $$\begin{aligned} \partial_t u + \mathcal{A}(u)u &= Bq \quad \text{in } (0,T)\times\Omega, \\ u(0) &= u_0 \qquad \text{in } \Omega, \end{aligned}$$

(1.1c) $$q \in Q_{ad} \subset Q,$$

where

$$\mathcal{A}(u) = -\nabla\cdot\xi(u)\mu\nabla$$

with ξ being a scalar function and μ a spatially dependent coefficient function. Boundary conditions are implicitly included in the definition of the differential operator in (1.1b); see Section 2 for the precise assumptions. The optimal control of

Key words and phrases. Optimal control, second order optimality conditions, quasilinear parabolic partial differential equation, nonautonomous equation, maximal parabolic regularity.

The first author is supported by the International Research Training Group IGDK, funded by the German Science Foundation (DFG) and the Austrian Science Fund (FWF).

quasilinear parabolic PDEs of this type has many important applications, for example heat conduction in electrical engineering [44] and semiconductors [55]. As we will see, both distributed control in two and three spatial dimensions as well as Neumann boundary control in two dimensions is included in this setting. Boundary control in three dimensions can be considered for purely time-dependent controls.

In the literature there are many contributions to optimal control of nonlinear parabolic equations that may be distinguished by the differential operator being either monotone or nonmonotone. Existence of solutions to optimal control problems governed by parabolic equations of *monotone* type has been proved in [46]. Necessary optimality conditions have further been established in [3, 52, 60]; cf. also the introduction in [30]. Less abstract hypotheses have been used in [13] to show first order optimality conditions. Concerning optimal control of *nonmonotone* parabolic equations, fewer results have been published. Existence of solutions has been considered in [49, 50] for distributed control. The studies impose pointwise control constraints and the control enters nonlinearly in the state equation. More recently, first order necessary conditions for a quasilinear equation subject to integral state constraints have been proved in [30] with distributed controls in $L^2((0,T)\times\Omega)$. It is worth mentioning, that in the latter study all the coefficients of the elliptic operator may depend on u, ∇u and the control q. However, the derivatives of the coefficients have to satisfy certain growth bounds, so that our prototype problem does not comply with assumption (2.5) in [30]. To our best knowledge, there are no published results on second order optimality conditions for quasilinear parabolic equations. However, optimal control of semilinear parabolic equations even with pointwise state constraint is well-investigated; see for instance [18, 19, 43, 54]. Concerning the case of quasilinear elliptic equations, first- and second order optimality conditions have been established in [12, 15].

Recently, uniform Hölder estimates for linear parabolic equations subject to mixed boundary conditions and rough domains have been established in [48], which in turn implies that the state belongs to

$$W^{1,s}((0,T);W_D^{-1,p})\cap L^s((0,T);W_D^{1,p})\hookrightarrow_c C^\alpha((0,T);C^\kappa(\Omega)),$$

for any right-hand side in $L^s((0,T);W_D^{-1,p})$. This is the starting point for our investigation. Adapting the ideas of Casas and Tröltzsch from [15, 17], we prove second order necessary as well as sufficient optimality conditions for Neumann boundary control in spatial dimension two and purely time-dependent control and distributed control in dimensions two and three. The main difficulty is that this requires the first and second derivative of the reduced objective functional to be extended to $L^2(\Lambda,\varrho)$, but the linearized state equation contains an additional term involving the gradient of u (cf. Proposition 4.4). We overcome this issue by a careful regularity analysis of the state equation based on the results in [22].

Moreover, in applications it is often required to guarantee uniform boundedness of the solution operator to certain linearized-type equations, cf. Lemma 5.1. Once existence of a solution to the nonlinear state equation is established in an appropriate function space, improved regularity results can be transferred from the linear to the nonlinear setting by plugging the solution into the nonlinear differential operator and applying linear regularity theory. Then, however, it is crucial to track the explicit dependence of, e.g., the constants in all appearing stability estimates on the solution u, or $\mathcal{A}(u)$, respectively. We eventually prove that the time-dependent operator $-\nabla\cdot\xi(u)\mu\nabla$ exhibits maximal parabolic regularity not only on $L^s((0,T);W_D^{-1,p})$ but on $L^s(I;H_D^{-\zeta,p}(\Omega))$. Whence, the state belongs to

$$W^{1,s}((0,T);H_D^{-\zeta,p}(\Omega)) \cap L^s((0,T);\mathcal{D}_{H_D^{-\zeta,p}(\Omega)}(-\nabla\cdot\mu\nabla)) \hookrightarrow_c C^\alpha((0,T);W_D^{1,p}(\Omega)),$$

where $H_D^{-\zeta,p}(\Omega)$ is the Bessel-potential space that can be obtained by complex interpolation between $L^p(\Omega)$ and $W_D^{-1,p}(\Omega)$, and $\mathcal{D}_{H_D^{-\zeta,p}(\Omega)}(-\nabla\cdot\mu\nabla)$ is the domain of $-\nabla\cdot\mu\nabla$ considered on $H_D^{-\zeta,p}(\Omega)$. Indeed, the norm of the solution in the space of maximal regularity can be explicitly estimated in terms of the problem data. Of course the compact embedding yields uniformity of the solution operator to the linearized equation. This functional analytic setting has been proposed in [38] and covers rough domains and mixed boundary conditions. Furthermore, the space $H_D^{-\zeta,p}(\Omega)$ allows for distributional objects such as surface charge densities or thermal sources concentrated on hypersurfaces, cf. [38, Theorems 3.6 and 6.9]. We point out that in [38] the authors proved local-in-time existence of solutions in the $H_D^{-\zeta,p}(\Omega)$ setting, but not existence on the whole time interval $(0,T)$.

The paper is organized as follows. In Section 2 we state the precise assumptions of the problem and collect specific examples of control settings that are covered by Problem (1.1). Maximal parabolic regularity on the mentioned space $L^s(I;H_D^{-\zeta,p})$ is proved in Section 3. Section 4 is devoted to the analysis of the optimal control problem including second order necessary and sufficient optimality conditions. In Section 5 we investigate stability of optimal solutions with respect to perturbations on ξ using both the improved regularity and the second order optimality conditions. Some interesting but technical results are collected in the appendix.

2. Notation and assumptions

We now give the precise assumptions concerning the geometry, the operators and the problem data.

2.1. Notation. For $\Omega\subset\mathbb{R}^d$ a Lipschitz domain, $\theta\in(0,1]$, and $p\in(1,\infty)$ we define the space $H_D^{\theta,p}(\Omega)$ as the closure of

$$C_D^\infty(\Omega) = \left\{\psi|_\Omega\colon \psi\in C^\infty(\mathbb{R}^d),\ \operatorname{supp}(\psi)\cap\Gamma_D=\emptyset\right\}$$

in the Bessel-potential space $H^{\theta,p}(\Omega)$, i.e.

$$H_D^{\theta,p}(\Omega) = \overline{C_D^\infty(\Omega)}^{H^{\theta,p}(\Omega)}.$$

If $\theta=1$, then the space $H_D^{\theta,p}(\Omega)$ coincides with the usual Sobolev space that we denote by $W_D^p(\Omega)$. Of course, if $\Gamma_N=\emptyset$, then $H_D^{\theta,p}(\Omega)=H_0^{\theta,p}(\Omega)$ and if $\Gamma_N=\partial\Omega$, then $H_D^{\theta,p}(\Omega)=H^{\theta,p}(\Omega)$. The corresponding dual space of $W_D^{1,p}(\Omega)$ is denoted by $W^{-1,p'}(\Omega)$, and the dual space of $H_D^{\theta,p}(\Omega)$ is denoted by $H_D^{-\theta,p'}(\Omega)$, where p' stands for the conjugate Sobolev exponent, i.e. $1=1/p+1/p'$. If ambiguity is not to be expected, we drop the spatial domain Ω from the notation of the spaces. The domain of a linear (possibly unbounded) operator A on a Banach space X is denoted by $\mathcal{D}_X(A)$. As usual $R(z,A)=(z-A)^{-1}$ denotes the *resolvent* of an operator A.

By the symbol $\mathcal{M}_d(\mu_\bullet,\mu^\bullet)$ we denote the set of measurable mappings $\mu\colon\Omega\to\mathbb{R}^{d\times d}$ having values in the set of real-valued matrices which satisfy the uniform ellipticity condition

$$\mu_\bullet\|z\|^2 \le \sum_{i,j=1}^d \mu_{ij}(x)z_jz_i,\quad z\in\mathbb{R}^d,\text{ a.a. } x\in\Omega,$$

and $\|\mu_{ij}\|_{L^\infty} \le \mu^\bullet$, $i,j = 1,\dots,d$, for some constants $\mu^\bullet, \mu_\bullet > 0$. Last, c is a generic constant that may have different values at different appearances.

2.2. **Assumptions.**

Assumption 1. *Let $\Omega \subset \mathbb{R}^d$ with $d \in \{2,3\}$ be a bounded domain with boundary $\partial\Omega$ and Γ_N is a relatively open subset of $\partial\Omega$ denoting the Neumann boundary part and $\Gamma_D = \partial\Omega \setminus \Gamma_N$ the Dirichlet boundary part. Moreover, we assume that $\Omega \cup \Gamma_N$ is Gröger regular; see Definition A.1. In addition, each mapping ϕ_x in Definition A.1 is volume-preserving. We consider a fixed time interval $I = (0,T)$ with $T > 0$.*

Remark 2.1.
(i) If $\Omega \cup \Gamma_N$ is regular in the sense of Gröger, then Ω is a Lipschitz domain; see [36, Theorem 5.1]. Conversely, if Ω is a Lipschitz domain, then Ω and $\Omega \cup \partial\Omega$ are Gröger regular, cf. Definition 1.2.1.2 in [33].
(ii) For simplified characterizations of regular sets see [36, Theorems 5.2 and 5.4].
(iii) We do not exclude the cases $\Gamma_D = \emptyset$ or $\Gamma_D = \partial\Omega$.
(iv) The additional requirement of volume-preserving bi-Lipschitz transformations is satisfied in many practical situations. In spatial dimension three, two crossing beams allow for a volume-preserving bi-Lipschitz transformation; see Section 7.3 in [38]. In particular, domains with Lipschitz boundary satisfy Assumption 1; see Remark 3.3 in [38].
(v) Note that the assumption of volume-preserving ϕ_x is only used for the existence result [48] and the characterization of $H_D^{-\zeta,p}$ to be an interpolation space between L^p and $W_D^{-1,p}$; see (3.8).

Assumption 2. *Let ξ be real-valued, twice continuously differentiable, and ξ'' Lipschitz continuous on bounded sets. For fixed $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$ define*

$$\langle \mathcal{A}(u)\varphi, \psi\rangle_{L^2(I;W_D^{1,2})} = \int_I \int_\Omega \xi(u)\mu\nabla\varphi\cdot\nabla\psi \,\mathrm{d}x\,\mathrm{d}t, \qquad \varphi,\psi \in L^2(I;W_D^{1,2}),$$

with constants $0 < \mu_\bullet < \mu^\bullet$. Moreover, suppose there are $\xi^\bullet, \xi_\bullet > 0$ such that

$$0 < \xi_\bullet \le \xi(u) \le \xi^\bullet \quad \forall u \in \mathbb{R}.$$

Assumption 3. *There is $p \in (d,4)$ such that*

$$-\nabla\cdot\mu\nabla + 1\colon W_D^{1,p} \to W_D^{-1,p}$$

provides a topological isomorphism.

Remark 2.2. The function ξ might also depend on t, if the dependence is sufficiently smooth (e.g. Lipschitz) and all results hold with obvious modifications.

Remark 2.3. Assumption 3 is a further restriction on the spatial domain, the coefficient function μ, and the boundary conditions. If $d = 2$, then the assumption is always fulfilled for only Gröger regular sets [34, Theorem 1]. Indeed, this is true under less restrictive assumptions on the domain; cf. [37, Theorem 5.6]. In general p exceeds 2 by only an arbitrary small number; see [28, Chapter 4]. Even if the coefficient function and the domain are smooth, $p \ge 4$ cannot be expected in case of mixed boundary conditions; see the example of [56, p. 151].

However, in many practical situations with $d = 3$ the isomorphism property holds with $p > 3$; see [23]. In particular, if Ω is of class C^1, μ is uniformly continuous and $\Gamma_N = \emptyset$, then Assumption 3 holds for all $p \in (1,\infty)$. In the example of two crossing beams, cf. Remark 2.1, if μ is constant on each beam, then Assumption 3 holds with homogeneous Dirichlet boundary conditions.

Last, we define the setting for the controls. To have one notation for different control situations, we define a measure space (Λ, ϱ) and the control operator B mapping $L^\infty(\Lambda, \varrho)$ into $L^s(I; W_D^{-1,p})$. Concrete examples are given below.

Assumption 4. *Let $\zeta \in (0,1)$ be fixed such that $\max\{1-\frac{1}{p}, \frac{d}{p}\} < \zeta$ and choose $s > \frac{2}{\zeta - d/p}, \frac{2}{1-\zeta}$. Given a measure space (Λ, ϱ), define the control space as $Q = L^\infty(\Lambda, \varrho)$ and the set of admissible controls as*

$$Q_{ad} \coloneqq \{q \in Q \colon q_a \leq q \leq q_b \ \varrho\text{-almost everywhere in } \Lambda\}$$

for two fixed elements $q_a, q_b \in Q$, $q_a \leq q_b$ ϱ-almost everywhere. The control operator

$$B \colon L^\infty(\Lambda, \varrho) \to L^s(I; H_D^{-\zeta,p})$$

is linear and bounded. Moreover, B can be continuously extended to an operator

$$B \colon L^2(\Lambda, \varrho) \to L^2(I; W_D^{-1,p}).$$

The desired state $\hat{u}$ and the initial value u_0 satisfy $\hat{u} \in L^\infty(I; L^2)$, and $u_0 \in (H_D^{-\zeta,p}, \mathcal{D}_{H_D^{-\zeta,p}}(-\nabla \cdot \mu \nabla))_{1-\frac{1}{s},s}$, respectively, and $\lambda > 0$ is the regularization (or cost) parameter.

The extension property of the control operator stated in Assumption 4 is only needed for second order sufficient conditions. Therein, we require the linearized equation to be solvable for right-hand sides in $L^2(I; W_D^{-1,p})$. In order to ease readability, we denote the maximal restriction of $-\nabla \cdot \mu \nabla$, $\mathcal{A}(u)$ and B, respectively, to any of the spaces occurring in this article by the same symbol.

Before continuing the analysis, let us state typical situations covered by the general setting of Problem (1.1)

Example 2.4 (Neumann boundary control; $d = 2$)**.** Given $q_a, q_b \in L^\infty(I \times \Gamma_N)$ define the control space and set of admissible controls as

$$Q = L^\infty(I \times \Gamma_N), \qquad Q_{ad} = \{q \in Q \,|\, q_a \leq q \leq q_b \text{ a.e. in } I \times \Gamma_N\}.$$

According to [38, Theorem 6.9] the adjoint of the trace operator Tr^* is continuous from $L^2(\Gamma_N)$ to $H_D^{-\theta,2}$ for $\theta \in (1/2, 1)$. If the embedding $E \colon H_D^{-\theta,2} \to W_D^{-1,p}$ with p from Assumption 3 is continuous, then $B = E \circ \mathrm{Tr}^*$ satisfies Assumption 4. The embedding theorem yields $W_D^{1,p'} \hookrightarrow H_D^{\theta,2}$ if

$$1 - \frac{d}{p'} \geq \theta - \frac{d}{2}, \quad \text{or, equivalently,} \quad 1 - \frac{d}{2} + \frac{d}{p} \geq \theta.$$

For $d = 2$, we require $2/p \geq \theta$. As p does in general not exceed 4 (see Remark 2.3), there is $\theta \in (1/2, 2/p)$ such that B is continuous from $L^2(I \times \Gamma_N)$ to $L^2(I; W_D^{-1,p})$.

Note that for $d = 3$, we have to require $3/p - 1/2 \geq \theta$ and $\theta > 1/2$, but $3/p - 1/2 < 1/2$. This motivates the analysis of purely time dependent controls which are also interesting in practice, since distributed controls are usually difficult to implement; see [19] and references therein for applications.

Example 2.5 (Purely time dependent controls; $d = 2, 3$)**.** Let $e_1, \ldots, e_m \in H_D^{-\zeta,p}$ for ζ as in Assumption 4 be given, define the control operator as

$$B \colon L^\infty(I; \mathbb{R}^m) \to L^s(I; H_D^{-\zeta,p}), \qquad (Bq)(t) = \sum_{i=1}^m q^i(t) e^i.$$

The control space and the space of admissible controls, respectively, are given by

$$Q = L^\infty(I;\mathbb{R}^m), \qquad Q_{ad} = \{q \in Q \mid q_a \le q \le q_b \text{ a.e. in } I\},$$

where $q_a, q_b \in L^\infty(0,T;\mathbb{R}^m)$. The inequality above is understood componentwise. We note that Assumption 4 is satisfied due to the continuous embedding $H_D^{-\zeta,p} \hookrightarrow W_D^{-1,p}$. The measure space (Λ, ϱ) is defined via the product of the Lebesgue measure on I with the counting measure on $\{1,\dots,m\}$.

Example 2.6 (Distributed control; $d = 2,3$)**.** Let $\omega \subseteq \Omega$ and define the control space and the set of admissible controls by

$$Q = L^\infty(I \times \omega), \qquad Q_{ad} = \{q \in Q \mid q_a \le q \le q_b \text{ a.e. in } I \times \Omega\},$$

for fixed $q_a, q_b \in L^\infty(I \times \omega)$. The control operator B is defined as the embedding and extension operator from $L^\infty(I\times\omega)$ into $L^s(I; H_D^{-\zeta,p})$. The Sobolev embedding yields $L^2 \hookrightarrow W_D^{-1,p}$, if $p \le 6$. Take $\Lambda = I \times \omega$, and (Λ, ϱ) is the measure space equipped with the Lebesgue measure.

3. The quasilinear parabolic state equation

We start with the discussion of existence and regularity for the state equation (1.1b). First we introduce the concept of maximal parabolic regularity for nonautonomous operators from [24, Definition 2.1]. Note that if the operator is time-independent, then the definition coincides with the usual notion of maximal parabolic regularity for autonomous equations; cf. [7, Section III.1.5].

Definition 3.1. Let X, Y be Banach spaces such that $Y \hookrightarrow_d X$, and $[0,T] \ni t \mapsto A(t) \in \mathcal{L}(Y,X)$ be a bounded and measurable map. Moreover, $s \in (1,\infty)$ and $A(t)$ is a closed operator in X for each $t \in [0,T]$. Then A is said to satisfy *maximal parabolic regularity* on X, if for every $f \in L^s(I;X)$ and $w_0 \in (X,Y)_{1-1/s,s}$ there exists a unique solution $w \in W^{1,s}(I;X) \cap L^s(I;Y)$ satisfying

$$\partial_t w + Aw = f, \quad w(0) = w_0,$$

where the time derivative is taken in the sense of X-valued distributions on I; see Chapter III.1 in [7].

Proposition 3.2 ([4, Theorem 3], cf. also [22, Lemma 3.4])**.** *Let X, Y be Banach spaces such that $Y \hookrightarrow X$ and $s \in (1,\infty)$. If $\tau \in (1 - \frac{1}{s}, 1)$, then*

$$W^{1,s}(I;X) \cap L^s(I;Y) \hookrightarrow L^r(I;(X,Y)_{\tau,1}), \qquad 1 < r < \frac{s}{1-(1-\tau)s}. \tag{3.1}$$

If $\tau \in (0, 1-\frac{1}{s})$, then

$$W^{1,s}(I;X) \cap L^s(I;Y) \hookrightarrow C^\alpha(I;(X,Y)_{\tau,1}), \qquad 0 \le \alpha < 1 - \frac{1}{s} - \tau. \tag{3.2}$$

If in addition $Y \hookrightarrow_c X$, then the embeddings above are compact as well.

3.1. **Maximal parabolic regularity on $L^s((0,T);W_D^{-1,p})$.** This subsection is devoted to maximal parabolic regularity of the nonautonomous operator $\mathcal{A}(u)$ on $L^s(I;W_D^{-1,p})$. To this end, we first consider the time-independent operator $-\nabla\cdot\mu\nabla$ for an arbitrary coefficient function $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$.

Proposition 3.3. *Let $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$ with $0 < \mu_\bullet < \mu^\bullet$. The operator $-\nabla \cdot \mu \nabla$ exhibits maximal parabolic regularity on $W_D^{-1,p}$. If $1 < s \leq \frac{2p}{p-d}$, then*

$$W^{1,s}(I; W_D^{-1,p}) \cap L^s(I; W_D^{1,p}) \hookrightarrow_c L^r(I; C^\kappa), \quad 1 < r < \frac{2ps}{ds - p(s-2)}, \tag{3.3}$$

for $\kappa = \kappa(r) > 0$ sufficiently small and $\frac{1}{0} := \infty$. Otherwise, if $\frac{2p}{p-d} < s < \infty$, then

$$W^{1,s}(I; W_D^{-1,p}) \cap L^s(I; W_D^{1,p}) \hookrightarrow_c C^\alpha(I; C^\kappa), \tag{3.4}$$

with $\alpha = \frac{1}{2} - \frac{d}{2p} - \frac{1}{s} - \frac{\kappa}{2}$ and $\kappa > 0$ sufficiently small.

Proof. We apply [22, Theorem 4.6]. Note that all regularity requirements of [22] are fulfilled, see Propositions A.2 and A.3, and Kato's square root property holds due to Theorem 4.1 (cf. also Remark 2.4 (3)) in [27]. Hence $-\nabla \cdot \mu \nabla$ exhibits maximal parabolic regularity on $W_D^{-1,p}$ and $-\nabla \cdot \mu \nabla + 1$ is positive. Moreover, according to Lemma 4.8 in [22], we have

$$(W_D^{-1,p}, \mathcal{D}_{W_D^{-1,p}}(-\nabla \cdot \mu \nabla))_{\tau,1} \hookrightarrow C^\kappa, \quad \frac{1}{2} + \frac{d}{2p} + \frac{\kappa}{2} < \tau < 1, \tag{3.5}$$

for $\kappa > 0$ sufficiently small. Both embeddings (3.3) and (3.4) follow from Proposition 3.2 with $X = W_D^{-1,p}$ and $Y = W_D^{1,p}$: For (3.3) we have to guarantee that $1 - 1/s < \tau < 1$. It holds

$$s \leq \frac{2p}{p-d} \quad \Leftrightarrow \quad 1 - \frac{1}{s} \leq \frac{1}{2} + \frac{d}{2p}. \tag{3.6}$$

Hence, taking $1/2 + d/2p < \tau$ yields $1 - 1/s < \tau$. If inequality (3.6) is strict, then we estimate the upper bound for r in (3.1) as

$$r < \frac{s}{1 - (1-\tau)s} < \frac{s}{1 - (\frac{1}{2} - \frac{d}{2p})s} = \frac{2ps}{ds - p(s-2)}.$$

Otherwise, if equality holds in (3.6), we may choose τ arbitrarily close to $1 - 1/s$. In both cases, embedding (3.3) follows from (3.1) and (3.5) with κ sufficiently small.

Concerning the second embedding, we have to ensure that $0 < \tau < 1 - 1/s$. The condition $s > 2p/(p-d)$ is equivalent to $1/2 + d/2p < 1 - 1/s$. Hence, there is $\kappa > 0$ sufficiently small such that

$$\frac{1}{2} + \frac{d}{2p} + \frac{\kappa}{2} < 1 - \frac{1}{s}. \tag{3.7}$$

Taking τ between both values in (3.7), both (3.5) and $0 < \tau < 1 - 1/s$ are satisfied, and we conclude embedding (3.4).

It remains to prove compactness of $W_D^{1,p} \hookrightarrow_c W_D^{-1,p}$. Due to Lemma 3.2 and Remark 3.3 (i) in [11], there is a continuous extension operator mapping $W_D^{1,p}(\Omega)$ into $W^{1,p}(\mathcal{B})$, where $\mathcal{B} \subset \mathbb{R}^d$ is an open ball containing $\overline{\Omega}$. Employing Theorem 2.8.1 in [61], we see that $W^{1,p}(\mathcal{B}) \hookrightarrow_c L^p(\mathcal{B})$. By means of the restriction from $L^p(\mathcal{B})$ into $L^p(\Omega)$, we obtain $W_D^{1,p} \hookrightarrow_c W_D^{-1,p}$. □

Lemma 3.4. *Let $s \in (1, \infty)$ and $u \in C(\overline{I \times \Omega})$ be given. The operator $-\nabla \cdot \xi(u) \mu \nabla$ exhibits maximal parabolic regularity on $L^s(I; W_D^{-1,p})$.*

Proof. According to Proposition 3.3 the operator $-\nabla \cdot \xi(u(t)) \mu \nabla$ has maximal parabolic regularity on $W_D^{-1,p}$ for every $t \in [0, T]$. Since maximal regularity is preserved under relative compact perturbations, see, e.g., [41, Theorem IV. 5.3 5.26], we may add or subtract the embedding $1\colon W^{1,p}_D \to W^{-1,p}_D$ to the operator, so $-\nabla\cdot\xi(u(t))\mu\nabla + 1$ has maximal regularity. Moreover, $-\nabla\cdot\xi(u(t))\mu\nabla + 1$ provides a topological isomorphism between $W^{1,p}_D$ and $W^{-1,p}_D$ for all t due to [23, Lemma 6.2]. In particular, the elliptic operator $-\nabla\cdot\xi(u(t))\mu\nabla + 1$ has constant domain with respect to time. Since the mapping $t \mapsto -\nabla\cdot\xi(u(t))\mu\nabla + 1 \in \mathcal{L}(W^{1,p}_D, W^{-1,p}_D)$ is continuous, maximal parabolic regularity transfers to the time-dependent operator $-\nabla\cdot\xi(u)\mu\nabla + 1$ according to [5, Theorem 7.1]. Subtracting the embedding $1\colon W^{1,p}_D \to W^{-1,p}_D$ again concludes the proof. □

The starting point for our following investigation is the existence result for quasilinear parabolic equations subject to mixed boundary conditions:

Proposition 3.5 ([48, Corollary 5.8]). *Let $q \in L^s(I; W^{-1,p}_D)$, then there exists a unique solution $u \in W^{1,s}(I; W^{1,p}_D) \cap L^s(I; W^{-1,p}_D)$ satisfying the state equation (1.1b).*

Since any Hölder continuous function on $I \times \Omega$ is uniformly continuous, there is a unique uniformly continuous extension to the closure $\overline{I \times \Omega}$. Furthermore, this extension is Hölder continuous with the same exponent. Therefore, the solution of the state equation which exists due to Proposition 3.5 satisfies $u \in C^\alpha([0,T]; C^\kappa(\overline{\Omega}))$ with $\alpha, \kappa > 0$ from Proposition 3.3.

3.2. **Maximal parabolic regularity on $L^s((0,T); H^{-\zeta,p}_D)$.** Given any solution u to the state equation, we consider the nonautonomous operator $\mathcal{A}(u) = -\nabla\cdot\xi(u)\mu\nabla$. To prove maximal parabolic regularity of $\mathcal{A}(u)$ on $L^s(I; H^{-\zeta,p}_D)$, we establish two ingredients: First, we show that $-\nabla\cdot\mu\nabla$ is uniformly $\mathcal{R}$-sectorial with respect to μ; see (3.12). Second, we verify that $\mathcal{A}(u)$ satisfies the Acquistapace-Terreni condition; see (3.19). We consider the operators on L^p and on $W^{-1,p}_D$ separately and obtain the results on $H^{-\zeta,p}_D$ by interpolation. Note that in general we cannot show that $\mathcal{A}(u)$ fulfills the Acquistapace-Terreni condition on L^p, but on the space $H^{-\zeta,p}_D$ that can be expressed by complex interpolation between L^p and $W^{-1,p}_D$. Precisely, we have [32, Theorem 3.1]

$$H^{-\zeta,p}_D = [L^p, W^{-1,p}_D]_\zeta. \tag{3.8}$$

We collect properties of the operator $-\nabla\cdot\mu\nabla$ on $H^{-\zeta,p}_D$ from [38].

Proposition 3.6 (Lemmas 6.6 and 6.7 in [38]). *Let ζ be as in Assumption 4.*

(i) For each $\tau \in (\frac{1+\zeta}{2}, 1)$, there exists a continuous embedding

$$(H^{-\zeta,p}_D, \mathcal{D}_{H^{-\zeta,p}_D}(-\nabla\cdot\mu\nabla))_{\tau,1} \hookrightarrow W^{1,p}_D. \tag{3.9}$$

(ii) The embedding $\mathcal{D}_{H^{-\zeta,p}_D}(-\nabla\cdot\mu\nabla) \hookrightarrow W^{1,p}_D$ holds, and the linear mapping

$$W^{1,p}_D \to \mathcal{L}(\mathcal{D}_{H^{-\zeta,p}_D}(-\nabla\cdot\mu\nabla), H^{-\zeta,p}_D), \quad \xi \mapsto -\nabla\cdot\xi\mu\nabla, \tag{3.10}$$

is continuous.

Combining Proposition 3.2 and (3.9) yields the following embedding, where compactness is due to $\mathcal{D}_{H^{-\zeta,p}_D}(-\nabla\cdot\mu\nabla) \hookrightarrow W^{1,p}_D \hookrightarrow_c L^p \hookrightarrow H^{-\zeta,p}_D$.

Corollary 3.7 ([38, Corollary 6.16]). *If $s > \frac{2}{1-\zeta}$, then there is $\alpha > 0$ such that*

$$W^{1,s}(I; H^{-\zeta,p}_D) \cap L^s(I; \mathcal{D}_{H^{-\zeta,p}_D}(-\nabla\cdot\mu\nabla)) \hookrightarrow_c C^\alpha(I; W^{1,p}_D). \tag{3.11}$$

Let Σ_θ denote the open sector in the complex plane with vertex 0 and opening angle 2θ, which is symmetric with respect to the positive real half-axis, i.e.

$$\Sigma_\theta \coloneqq \{z \in \mathbb{C} \setminus \{0\} \colon |\arg z| < \theta\} .$$

Proposition 3.8. *Let $0 < \mu_\bullet < \mu^\bullet$ and $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$. There exists $\theta \in [\pi/4, \pi/2)$ such that the spectrum of $-\nabla \cdot \mu\nabla + 1$ considered on L^p, $W_D^{-1,p}$, and $H_D^{-\zeta,p}$ is contained in $\overline{\Sigma_\theta}$ uniformly in μ, i.e. $\sigma(-\nabla \cdot \mu\nabla + 1) \subset \overline{\Sigma_\theta}$.*

Proof. Due to [11, Proposition 4.6 (ii), Theorem 11.5 (ii)], the operator $-\nabla\cdot\mu\nabla+1$ admits bounded H^∞-calculi on L^p and $W_D^{-1,p}$ with H^∞-angle

$$\phi^\infty_{-\nabla\cdot\mu\nabla+1} \le \arctan\left(\frac{\|\mu\|_{L^\infty}}{\mu_\bullet(\mu)}\right) \le \arctan\left(\frac{\mu^\bullet}{\mu_\bullet}\right) \in [\pi/4, \pi/2),$$

where $\mu_\bullet(\mu)$ is the coercivity constant of μ. In particular, the spectra are contained in a sector $\overline{\Sigma_\theta}$. Using (3.8) this carries over to the operator considered on $H_D^{-\zeta,p}$. □

3.2.1. *Uniform $\mathcal{R}$-sectoriality in $L^p(\Omega)$.* We first introduce the concept of $\mathcal{R}$-boundedness. For more details we refer to [20, Sections 3 and 4] and [45, Section I.2].

Definition 3.9. Let X and Y be Banach spaces. A family of linear operators $\mathcal{T} \subset \mathcal{L}(X, Y)$ is called *$\mathcal{R}$-bounded*, if for $p \in [1, \infty)$ there is a constant $C > 0$ such that for all $N \in \mathbb{N}$, $T_j \in \mathcal{T}$, $\varphi_j \in X$ and for all independent, symmetric, $\{-1, 1\}$-valued random variables ε_j on a probability space M the inequality

$$\|\sum_{j=1}^N \varepsilon_j T_j \varphi_j\|_{L^p(M;Y)} \le C\|\sum_{j=1}^N \varepsilon_j \varphi_j\|_{L^p(M;X)}$$

holds. The smallest such C is called $\mathcal{R}$-bound of $\mathcal{T}$ and denoted by $\mathcal{R}(\mathcal{T})$.

Remark 3.10. The $\mathcal{R}$-bound has the following properties.

(i) If $\mathcal{T} \subset \mathcal{L}(X, Y)$ is $\mathcal{R}$-bounded, then it is uniformly bounded in $\mathcal{L}(X, Y)$ with bound $\sup_{T\in\mathcal{T}} \|T\|_{\mathcal{L}(X,Y)} \le \mathcal{R}(\mathcal{T})$.

(ii) If $\mathcal{T} \subset \mathcal{L}(X, Y)$ is $\mathcal{R}$-bounded for some $p \in [1, \infty)$, then it is $\mathcal{R}$-bounded for every $p \in [1, \infty)$ due to the inequality of Kahane. However, the $\mathcal{R}$-bound depends on p.

(iii) Let X, Y be Banach spaces and $\mathcal{S}, \mathcal{T} \subset \mathcal{L}(X, Y)$ be $\mathcal{R}$-bounded, then the set $\mathcal{S}+\mathcal{T} = \{S+T \colon S \in \mathcal{S},\, T \in \mathcal{T}\}$ is $\mathcal{R}$-bounded with $\mathcal{R}(\mathcal{S}+\mathcal{T}) \le \mathcal{R}(\mathcal{S})+\mathcal{R}(\mathcal{T})$.

(iv) Let X, Y, Z be Banach spaces. If $\mathcal{T} \subset \mathcal{L}(X, Y)$ and $\mathcal{S} \subset \mathcal{L}(Y, Z)$ are both $\mathcal{R}$-bounded, then the set $\mathcal{S}\mathcal{T} = \{\, ST \colon S \in \mathcal{S},\, T \in \mathcal{T}\,\}$ is $\mathcal{R}$-bounded and it holds $\mathcal{R}(\mathcal{S}\mathcal{T}) \le \mathcal{R}(\mathcal{S})\mathcal{R}(\mathcal{T})$.

An operator A on a Banach space X is called *$\mathcal{R}$-sectorial* of angle $\theta \in (0, \pi/2)$ if

$$\mathcal{R}\left(\left\{zR(z, A) \colon z \in \mathbb{C} \setminus \overline{\Sigma_\theta}\right\}\right) < \infty. \tag{3.12}$$

To prove uniform $\mathcal{R}$-sectoriality of $-\nabla\cdot\mu\nabla$ on L^p, we first consider Gaussian bounds of the heat kernels associated with the respective semigroups. Using an argument due to Davies, the Gaussian bound may be extended to hold on a sector Σ_θ. Since $\mathcal{R}$-boundedness is inherited by domination, we obtain $\mathcal{R}$-boundedness of the semigroup operators and, thus, $\mathcal{R}$-boundedness of the resolvents employing the Laplace transformation. This is a well-established idea originating from [63, Section 4e].

Proposition 3.11. *Let $0 < \mu_\bullet < \mu^\bullet$ and let $S_{-\nabla\cdot\mu\nabla+1}$ denote the semigroup generated by $\nabla\cdot\mu\nabla - 1$ for $\mu \in \mathcal{M}_d(\mu_\bullet,\mu^\bullet)$. The operators $S_{-\nabla\cdot\mu\nabla+1}(t)$ have positive kernels K_t satisfying upper Gaussian estimates, i.e. there exist $b, c > 0$ such that*

$$0 \le K_t(x,y) \le ct^{-d/2}\mathrm{e}^{-b\frac{|x-y|^2}{t}}, \quad \text{for a.a. } x,y \in \Omega,$$

uniformly in $t > 0$ and $\mu \in \mathcal{M}_d(\mu_\bullet,\mu^\bullet)$.

Proof. This is a special case of [8, Theorem 4.4], cf. also [58, Theorem 7.5]. The assumptions on the space $W^{1,2}_D$ are verified in the proof of Theorem 3.1 in [59]. The constants in [8] are constructive and can be chosen uniformly with respect to μ due to $\mu \in \mathcal{M}_d(\mu_\bullet,\mu^\bullet)$. This yields the Gaussian bound

$$0 \le K_t(x,y) \le ct^{-d/2}\mathrm{e}^{-b\frac{|x-y|^2}{t}}\mathrm{e}^{\omega t}, \quad \text{for a.a. } x,y \in \Omega,$$

for some constants $b, c > 0$, and $\omega \in \mathbb{R}$. Since we consider the operator $-\nabla\cdot\mu\nabla+1$ instead of $-\nabla\cdot\mu\nabla$, all calculations in the proof of [8, Theorem 4.4] hold with ρ^2 instead of $1+\rho^2$ and the estimate above is valid with $\omega = 0$. ☐

Proposition 3.12 ([25, Proposition 3.3])**.** *Let S be a holomorphic semigroup on L^2 with holomorphy sector Σ_{θ_0} and uniform bound $\|S(z)\|_{L^2} \le C$ for all $z \in \Sigma_{\theta_0}$. Suppose S satisfies Gaussian bounds as in Proposition 3.11. Then the semigroup S is holomorphic on L^p for any $p \in [1,\infty]$ with holomorphy sector Σ_{θ_0}. Moreover, $S(z)$ has a kernel K_z satisfying upper Gaussian estimates. More precisely, for all $\varepsilon \in (0,1]$ and $\theta \in (0,\varepsilon\,\theta_0)$ there exists $c > 0$ such that*

$$|K_z(x,y)| \le c(\Re z)^{-d/2}\mathrm{e}^{-(1-\varepsilon)b\frac{|x-y|^2}{|z|}}, \quad \text{for a.a. } x,y \in \Omega,$$

uniformly in $z \in \Sigma_\theta$. The constant c depends exclusively on ε, C, the domain Ω, and the constant c of Proposition 3.11.

We remark that the doubling property and the uniformity in growth condition required for [25, Proposition 3.3, cf. also text after proof] are satisfied since Ω is a d-set; see Proposition A.2 in the appendix. Furthermore, the notation for kernel bounds used in [25] is equivalent to ours up to positive constants due to the uniformity condition.

Lemma 3.13. *Let $0 < \mu_\bullet < \mu^\bullet$. There exist $\theta \in (0,\pi/2)$ and $c > 0$ such that*

$$\mathcal{R}\left(\left\{zR(z,-\nabla\cdot\mu\nabla+1)\colon z \in \mathbb{C}\setminus\overline{\Sigma_\theta}\right\}\right) \le c, \tag{3.13}$$

in $\mathcal{L}(L^p)$ uniformly in $\mu \in \mathcal{M}_d(\mu_\bullet,\mu^\bullet)$.

Proof. Due to Proposition B.1, the semigroup $S_{-\nabla\cdot\mu\nabla+1}$ generated by $\nabla\cdot\mu\nabla-1$ on L^2 is uniformly bounded. Hence, according to Proposition 3.12, the Gaussian estimate of Proposition 3.11 associated with $S_{-\nabla\cdot\mu\nabla+1}$ can be extended to a sector Σ_θ. Clearly, each operator $S_{-\nabla\cdot\mu\nabla+1}(z)$ is dominated by the Gaussian bound, i.e.

$$|(S_{-\nabla\cdot\mu\nabla+1}(z)f)(x)| \le c(\Re z)^{-d/2}\int_\Omega \mathrm{e}^{-b\frac{|x-y|^2}{|z|}}|f(y)|dy,$$

for almost all $x \in \Omega$ and all $z \in \Sigma_\theta$. Due to [25, Proposition 2.4], the latter can be bounded by the Hardy-Littlewood maximal function which defines a bounded linear operator M on L^p; see [25, p. 97]. Thus,

$$\mathcal{R}\left(\{S_{-\nabla\cdot\mu\nabla+1}(z)\colon z \in \Sigma_\theta\}\right) \le c\|M\|_{\mathcal{L}(L^p)}$$

holds uniformly with respect to $t \in [0,T]$. From [64, Theorem 4.2] we conclude

$$\mathcal{R}\left(\left\{zR(z,-\nabla\cdot\mu\nabla+1)\colon z\in\mathbb{C}\setminus\overline{\Sigma_\theta}\right\}\right)\le 2\mathcal{R}\left(\{S_{-\nabla\cdot\mu\nabla+1}(z)\colon z\in\Sigma_\theta\}\right),$$

where the bound follows from the proof of [64, Theorem 2.10]. $\square$

Lemma 3.14. *Let $0<\mu_\bullet<\mu^\bullet$. There exist $\theta\in(0,\pi/2)$ and $c>0$ such that*

$$\mathcal{R}\left(\left\{R(z,-\nabla\cdot\mu\nabla+1)\colon z\in\mathbb{C}\setminus\overline{\Sigma_\theta}\right\}\right)\le c, \tag{3.14}$$

in $\mathcal{L}(L^p)$ uniformly in $\mu\in\mathcal{M}_d(\mu_\bullet,\mu^\bullet)$.

Proof. Let $S_{-\nabla\cdot\mu\nabla+1}$ denote the semigroup generated by $\nabla\cdot\mu\nabla-1$. From Proposition B.2 we conclude that the semigroup is exponentially stable in L^p, i.e. there exists $\omega>0$ such that

$$\|S_{-\nabla\cdot\mu\nabla+1}(z)\|_{\mathcal{L}(L^p)}\le \mathrm{e}^{-\omega z},\quad z>0,$$

uniformly in $\mu\in\mathcal{M}_d(\mu_\bullet,\mu^\bullet)$. According to [9, Corollary 2.4], we infer

$$\mathcal{R}\left(\{R(z,-\nabla\cdot\mu\nabla+1)\colon \Re z\le 0\}\right)\le 2/\omega. \tag{3.15}$$

Indeed, we may extend the $\mathcal{R}$-bound (3.15) to a sector $\mathbb{C}\setminus\overline{\Sigma_\theta}$ for some $\theta\in(0,\pi/2)$. Using the power series expansion for $\lambda\in\mathbb{R}$, see f.i. [35, Proposition A.2.3], it holds

$$R(\lambda\mathrm{e}^{\mathrm{i}\varphi+\mathrm{i}\pi/2},-\nabla\cdot\mu\nabla+1)=R(\mathrm{i}\lambda,-\nabla\cdot\mu\nabla+1)\sum_{m=0}^{\infty}(1-\mathrm{e}^{\mathrm{i}\varphi})^m[\mathrm{i}\lambda R(\mathrm{i}\lambda,-\nabla\cdot\mu\nabla+1)]^m.$$

Let C denote the $\mathcal{R}$-bound (3.13). Choose $\theta_0\in(0,\pi/2)$ sufficiently small such that $|1-\mathrm{e}^{\mathrm{i}\varphi}|<(2C)^{-1}$ for $|\varphi|<\theta_0$. Due to (3.15) for any $\varphi\in[-\theta_0,\theta_0]$ we have

$$\begin{aligned}&\mathcal{R}(\{R(\lambda\mathrm{e}^{\mathrm{i}\varphi+\mathrm{i}\pi/2},-\nabla\cdot\mu\nabla+1)\colon\lambda\in\mathbb{R}\})\\&\qquad\le\frac{2}{\omega}\sum_{m=0}^{\infty}(2C)^{-m}\mathcal{R}(\{\mathrm{i}\lambda R(\mathrm{i}\lambda,-\nabla\cdot\mu\nabla+1)\colon\lambda\in\mathbb{R}\})^m\le\frac{2}{\omega}\sum_{m=0}^{\infty}2^{-m}\le\frac{4}{\omega},\end{aligned}$$

where we have used Remark 3.10. Lemma 2.7 (cf. also Example 2.9) in [64] shows

$$\mathcal{R}\left(\left\{R(z,-\nabla\cdot\mu\nabla+1)\colon z\in\mathbb{C}\setminus\overline{\Sigma_\theta}\right\}\right)\le 2\mathcal{R}(\{R(\lambda\mathrm{e}^{\mathrm{i}\varphi+\mathrm{i}\pi/2},-\nabla\cdot\mu\nabla+1)\colon\lambda\in\mathbb{R}\}),$$

hence (3.14) holds with $c=8/\omega$ and $\theta=\pi/2-\theta_0$. $\square$

Clearly, from the estimates (3.13) and (3.14) and using Remark 3.10, we infer the resolvent estimate

$$\|R(z,-\nabla\cdot\mu\nabla+1)\|_{\mathcal{L}(L^p)}\le\frac{c}{1+|z|},\quad z\in\mathbb{C}\setminus\overline{\Sigma_\theta}, \tag{3.16}$$

uniformly in $\mu\in\mathcal{M}_d(\mu_\bullet,\mu^\bullet)$.

3.2.2. *Uniform $\mathcal{R}$-sectoriality in $W^{-1,p}_D(\Omega)$.* We next establish uniform $\mathcal{R}$-sectoriality on $W^{-1,p}_D$. Since $(-\nabla\cdot\mu\nabla+1)^{1/2}$ provides an isomorphism from L^p onto $W^{-1,p}_D$ and commutes with the resolvent of $-\nabla\cdot\mu\nabla+1$, the result on $W^{-1,p}_D$ follows from the result on L^p, as the square root operators are uniformly bounded.

Proposition 3.15 ([11, Theorem 5.1]; for the uniformity [26])**.** *Let $0<\mu_\bullet<\mu^\bullet$ and $\mu\in\mathcal{M}_d(\mu_\bullet,\mu^\bullet)$ and $p\in[2,\infty)$. The operator $(-\nabla\cdot\mu\nabla+1)^{1/2}$ is an isomorphism from L^p onto $W^{-1,p}_D$. Moreover, the operator norms of $(-\nabla\cdot\mu\nabla+1)^{1/2}$ and $(-\nabla\cdot\mu\nabla+1)^{-1/2}$ are uniformly bounded with respect to $\mu\in\mathcal{M}_d(\mu_\bullet,\mu^\bullet)$.*

We emphasize that the regularity requirements of [11] are considerably weaker than Gröger regular; see Propositions A.2 and A.3 in the appendix and for Kato's square root property [27, Theorem 4.1, Remark 2.4 (3)]. The following lemma is a direct consequence of Proposition 3.15 and the corresponding results (3.13) and (3.16) on L^p.

Lemma 3.16. *Let $0 < \mu_\bullet < \mu^\bullet$. There exist $\theta \in (0, \pi/2)$ and $c > 0$ such that*

$$\mathcal{R}\left(\left\{ zR(z, -\nabla \cdot \mu\nabla + 1) \colon z \in \mathbb{C} \setminus \overline{\Sigma_\theta} \right\}\right) \le c, \tag{3.17}$$

in $\mathcal{L}(W_D^{-1,p})$ and

$$\|R(z, -\nabla \cdot \mu\nabla + 1)\|_{\mathcal{L}(W_D^{-1,p})} \le \frac{c}{1 + |z|}, \quad z \in \mathbb{C} \setminus \overline{\Sigma_\theta}, \tag{3.18}$$

uniformly in $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$.

3.2.3. *Acquistapace-Terreni condition.* As the next step towards our regularity result Theorem 3.20 we verify the so-called *Acquistapace-Terreni* (AT) condition. A family of operators $\{ A(t) \colon t \in [0, T] \}$ on a Banach space X satisfies the (AT) condition if there are constants $0 \le \beta < \alpha < 1$, $\theta \in (0, \pi/2)$ and $c > 0$ such that

$$\|A(t)R(z, A(t))[A(t)^{-1} - A(s)^{-1}]\|_{\mathcal{L}(X)} \le c|t - s|^{\alpha}|z|^{\beta - 1} \tag{3.19}$$

for all $t, s \in [0, T]$ and $z \in \mathbb{C} \setminus \overline{\Sigma_\theta}$. We verify this condition for $X = H_D^{-\zeta,p}$. To this end, we first consider the differential operators on L^p, and thereafter on $W_D^{-1,p}$. Recall that given any control $q \in L^s(I; H_D^{-\zeta,p}) \hookrightarrow L^s(I; W_D^{-1,p})$, there is a unique solution $u \in W^{1,s}(I; W_D^{1,p}) \cap L^s(I; W_D^{-1,p})$ to the state equation (1.1b). We set

$$A(t) \coloneqq \mathcal{A}(u(t)) = -\nabla \cdot \xi(u(t))\mu\nabla.$$

Proposition 3.17. *For $\hat{\beta} \in (0, 1/2)$, there are $\theta \in (0, \pi/2)$ and $c > 0$ such that*

$$\|(A(t) + 1)R(z, A(t) + 1)[(A(t) + 1)^{-1} - (A(s) + 1)^{-1}]\|_{\mathcal{L}(L^p)} \le c|t - s|^{\alpha}|z|^{-\hat{\beta}},$$

for all $z \in \mathbb{C} \setminus \overline{\Sigma_\theta}$ and $s, t \in [0, T]$.

Proof. Unlike on $W_D^{-1,p}$, the domains of the elliptic operators on L^p may depend on $t \in [0, T]$. However, some intermediate spaces between L^p and its domains are constant. According to [22, Lemma 4.7] and Assumption 3, for $\hat{\beta} \in (0, 1/2)$ it holds

$$(L^p, \mathcal{D}_{L^p}(A(t)))_{\hat{\beta},\infty} = (W_D^{-1,p}, W_D^{1,p})_{\hat{\beta}+1/2,\infty}, \tag{3.20}$$

i.e. the real interpolation space $(L^q, \mathcal{D}_{L^p}(A(t) + 1))_{\hat{\beta},\infty}$ is constant with respect to time. Moreover, the resolvent identity yields

$$(A(t) + 1)^{-1} - (A(s) + 1)^{-1} = (A(t) + 1)^{-1}[A(s) - A(t)](A(s) + 1)^{-1}.$$

Thus, Hölder continuity of $t \mapsto A(t) + 1 \in \mathcal{L}(W_D^{1,p}, W_D^{-1,p})$ and the continuous injections $L^p \hookrightarrow W_D^{-1,p}$ as well as $W_D^{1,p} \hookrightarrow (W_D^{-1,p}, W_D^{1,p})_{\hat{\beta}+1/2,\infty}$ imply

$$\begin{aligned} &\|(A(t) + 1)^{-1} - (A(s) + 1)^{-1}\|_{\mathcal{L}(L^p, (W_D^{-1,p}, W_D^{1,p})_{\hat{\beta}+1/2,\infty})} \\ &\le c\|(A(t) + 1)^{-1}\|_{\mathcal{L}(W_D^{-1,p}, W_D^{1,p})}|t - s|^{\alpha}\|(A(s) + 1)^{-1}\|_{\mathcal{L}(W_D^{-1,p}, W_D^{1,p})} \\ &\le c|t - s|^{\alpha}. \end{aligned} \tag{3.21}$$

In the last step we have used that smoothness of the inversion mapping and continuity of $t \mapsto A(t) + 1 \in \mathcal{L}(W_D^{1,p}, W_D^{-1,p})$ yield a constant $c > 0$ independent of s such that it holds

$$\|(A(s) + 1)^{-1}\|_{\mathcal{L}(W_D^{-1,p}, W_D^{1,p})} \le c \quad \forall s \in [0, T]. \tag{3.22}$$

Since the operators are uniformly sectorial, see Proposition 3.8 and Lemma 3.16, we use Proposition C.2 characterizing the real interpolation space $(L^p, \mathcal{D}_{L^p}(A(t)))_{\hat{\beta},\infty}$ with uniform equivalence of norms. Thus, (3.20) and (3.21) imply

$$\begin{aligned}
&\|(A(t) + 1)R(z, A(t) + 1)[(A(t) + 1)^{-1} - (A(s) + 1)^{-1}]\varphi\|_{L^p} \\
&\le c|z|^{-\hat{\beta}}\|[(A(t)+1)^{-1} - (A(s)+1)^{-1}]\varphi\|_{(W_D^{-1,p}, W_D^{1,p})_{\hat{\beta}+1/2,\infty}} \le c|t-s|^{\alpha}|z|^{-\hat{\beta}}\|\varphi\|_{L^p},
\end{aligned}$$

for all $\varphi \in L^p$, $z \in \mathbb{C} \setminus \overline{\Sigma_\theta}$ and $s, t \in [0, T]$; cf. also Hypothesis 7.3 in [2]. $\square$

Proposition 3.18. *There exist $\theta \in (0, \pi/2)$ and $c > 0$ such that*

$$\|(A(t) + 1)R(z, A(t) + 1)[(A(t) + 1)^{-1} - (A(s) + 1)^{-1}]\|_{\mathcal{L}(W_D^{-1,p})} \le c|t - s|^{\alpha}|z|^{-1},$$

for all $z \in \mathbb{C} \setminus \overline{\Sigma_\theta}$ and $s, t \in [0, T]$.

Proof. Since the operator $A(t)$ is an isomorphism from $W_D^{1,p}$ onto $W_D^{-1,p}$, in particular $A(t)$ has constant domain with respect to t. The resolvent identity yields

$$\begin{aligned}
&(A(t) + 1)R(z, A(t) + 1)[(A(t) + 1)^{-1} - (A(s) + 1)^{-1}] \\
&\qquad = R(z, A(t) + 1)[A(t) - A(s)](A(s) + 1)^{-1}.
\end{aligned} \tag{3.23}$$

Using Hölder continuity of u and Lipschitz continuity of ξ, we have

$$\|A(t) - A(s)\|_{\mathcal{L}(W_D^{1,p}; W_D^{-1,p})} \le \|\xi(u(t))\mu - \xi(u(s))\mu\|_{L^\infty(\Omega)} \le c|t - s|^{\alpha}. \tag{3.24}$$

Taking the $\mathcal{L}(W_D^{-1,p})$ norm in (3.23) and employing (3.22), (3.24) and the resolvent estimate (3.18), we obtain the assertion. $\square$

3.2.4. *Maximal parabolic regularity.* We now prove the main result of this section by employing Lemma D.1. Since this requires UMD spaces we need

Proposition 3.19. *The space $H_D^{-\zeta,p}$ is an UMD space* [7, Section III.4.4].

Proof. L^p is an UMD space for all $p \in (1, \infty)$; see, e.g., [45, Section 3.14]. Whence the result follows from [7, Theorem III.4.5.2]. $\square$

Theorem 3.20. *Let $s \in (1, \infty)$. The nonautonomous operator $\mathcal{A}(u)$ has maximal parabolic regularity on $L^s(I; H_D^{-\zeta,p})$. Furthermore, consider $\Xi \subset C(\mathbb{R})$ such that*

(i) For all $\xi \in \Xi$ it holds $\xi_\bullet \le \xi \le \xi^\bullet$, and
(ii) Ξ is equi-Lipschitz continuous on bounded sets, i.e. for all $K > 0$ there is a constant $C_K > 0$ such that

$$|\xi(x) - \xi(y)| \le C_K|x - y| \quad \forall\, |x|, |y| < K, \forall\, \xi \in \Xi.$$

The norm of the solution operator

$$(\partial_t - \nabla \cdot \xi(u)\mu\nabla, \gamma_0)^{-1}$$

from $L^s(I; H_D^{-\zeta,p}) \times (H_D^{-\zeta,p}, D)_{1-1/s,s}$ into $W^{1,s}(I; H_D^{-\zeta,p}) \cap L^s(I; D)$ is uniformly bounded with respect to $\xi \in \Xi$ and $u \in S(Q_{ad})$, where $S(Q_{ad})$ denotes the set of solutions to the state equation (1.1b) with control $q \in Q_{ad}$, and $D := \mathcal{D}_{H_D^{-\zeta,p}}(-\nabla \cdot \mu\nabla)$.

Proof. We will verify the supposition of Lemma D.1.

Step 1. We use complex interpolation to combine the results on $W_D^{-1,p}$ and on L^p that we obtained in the preceding subsections. From Lemma 3.13 and Lemma 3.16, we immediately infer uniform resolvent estimates and uniform $\mathcal{R}$-sectoriality with respect to u. Concerning the (AT) condition, we conclude from Propositions 3.17 and 3.18 that for all $\hat{\beta} \in (0, 1/2)$ there is $c > 0$ such that

$$\|(A(t)+1)R(z, A(t)+1)[(A(t)+1)^{-1} - (A(s)+1)^{-1}]\|_{\mathcal{L}(H_D^{-\zeta,p})} \leq c|t-s|^{\alpha}|z|^{-\zeta-(1-\zeta)\hat{\beta}}$$

and set $\beta = 1-\zeta-(1-\zeta)\hat{\beta} = (1-\zeta)(1-\hat{\beta})$. We have to find $\hat{\beta} \in (0, 1/2)$ such that $0 \leq \beta < \alpha$. Clearly, $\beta > 0$. Moreover, $\beta < \alpha$ if and only if $(1-\zeta-\alpha)(1-\zeta)^{-1} < \hat{\beta}$. Moreover, $(1-\zeta-\alpha)(1-\zeta)^{-1} < 1/2$ if and only if

$$\zeta > 1 - 2\alpha = \frac{d}{p} + \frac{2}{s} + \kappa, \tag{3.25}$$

where the latter equality is due to $\alpha = 1/2 - d/2p - 1/s - \kappa/2$. Condition (3.25) is satisfied for some $\kappa > 0$ sufficiently small, due to $s > 2(\zeta - d/p)^{-1}$; see Assumption 4. Thus, there is $\hat{\beta} \in (0, 1/2)$ such that $\beta < \alpha$ and condition (3.19) is satisfied on $H_D^{-\zeta,p}$. Further, due to (3.13) and (3.17) the operators $A(t)+1$ are $\mathcal{R}$-sectorial on $H_D^{-\zeta,p}$.

Step 2. It remains to argue uniformity of the (AT) condition. Suppose for the moment that $u_0 = 0$. According to [48, Theorem 2.13 ii)], there is $\alpha > 0$ such that the mappings

$$(\partial_t + \mathcal{A}(u), \gamma_0)^{-1} : L^s(I; W_D^{-1,p}) \times \{0\} \to C^{\alpha}(I \times \Omega)$$

are equicontinuous for all $u \in C(\overline{I \times \Omega})$, due to the lower and upper bound on ξ and μ; see Assumption 2. Whence, $(\partial_t + \mathcal{A}(u), \gamma_0)^{-1}(BQ_{ad}, 0)$ is contained in a compact subset of $C(\overline{I \times \Omega})$, because of boundedness of Q_{ad}.

If $u_0 \neq 0$, we set $v_1(t) = S_{-\nabla\cdot\mu\nabla}(t)u_0$ and find $v_1 \in W^{1,s}(I; W_D^{1,p}) \cap L^s(I; W_D^{-1,p})$; see, e.g., [7, Proposition III.4.10.3]. Thus, it holds

$$(\partial_t + \mathcal{A}(u), \gamma_0)^{-1}(Bq, u_0) = v_1 + v_2$$

for any $q \in Q_{ad}$, where v_2 solves

$$(\partial_t + \mathcal{A}(u))\, v_2 = Bq - (\partial_t + \mathcal{A}(u))\, v_1, \quad v_2(0) = 0.$$

Now we are in the situation with homogeneous initial conditions as before, and since

$$\|(\partial_t + \mathcal{A}(u))\, v_1\|_{L^s(I; W_D^{-1,p})} \leq c\|u_0\|_{(W_D^{-1,p}, W_D^{1,p})_{1-1/s,s}}$$

with uniform constant, we find that v_2 is contained in a compact subset of $C(\overline{I \times \Omega})$. Employing embedding (3.4) we in summary infer that $(\partial_t + \mathcal{A}(u), \gamma_0)^{-1}(BQ_{ad}, u_0)$ is a compact subset of $C(\overline{I \times \Omega})$. Furthermore, $u \mapsto (\partial_t + \mathcal{A}(u), \gamma_0)^{-1}$ is continuous from $C(\overline{I \times \Omega})$ into

$$\mathcal{L}(L^s(I; W_D^{-1,p}) \times (W_D^{-1,p}, W_D^{1,p})_{1-1/s,s}; W^{1,s}(I; W_D^{1,p}) \cap L^s(I; W_D^{-1,p})).$$

Hence, even the operators $(\partial_t + \mathcal{A}(u), \gamma_0)^{-1}$ mapping from

$$L^s(I; W_D^{-1,p}) \times (W_D^{-1,p}, W_D^{1,p})_{1-1/s,s} \text{ to } W^{1,s}(I; W_D^{1,p}) \cap L^s(I; W_D^{-1,p})$$

are uniformly bounded with respect to u. For this reason and using again boundedness of Q_{ad} with the embedding (3.4), we obtain that u is uniformly bounded in $C^{\alpha}(I; C^{\kappa}(\Omega))$, where α is as in (3.4). With the resolvent estimates (3.16) and (3.18), we see that the (AT) conditions in Proposition 3.17 and Proposition 3.18 are uniform with respect to u.

Step 3. Finally, Lemma D.1 guarantees maximal parabolic regularity of $A(\cdot)$ on $L^s(I; H_D^{-\zeta,p})$, since $H_D^{-\zeta,p}$ is an UMD space; see Proposition 3.19. Indeed, according to [21, Corollary 2.3] we have

$$u(t) \in (H_D^{-\zeta,p}, \mathcal{D}_{H_D^{-\zeta,p}}(A(t)))_{1-1/s,s} \quad \forall t \in [0,T].$$

Hence, for $\tau \in (\frac{1+\zeta}{2}, 1-\frac{1}{s})$, we find

$$(H_D^{-\zeta,p}, \mathcal{D}_{H_D^{-\zeta,p}}(A(t)))_{1-1/s,s} \hookrightarrow (H_D^{-\zeta,p}, \mathcal{D}_{H_D^{-\zeta,p}}(A(t)))_{\tau,1} \subset W_D^{1,p};$$

see [61, Theorem 1.3.3 (e)] and (3.9). Therefore, (3.10) implies that

$$\mathcal{D}_{H_D^{-\zeta,p}}(A(t)) = \mathcal{D}_{H_D^{-\zeta,p}}(-\nabla \cdot \mu \nabla) = D,$$

i.e. the operators $A(t)$ have constant domain with respect to time. □

4. Optimal control problem

After the detailed discussion of the state equation we now return to the optimal control problem. By means of Theorem 3.20 it is justified to introduce the control-to-state mapping

$$S\colon Q \to W^{1,s}(I; H_D^{-\zeta,p}) \cap L^s(I; \mathcal{D}_{H_D^{-\zeta,p}}(-\nabla \cdot \mu \nabla)), \quad S(q) = u,$$

where u denotes the solution of (1.1b) for any control $q \in Q = L^\infty(\Lambda, \varrho)$. Recall that $p \in (d,4)$ from Assumption 3 is close to the spatial dimension d and the numbers ζ and s are chosen according to Assumption 4 with ζ very close to 1 and $s > \frac{2}{\zeta - d/p}, \frac{2}{1-\zeta}$ approaching $+\infty$.

The control-to-state mapping S leads to the reduced objective function

$$j\colon Q \to \mathbb{R}_0^+,\ q \mapsto J(q, S(q)).$$

Here and in the following we omit trivial embedding operators to improve readability. Then, the optimal control problem (1.1) is equivalent to

$$(P) \qquad \text{Minimize } j(q) \text{ subject to } q \in Q_{ad}.$$

Since the set of admissible controls Q_{ad} is not empty due to Assumption 4, we obtain by standard arguments, see, e.g., [62], the following existence result for optimal controls. In particular, we use compactness of the embedding $W^{1,s}(I; W_D^{-1,p}) \cap L^s(I; W_D^{1,p}) \hookrightarrow_c C^\alpha(I; C^\kappa)$ and continuity of the mapping $u \mapsto \mathcal{A}(u)$ from $C(\overline{I \times \Omega})$ to $\mathcal{L}(W_D^{1,p}, W_D^{-1,p})$; see also [48, Section 6].

Lemma 4.1. *The optimal control problem (P) admits at least one globally optimal control $\bar{q} \in Q_{ad}$ with associated optimal state $\bar{u} = S(\bar{q})$.*

We point out that the reduced objective function is not necessarily convex due to the nonlinear state equation and introduce the notation of local solutions.

Definition 4.2. A control $\bar{q} \in Q_{ad}$ is called a *local solution* of (P) in the sense of $L^2(\Lambda, \varrho)$ if there exists a constant $\varepsilon > 0$ such that the inequality

$$j(q) \geq j(\bar{q})$$

is satisfied for all $q \in Q_{ad}$ with $\|\bar{q} - q\|_{L^2(\Lambda,\varrho)} \leq \varepsilon$.

4.1. **Differentiability of the control-to-state mapping.** We first prove differentiability of the control-to-state mapping S and thereafter derive first and second order optimality conditions. To ease readability, we introduce the following notation

$$\mathcal{A}'(u)v \coloneqq -\nabla \cdot \xi'(u)v\mu\nabla u,$$
$$\mathcal{A}''(u)[v_1, v_2] \coloneqq -\nabla \cdot \left[\xi'(u)(v_1\mu\nabla v_2 + v_2\mu\nabla v_1) + \xi''(u)v_1v_2\mu\nabla u\right],$$

for $v, v_1, v_2 \in W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$ and $r \in (1, \infty)$.

Proposition 4.3. *Let $u \in W^{1,s}(I; W^{-1,p}_D) \cap L^s(I; W^{1,p}_D)$. For any $r \in (1, \infty)$ the mapping $v \mapsto \mathcal{A}'(u)v$ is linear and completely continuous from $W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$ into $L^r(I; W^{-1,p}_D)$. Moreover, $\mathcal{A}''(u)[\cdot, \cdot]$ is linear and completely continuous in each component from $W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$ into $L^r(I; W^{-1,p}_D)$ for $r > 2p/(p - d)$.*

Proof. We consider

$$\langle \mathcal{A}'(u)v, \varphi\rangle = \int_I \int_\Omega \xi'(u)v\mu\nabla u \cdot \nabla\varphi, \quad \varphi \in L^{r'}(I; W^{1,p'}_D). \tag{4.1}$$

If $r > 2p/(p-d)$, then it holds the embedding $v \in W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D) \hookrightarrow L^\infty(I; L^\infty)$, see (3.4), and (4.1) is bounded by the norm of v. Otherwise, if $r \le 2p/(p - d)$, then we have to require $v \in L^{r_1}(I; L^\infty)$ with $r_1 = \frac{rs}{s-r}$ and $1 < r < s$. It holds $v \in W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D) \hookrightarrow L^{r_1}(I; L^\infty)$ provided that

$$r_1 = \frac{rs}{s - r} < \frac{2pr}{dr - p(r - 2)},$$

see (3.3), which is equivalent to $s > \frac{2p}{p-d}$. Due to Assumption 4 we have

$$s > \frac{2}{1 - \zeta} > \frac{2}{1 - d/p} = \frac{2p}{p - d} \ge r.$$

Hence, the embedding (3.3) is valid, proving continuity in case $r \le 2p/(p-d)$. Since the embeddings (3.3) and (3.4) are compact, we even have complete continuity.

By similar arguments we observe that $\mathcal{A}''(u)[\cdot, \cdot]$ is continuous in each component from $L^\infty(I; L^\infty) \cap L^r(I; W^{1,p}_D)$ into $L^r(I; W^{-1,p}_D)$. Thus, we have to require $r > \frac{2p}{p-d}$ to employ embedding (3.4) and obtain continuity from $W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$ to $L^r(I; W^{-1,p}_D)$. □

Proposition 4.4. *Let $r \in (1, \infty)$ and $u = S(q)$ be the state associated with a control $q \in Q_{ad}$. For each right-hand side $f \in L^r(I; W^{-1,p}_D)$ there exists a unique solution $v \in W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$ to the equation*

$$\partial_t v + \mathcal{A}(u)v + \mathcal{A}'(u)v = f, \quad v(0) = 0. \tag{4.2}$$

Moreover, if $f_n \rightharpoonup f$ in $L^2(I; W^{-1,p}_D)$, then $v_n \to v$ in $L^{r_1}(I; C^\kappa)$ for $r_1 \in (1, 2p/d)$, where v_n denotes the corresponding solution to (4.2) with right-hand side f_n.

Proof. According to Lemma 3.4 the linear mapping

$$\partial_t + \mathcal{A}(u) \colon W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D) \cap \{\varphi(0) = 0\} \to L^r(I; W^{-1,p}_D)$$

provides a topological isomorphism for all $r \in (1, \infty)$ and, in particular, $\partial_t + A(u)$ defines a Fredholm operator of index 0. We will prove that $\mathcal{A}'(u)$ is relatively compact with respect to $\partial_t + \mathcal{A}(u)$ to apply a perturbation result. To this end, let $(v_n)_n$ be a sequence that is bounded in $W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$. Due to complete continuity of $\mathcal{A}'(u)$, the sequence $\mathcal{A}'(u)v_n$ converges in $L^r(I; W_D^{-1,p})$. Thus, $\mathcal{A}'(u)$ is relatively compact and, using the perturbation result [41, Theorem IV. 5.3 5.26], we infer that $\partial_t + \mathcal{A}(u) + \mathcal{A}'(u)$ is a Fredholm operator of index 0. Hence, $\partial_t + \mathcal{A}(u) + \mathcal{A}'(u)$ is a topological isomorphism, which proves the first statement. The second assertion is again a consequence of the compact injection $W^{1,2}(I; W_D^{-1,p}) \cap L^2(I; W_D^{1,p}) \hookrightarrow L^{r_1}(I; C^\kappa)$; see embedding (3.3). □

Lemma 4.5. *The control-to-state mapping S is twice continuously Fréchet-differentiable from $L^\infty(\Lambda, \varrho)$ to $W^{1,s}(I; W_D^{-1,p}) \cap L^s(I; W_D^{1,p})$, for $s > 2p/(p-d)$. Moreover, $v = S'(q)\delta q$ is the unique solution of the linearized state equation*

$$\partial_t v + \mathcal{A}(u)v + \mathcal{A}'(u)v = B\delta q, \quad v(0) = 0,$$

with $u = S(q)$, and $w = S''(q)(\delta q_1, \delta q_2)$ is the unique solution of

$$\partial_t w + \mathcal{A}(u)w + \mathcal{A}'(u)w = \mathcal{A}''(u)[v_1, v_2], \quad w(0) = 0,$$

where $v_i = S'(q)\delta q_i$, $i \in \{1, 2\}$, and $u = S(q)$. Furthermore, $S'(q)$ can be uniquely extended to a continuous mapping from $L^2(\Lambda, \varrho)$ to $W^{1,2}(I; W_D^{-1,p}) \cap L^2(I; W_D^{1,p})$.

Proof. This follows from the implicit function theorem and Proposition 4.4. □

4.2. **First order optimality conditions.**

Lemma 4.6. *Let $r \in (1, \infty)$ and $u = S(q)$ be the state corresponding to a control $q \in L^\infty(\Lambda, \varrho)$. There is a unique* adjoint state *$z = z(q) \in L^r(I; W_D^{1,p'})$ such that*

$$j'(q)(\delta q) = (\lambda q + B^* z, \delta q)_{L^2(\Lambda,\varrho)}, \quad \delta q \in L^2(\Lambda, \varrho). \tag{4.3}$$

Furthermore, z has the improved regularity $z \in W^{1,r}(I; W_D^{-1,p'}) \cap L^r(I; W_D^{1,p'})$ and

$$-\partial_t z + \mathcal{A}(u)^* z + \mathcal{A}'(u)^* z = u - \hat{u}, \quad z(T) = 0.$$

As usual, B^* is the adjoint operator of B, i.e. $B^* \colon L^2(I; W_D^{1,p'}) \to L^2(\Lambda, \varrho)$.

Proof. According to Lemma 4.5, the control-to-state mapping S is continuously Fréchet-differentiable and the chain rule yields

$$\begin{aligned} j'(q)(\delta q) &= (S(q) - \hat{u}, S'(q)\delta q)_{L^2(I\times\Omega)} + \lambda(q, \delta q)_{L^2(\Lambda,\varrho)} \\ &= (S'(q)^*(S(q) - \hat{u}), \delta q)_{L^2(I\times\Omega)} + \lambda(q, \delta q)_{L^2(\Lambda,\varrho)}. \end{aligned} \tag{4.4}$$

To find an expression for j', we have to make sense of $S'(q)^*$. First, we observe $S'(q)\delta q = (\partial_t + \mathcal{A}(u) + \mathcal{A}'(u), \gamma_0)^{-1} (B\delta q, 0)$, where γ_0 denotes the trace mapping. According to Proposition 4.4 the linear mapping

$$(\partial_t + \mathcal{A}(u) + \mathcal{A}'(u), \gamma_0) : W^{1,r'}(I; W_D^{-1,p}) \cap L^{r'}(I; W_D^{1,p}) \to L^{r'}(I; W_D^{-1,p}) \times \{0\}$$

provides a topological isomorphism for all $r' \in (1, \infty)$, i.e. $\mathcal{A}(u) + \mathcal{A}'(u)$ satisfies maximal parabolic regularity on $L^r(I; W_D^{-1,p})$; see [5, Proposition 3.1]. Clearly,

$$(\partial_t + \mathcal{A}(u) + \mathcal{A}'(u), \gamma_0)^* : L^r(I; W_D^{1,p'}) \times \{0\} \to \left(W^{1,r'}(I; W_D^{-1,p}) \cap L^{r'}(I; W_D^{1,p})\right)^* \tag{4.5}$$

is a topological isomorphism as well. Since $W^{1,r'}(I; W_D^{-1,p}) \cap L^{r'}(I; W_D^{1,p})$ is dense both in $W^{1,r'}(I; W_D^{-1,p})$ and $L^{r'}(I; W_D^{1,p})$, and these spaces continuously embed into $L^{r'}(I; W_D^{-1,p})$, we have [31, Satz 1.5.13]

$$\left(W^{1,r'}(I; W_D^{-1,p}) \cap L^{r'}(I; W_D^{1,p})\right)^* = (W^{1,r'}(I; W_D^{-1,p}))^* + L^r(I; W_D^{-1,p'}).$$

Because of $W^{1,p}_D \hookrightarrow L^2 \hookrightarrow W^{-1,p'}_D$, we may identify $u - \hat{u}$ with an element in $L^r(I; W^{-1,p'}_D)$. Furthermore, we identify the adjoint of the trace mapping

$$\gamma_T \colon C([0,T]; (W^{-1,p}_D, W^{1,p}_D)_{1-1/r',r'}) \to (W^{-1,p}_D, W^{1,p}_D)_{1-1/r',r'}$$

by the Dirac measure δ_T in T, i.e.

$$\delta_T \otimes z_T \in \mathcal{M}\left(I; (W^{-1,p'}_D, W^{1,p'}_D)_{1-1/r,r}\right)$$

for any $z_T \in (W^{-1,p'}_D, W^{1,p'}_D)_{1-1/r,r}$, and in particular for $z_T = 0$. Using the embedding $W^{1,r'}(I; W^{-1,p}_D) \cap L^{r'}(I; W^{1,p}_D) \hookrightarrow C([0,T]; (W^{-1,p}_D, W^{1,p}_D)_{1-1/r,r})$, we observe

$$u - \hat{u} + \delta_T \otimes 0 \in \left(W^{1,r}(I; W^{-1,p'}_D) \cap L^r(I; W^{1,p'}_D)\right)^*.$$

The initial condition is implicitly contained in $\left(W^{1,r}(I; W^{-1,p'}_D) \cap L^r(I; W^{1,p'}_D)\right)^*$. The isomorphism (4.5) yields the existence of $z \in L^r(I; W^{1,p'}_D)$ satisfying

$$\begin{aligned}\int_0^T \langle (\partial_t + \mathcal{A}(u) + \mathcal{A}'(u))^* z, \varphi\rangle &= \int_0^T \langle u - \hat{u} + \delta_T \otimes z_T, \varphi\rangle + \langle 0, \varphi(0)\rangle_{(W^{-1,p}_D, W^{1,p}_D)} \\ &= \int_0^T \langle u - \hat{u}, \varphi\rangle - \langle z_T, \varphi(T)\rangle_{(W^{-1,p}_D, W^{1,p}_D)} = \int_0^T \langle u - \hat{u}, \varphi\rangle\end{aligned}$$

for all $\varphi \in \mathcal{D}([0,T); W^{-1,p}_D)$. As $\mathcal{D}([0,T); W^{-1,p}_D) \hookrightarrow_d W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$ we verified that $z \in L^r(I; W^{1,p'}_D)$ is the weak $L^r(W^{1,p'}_D)$ solution; cf. [6, Section 6]. In summary, since $z \in L^r(I; W^{1,p'}_D)$, the expression $B^* z$ is well-defined and identity (4.4) with

$$B^* z = B^* \left(\partial_t + \mathcal{A}(u) + \mathcal{A}'(u), \gamma_0\right)^{-*} (u - \hat{u} + \delta_T \otimes 0) = S'(q)^*(S(q) - \hat{u}),$$

proves (4.3). It remains to show the improved regularity of the adjoint state. Since $\nabla u \in C([0,T]; L^p)$, the mapping $t \mapsto \mathcal{A}(u(T-t))^* + \mathcal{A}'(u(T-t))^*$ is continuous from $[0,T]$ into $\mathcal{L}(W^{1,p'}_D, W^{-1,p'}_D)$. Furthermore, similar as in Proposition 4.4, we find that for each fixed $t \in [0,T]$ the operator $\mathcal{A}(u(t)) + \mathcal{A}'(u(t))$ has maximal parabolic regularity on $W^{-1,p}_D$. Whence, $\mathcal{A}(u(t))^* + \mathcal{A}'(u(t))^*$ satisfies maximal parabolic regularity on $W^{-1,p'}_D$ for each $t \in [0,T]$. Employing [5, Theorem 7.1] we find that the nonautonomous operator $\mathcal{A}(u)^\bullet + \mathcal{A}'(u)^\bullet$ exhibits maximal parabolic regularity on $L^{s'}(I; W^{-1,p'}_D)$. Here we use the notation $\mathcal{A}(u)^\bullet = (t \mapsto \mathcal{A}(u(T-t))^*)$. [6, Proposition 6.1] yields the improved regularity of z.

For the terminal value, let $\varphi \in W^{1,r}(I; W^{-1,p}_D) \cap L^r(I; W^{1,p}_D)$ with $\varphi(0) = 0$. Using integration by parts [6, Proposition 5.1], we obtain

$$\begin{aligned}0 &= \int_0^T \langle z, \partial_t\varphi\rangle - \int_0^T \langle \partial_t^* z, \varphi\rangle = \int_0^T \langle z, \partial_t\varphi\rangle - \int_0^T \langle -\partial_t z, \varphi\rangle \\ &= \langle z(T), \varphi(T)\rangle_{(W^{-1,p}_D, W^{1,p}_D)} - \langle z(0), \varphi(0)\rangle_{(W^{-1,p}_D, W^{1,p}_D)} = \langle z(T), \varphi(T)\rangle_{(W^{-1,p}_D, W^{1,p}_D)}.\end{aligned}$$

Since the trace mapping is surjective, we conclude $z(T) = 0$. $\square$

Lemma 4.7. *Let $\bar{q} \in Q_{ad}$ be a local solution of (P). Then it holds*

$$j'(\bar{q})(p - \bar{q}) \geq 0, \quad \forall p \in Q_{ad}. \tag{4.6}$$

We refer to, e.g., Lemma 2.21 in [62], for a proof of the variational inequality. Employing the adjoint state z associated with $\bar{q}$ of Lemma 4.6 the first order necessary condition (4.6) can be expressed as

$$(\lambda \bar{q} + B^* z, p - \bar{q})_{L^2(\Lambda,\varrho)} \geq 0 \quad \forall p \in Q_{ad}.$$

Using the pointwise projection $P_{Q_{ad}}$ on the admissible set Q_{ad}, defined by

$$P_{Q_{ad}} \colon L^2(\Lambda,\varrho) \to Q_{ad}, \quad P_{Q_{ad}}(r)(t,x) = \max\{q_a, \min\{q_b, r(t,x)\}\},$$

then as in, e.g., Theorem 3.20 of [62], the optimality condition simplifies further to

$$\bar{q} = P_{Q_{ad}}\left(-\frac{1}{\lambda}B^* z(\bar{q})\right).$$

For the discussion of second order optimality conditions, we will need the following continuity result concerning the linearized and adjoint state.

Proposition 4.8. *Let $q_n \to q$ in $L^s(\Lambda,\varrho)$. Then the corresponding sequence of solution operators to the linearized equation converges in the operator norm, i.e. $S'(q_n) \to S'(q)$ in $\mathcal{L}(L^2(I;W_D^{-1,p}); W^{1,2}(I;W_D^{-1,p}) \cap L^2(I;W_D^{1,p}))$. Furthermore, the associated adjoint states satisfy $z_n \to z$ in $L^{r'}(I;W_D^{1,p'})$ for any $r' \in (1,\infty)$.*

Proof. Set $S(q_n) = u_n$. Clearly, continuity of the control-to-state mapping implies $u_n \to u$ in $W^{1,s}(I;W_D^{-1,p}) \cap L^s(I;W_D^{1,p})$. According to Proposition 4.4 the mapping

$$\partial_t + \mathcal{A}(u_n) + \mathcal{A}'(u_n) \colon W^{1,r}(I;W_D^{-1,p}) \cap L^r(I;W_D^{1,p}) \cap \{\varphi(0) = 0\} \to L^r(I;W_D^{-1,p})$$

is a topological isomorphism for each u_n. Whence, due to smoothness of the inversion mapping, for the first assertion it suffices to show

$$\mathcal{A}(u_n) + \mathcal{A}'(u_n) \to \mathcal{A}(u) + \mathcal{A}'(u) \tag{4.7}$$

in $\mathcal{L}(W^{1,r}(I;W_D^{-1,p}) \cap L^r(I;W_D^{1,p}), L^r(I;W_D^{-1,p}))$ with $r = 2$. Let $r \leq 2p/(p-d)$ and take $v \in W^{1,r}(I;W_D^{-1,p}) \cap L^r(I;W_D^{1,p})$. Then,

$$\begin{aligned}\|[\mathcal{A}(u_n) - \mathcal{A}(u)]v\|_{L^r(I;W_D^{-1,p})} &\leq c\|\xi(u_n) - \xi(u)\|_{L^\infty(I\times\Omega)}\|v\|_{L^r(I;W_D^{1,p})} \\ &\leq c\|u_n - u\|_{W^{1,s}(I;W_D^{-1,p})\cap L^s(I;W_D^{1,p})}\|v\|_{L^r(I;W_D^{1,p})},\end{aligned}$$

where we have used Lipschitz continuity of ξ from Assumption 2. Similarly, for the second term in (4.7) we calculate

$$\|[\mathcal{A}'(u_n) - \mathcal{A}'(u)]v\|_{L^r(I;W_D^{-1,p})} \leq c\|u_n - u\|_{W^{1,s}(I;W_D^{-1,p})\cap L^s(I;W_D^{1,p})}\|v\|_{L^{\frac{rs}{s-r}}(I;L^\infty)}.$$

Hence, using embedding (3.3) for v, (4.7) is a consequence of $u_n \to u$.

The second assertion follows from the first one as $z_n = S'(q_n)^*(u_n - \hat{u})$ for $r' \geq 2p/(p+d)$, hence for any $r' \in (1,\infty)$. □

4.3. **Second order optimality conditions.** We now discuss second order necessary and thereafter sufficient optimality conditions using a cone of critical directions. The analysis substantially relies on the following expression for the second derivative of the reduced objective functional employing again the adjoint state from the preceding subsection.

Proposition 4.9. *Let $q \in Q$. For all $\eta_i \in L^2(\Lambda,\varrho)$ we have*

$$\begin{aligned} j''(q)[\eta_1,\eta_2] = {}& \lambda(\eta_1,\eta_2)_{L^2(\Lambda,\varrho)} \\ &+ \int_I\int_\Omega v_1 v_2 - \nabla z(q)\cdot[\xi'(u)(v_1\mu\nabla v_2 + v_2\mu\nabla v_1) + \xi''(u)v_1 v_2\mu\nabla u], \end{aligned} \tag{4.8}$$

where $u = S(q)$, $v_i = S'(q)\eta_i$, $i = 1, 2$, and $z(q)$ denotes the adjoint state associated with the state $u = S(q)$. The mapping $\eta \mapsto j''(q)\eta^2$ is continuous and weakly lower semicontinuous on $L^2(\Lambda, \varrho)$. Moreover, if $q_n \to q$ in $L^s(\Lambda, \varrho)$ and $\eta_n \rightharpoonup \eta$ in $L^2(\Lambda, \varrho)$, then $j''(q_n)\eta_n^2 - j''(q)\eta_n^2 \to 0$ as $n \to \infty$.

Proof. Let $r > 2p/(p-d)$. We introduce the Lagrange function as

$$L(q, u, z) = J(u, q) - \int_I \langle \partial_t u + \mathcal{A}(u), z\rangle. \tag{4.9}$$

Since for $u = S(q)$ it holds $j(q) = L(u, q, z)$ for all $z \in L^{r'}(I; W_D^{1,p'})$, differentiating in (4.9) twice with respect to q in direction $\eta_i \in L^r(\Lambda, \varrho)$, and using Lemma 4.5 yields

$$\begin{aligned} j''(q)[\eta_1, \eta_2] = \lambda(\eta_1, \eta_2)_{L^2(\Lambda,\varrho)} + \int_I &\langle v_1, v_2\rangle + \langle u - \hat{u}, w\rangle \\ &- \langle \partial_t w + \mathcal{A}(u)w + \mathcal{A}'(u)w, z\rangle + \langle \mathcal{A}''(u)[v_1, v_2], z\rangle, \end{aligned}$$

where $v_i = S'(q)\eta_i$ and $w = S''(q)[\eta_1, \eta_2]$. Defining z to be the adjoint state as in Lemma 4.6, all terms involving w vanish and we obtain (4.8) for $\eta_i \in L^r(\Lambda, \varrho)$.

We would like to extend j'' to $L^2(\Lambda, \varrho)$, i.e. we have to argue that the expression in (4.8) is well-defined for $\eta_i \in L^2(\Lambda, \varrho)$. The only critical terms are those involving the adjoint state. According to Lemma 3.4 we have $u \in L^s(I; W_D^{1,p})$ with s from Assumption 4. Moreover, $z \in L^r(I; W_D^{1,p'})$ holds for any $r \in (1, \infty)$ due to Lemma 4.6.

For $\nabla z \cdot \xi'(u)v_1\mu\nabla v_2$ to be bounded, we have to require $v_1 \in L^{r_1}(I; L^\infty)$ with $r_1 = 2r/(r-2)$. Similarly, for $\nabla z \cdot \xi'(u)v_1v_2\mu\nabla u$ to be bounded, we have to require $v_1, v_2 \in L^{r_2}(I; L^\infty)$ with $r_2 = 2/(1 - 1/s - 1/r)$. Since $1 - 1/s > 1/2 + d/2p$, see Assumption 4, we infer $r_2 < 2/(1 - 1/s - 1/r)$.

Using the embedding (3.3), we deduce $v_1, v_2 \in L^{\tilde{r}}(I; L^\infty)$ with $\tilde{r} < 2p/d$. Thus, we need $r_1 < 2p/d$ and $2/(1 - 1/s - 1/r) < 2p/d$. Both inequalities are equivalent to $r > 2p/(p-d)$. Hence, for $z \in L^r(I; W_D^{1,p'})$ with $r > 2p/(p-d)$ the whole expression is well-defined, which proves continuity of $\eta \mapsto j''(q)\eta^2$ on $L^2(\Lambda, \varrho)$.

Moreover, if $\eta_n \rightharpoonup \eta$ in $L^2(\Lambda, \varrho)$, then $v_n = S'(q)\eta_n \rightharpoonup v$ in $L^2(I; W_D^{1,p})$ and $v_n \to v$ in $L^{r_1}(I; L^\infty)$ for $r_1 \in (1, 2p/d)$. Weak lower semicontinuity of $\|\cdot\|^2_{L^2(\Lambda,\varrho)}$ and continuity in v of the remaining parts yield

$$\begin{aligned} &\liminf_{n\to\infty} \lambda\|\eta_n\|^2_{L^2(\Lambda,\varrho)} + \int_I\int_\Omega v_n^2 - \nabla z(q) \cdot \left[2\xi'(\bar{u})v_n\mu\nabla v_n + \xi''(\bar{u})v_n^2\mu\nabla\bar{u}\right] \\ &\quad \geq \lambda\|\eta\|^2_{L^2(\Lambda,\varrho)} + \int_I\int_\Omega v^2 - \nabla z(q) \cdot \left[2\xi'(\bar{u})v\mu\nabla v + \xi''(\bar{u})v^2\mu\nabla\bar{u}\right] = j''(q)\eta^2. \end{aligned}$$

Concerning the last assertion, we find $S'(q_n)\eta_n \rightharpoonup S'(q)\eta_n$ in $W^{1,2}(I; W_D^{-1,p}) \cap L^2(I; W_D^{1,p})$ and $z(q_n) \to z(q)$ in $L^r(I; W_D^{1,p'})$ with $r > 2p/(p-d)$ as above, by means of Proposition 4.8. Hence, the compact embedding $W^{1,2}(I; W_D^{-1,p}) \cap L^2(I; W_D^{1,p}) \hookrightarrow_c L^{r_1}(I; L^\infty)$ with $r_1 \in (1, 2p/d)$ yields the result. □

For the second order optimality conditions we define the *critical cone* as

$$C_{\bar{q}} \coloneqq \left\{\eta \in L^2(\Lambda, \varrho) \colon \eta \text{ satisfies the sign condition (4.10) and } j'(\bar{q})\eta = 0\right\},$$

where the *sign condition* is given by

$$\eta \begin{Bmatrix} \leq 0 \text{ if } \bar{q} = q_b \\ \geq 0 \text{ if } \bar{q} = q_a \end{Bmatrix} \varrho\text{-almost everywhere in } \Lambda. \tag{4.10}$$

Then we obtain the usual second order necessary optimality condition.

Theorem 4.10. *Let $\bar{q} \in Q_{ad}$ be a local solution of (P). Then it holds*

$$j''(\bar{q})\eta^2 \geq 0, \quad \forall \eta \in C_{\bar{q}}.$$

Proof. The proof is completely analogous to the one of [15, Theorem 5.1] employing in particular Proposition 4.9. □

Second order sufficient optimality conditions are typically formulated using coercivity of j''. Indeed, for the given objective functional this is equivalent to the seemingly weaker positivity condition of j'', as already observed for semilinear parabolic PDEs in [18].

Theorem 4.11. *Let $\bar{q} \in Q_{ad}$ be given. The condition of positivity*

$$j''(\bar{q})\eta^2 > 0 \quad \forall \eta \in C_{\bar{q}} \setminus \{0\},$$

and the condition of coercivity

$$\exists \gamma > 0 \colon j''(\bar{q})\eta^2 \geq \gamma \|\eta\|^2_{L^2(\Lambda,\varrho)} \quad \forall \eta \in C_{\bar{q}},$$

are equivalent.

Proof. The proof is identical to the one of Theorem 4.11 in [18] except for the different structure of j'', where we use the formula given in Proposition 4.9. □

Theorem 4.12. *Let $\bar{q} \in Q_{ad}$. If $\bar{q}$ satisfies the first order necessary optimality conditions of Lemma 4.7 and in addition*

$$j''(\bar{q})\eta^2 > 0 \quad \forall \eta \in C_{\bar{q}} \setminus \{0\}, \tag{4.11}$$

then there exist constants $\varepsilon > 0$ and $\delta > 0$ such that the quadratic growth condition

$$j(\bar{q}) + \frac{\delta}{2}\|q - \bar{q}\|^2_{L^2(\Lambda,\varrho)} \leq j(q) \quad \forall q \in Q_{ad} \text{ s.t. } \|q - \bar{q}\|_{L^2(\Lambda,\varrho)} \leq \varepsilon, \tag{4.12}$$

is satisfied.

Proof. To prove this result, we apply [17, Theorem 2.3]. The delicate point is to verify assumption (A1), which is the continuous extension of j' and j'' to $L^2(\Lambda, \varrho)$. However, in our setting this is guaranteed due to Propositions 4.8 and 4.9. □

5. Application to stability analysis

As an application of the second order optimality conditions of Section 4 and the improved regularity of Section 3, we investigate the dependence of the optimal solution on perturbations of ξ. The stability analysis of optimal control problems is of independent interest, e.g., if the nonlinearity is not known exactly, cf. [55]. To this end, consider a family of perturbed nonlinearities $\xi_\varepsilon \in \Xi$ defined in Theorem 3.20 satisfying

$$\|\xi - \xi_\varepsilon\|_\infty \leq c\,\varepsilon, \quad \varepsilon > 0. \tag{5.1}$$

Note that due to uniform boundedness of the states in $C(\overline{I \times \Omega})$ Assumption (5.1) might be weakened to hold on compact subsets of $\mathbb{R}$. For ease of readability we rely on the stronger supposition. A similar problem subject to perturbations on the desired state $\hat{u}$ has been studied in [18, Section 4.4] for a semilinear heat equation. For any perturbed nonlinearity $\xi_\varepsilon \in \Xi$ fulfilling (5.1) let S_ε denote the associated control-to-state mapping. Define

$$j_\varepsilon \colon Q \to \mathbb{R}_0^+,\ q \mapsto J(q, S_\varepsilon(q)).$$

The perturbed optimal control problem reads as

$$(P_\varepsilon) \qquad \text{Minimize } j_\varepsilon(q) \text{ subject to } q \in Q_{ad}.$$

We first prove a general Lipschitz stability result of the control-to-state mapping.

Lemma 5.1. *There is a constant $c > 0$ independent of $\xi_\varepsilon \in \Xi$ such that for all $q \in Q_{ad}$ and $\eta \in L^2(\Lambda, \varrho)$ it holds*

$$\|S(q) - S_\varepsilon(q)\|_{W^{1,s}(I;W_D^{-1,p})\cap L^s(I;W_D^{1,p})} \le c\|\xi - \xi_\varepsilon\|_\infty. \tag{5.2}$$

If in addition $\xi_\varepsilon \in C^1(\mathbb{R})$, then

$$\|S'(q)\eta - S_\varepsilon'(q)\eta\|_{W^{1,2}(I;W_D^{-1,p})\cap L^2(I;W_D^{1,p})} \le c\left(\|\xi - \xi_\varepsilon\|_\infty + \|\xi' - \xi_\varepsilon'\|_\infty\right)\|\eta\|_{L^2(\Lambda,\varrho)}.$$

Proof. We denote in short $u = S(q)$ and $u_\varepsilon = S_\varepsilon(q)$. According to Theorem 3.20 all solutions u_ε are uniformly bounded in $W^{1,s}(I;H_D^{-\zeta,p}) \cap L^s(I;D)$, where we recall $D = \mathcal{D}_{H_D^{-\zeta,p}}(-\nabla\cdot\mu\nabla)$. Whence, all u_ε are contained in a compact subset of $C([0,T];W_D^{1,p}) \hookrightarrow C(\overline{I\times\Omega})$ due to embedding (3.11). Defining $w = u - u_\varepsilon$ we have

$$\partial_t w + \mathcal{A}(u)w - \nabla\cdot b_\varepsilon w\mu\nabla u_\varepsilon = \left[\mathcal{A}_\varepsilon(u_\varepsilon) - \mathcal{A}(u_\varepsilon)\right]u_\varepsilon, \quad w(0) = 0, \tag{5.3}$$

where

$$b_\varepsilon(t,x) \coloneqq \int_0^1 \xi'(u(t,x) + \tau(u_\varepsilon(t,x) - u(t,x)))\,\mathrm{d}\tau.$$

As in Proposition 4.4 we see that for each u_ε the left-hand side of (5.3) defines an isomorphism. Furthermore, using Lipschitz continuity of ξ' on bounded sets, we immediately infer that $u_\varepsilon \mapsto b_\varepsilon$ is continuous from $C(\overline{I\times\Omega})$ into itself. Whence, the mapping $u_\varepsilon \mapsto -\nabla\cdot b_\varepsilon\mu\nabla u_\varepsilon$ is continuous from $C([0,T];W_D^{1,p})$ into $\mathcal{L}(W^{1,s}(I;W_D^{-1,p}) \cap L^s(I;W_D^{1,p}); L^s(I;W_D^{-1,p}))$. Compactness of u_ε in $C([0,T];W_D^{1,p})$ yields uniformity of the norm of the solution operators to (5.3). Hence, we obtain the first assertion by estimating the right-hand side of (5.3) by

$$\|\left[\mathcal{A}_\varepsilon(u_\varepsilon) - \mathcal{A}(u_\varepsilon)\right]u_\varepsilon\|_{L^s(I;W_D^{-1,p})} \le \|\xi - \xi_\varepsilon\|_\infty \|\mu\|_{L^\infty(\Omega)}\|u_\varepsilon\|_{L^s(I;W_D^{1,p})},$$

as well as uniform boundedness of u_ε due to Theorem 3.20 and boundedness of Q_{ad}.

For the proof of the second estimate we set $\delta u = S'(q)\eta$ and $\delta u_\varepsilon = S_\varepsilon'(q)\eta$. Similar as above, compactness of u_ε in $C([0,T];W_D^{1,p})$ yields the uniform bound

$$\|\delta u_\varepsilon\|_{W^{1,2}(I;W_D^{-1,p})\cap L^2(I;W_D^{1,p})} \le c\|\eta\|_{L^2(\Lambda,\varrho)}.$$

The difference $\delta w = \delta u - \delta u_\varepsilon$ satisfies

$$\begin{aligned} \partial_t\delta w + \mathcal{A}(u)\delta w + \mathcal{A}'(u)\delta w &= \left[\mathcal{A}_\varepsilon(u_\varepsilon) - \mathcal{A}(u)\right]\delta u_\varepsilon + \left[\mathcal{A}_\varepsilon'(u_\varepsilon) - \mathcal{A}'(u)\right]\delta u_\varepsilon, \\ \delta w(0) &= 0, \end{aligned}$$

where $u = S(q)$ and $u_\varepsilon = S_\varepsilon(q)$. The terms on the right-hand side satisfy

$$\left[\mathcal{A}_\varepsilon(u_\varepsilon) - \mathcal{A}(u)\right]\delta u_\varepsilon = \left[\mathcal{A}_\varepsilon(u_\varepsilon) - \mathcal{A}(u_\varepsilon)\right]\delta u_\varepsilon - \nabla\cdot b_\varepsilon(u_\varepsilon - u)\mu\nabla\delta u_\varepsilon$$

and

$$[\mathcal{A}'_\varepsilon(u_\varepsilon) - \mathcal{A}'(u)]\,\delta u_\varepsilon = -\nabla \cdot \delta u_\varepsilon\,[\xi'_\varepsilon(u_\varepsilon) - \xi'(u_\varepsilon)]\,\mu\nabla u_\varepsilon \\ - \nabla \cdot b'_\varepsilon(u_\varepsilon - u)\mu\nabla u_\varepsilon - \nabla \cdot \delta u_\varepsilon \xi'(u)\mu\nabla\,[u_\varepsilon - u]\,,$$

where

$$b'_\varepsilon(t,x) \coloneqq \int_0^1 \xi''(u(t,x) + \tau(u_\varepsilon(t,x) - u(t,x)))\,\mathrm{d}\tau.$$

Whence, using (5.2) to bound $\|u_\varepsilon - u\|_{L^\infty(I\times\Omega)}$ we find

$$\|[\mathcal{A}_\varepsilon(u_\varepsilon) - \mathcal{A}(u)]\,\delta u_\varepsilon\|_{L^2(I;W_D^{-1,p})} \le c\|\xi_\varepsilon - \xi\|_\infty \|\delta u_\varepsilon\|_{L^2(I;W_D^{1,p})}$$

and, similarly,

$$\|[\mathcal{A}'_\varepsilon(u_\varepsilon) - \mathcal{A}'(u)]\,\delta u_\varepsilon\|_{L^2(I;W_D^{-1,p})} \le c\,(\|\xi_\varepsilon - \xi\|_\infty + \|\xi'_\varepsilon - \xi'\|_\infty)\,\|\delta u_\varepsilon\|_{L^r(I;L^\infty)},$$

where $1/2 = 1/r + 1/s$ and $r < 2p/d$ from embedding (3.3). Due to $s > 2p/(p-d)$, this is possible and maximal parabolic regularity yields the second bound. ☐

Applying a meanwhile standard localization argument, cf. [14], we introduce the auxiliary problem

$$\text{Minimize } j_\varepsilon(q) \text{ subject to } q \in Q_{ad} \cap \overline{\mathcal{B}_\rho(\bar q)}, \tag{5.4}$$

for $\rho > 0$ sufficiently small such that the second order sufficient optimality condition (4.11) holds. Existence of at least one solution follows by standard arguments.

Theorem 5.2. *Let $\bar q \in Q_{ad}$ be a locally optimal control of (P) satisfying the second order sufficient optimality conditions (4.11). There exist a sequence $(\bar q_\varepsilon)_\varepsilon$ of local solutions to (P_ε) and a constant $c > 0$ such that*

$$\|\bar q - \bar q_\varepsilon\|_{L^2(\Lambda,\varrho)} \le c\sqrt{\varepsilon}.$$

Proof. We set

$$F(q) \coloneqq \frac{1}{2}\|S(q) - \hat u\|^2_{L^2(I\times\Omega)}, \qquad F_\varepsilon(q) \coloneqq \frac{1}{2}\|S_\varepsilon(q) - \hat u\|^2_{L^2(I\times\Omega)}.$$

To begin with, let $(\bar q_\varepsilon)_\varepsilon$ denote a sequence of global solutions to (5.4). By optimality of $\bar q_\varepsilon$ for (5.4) and the quadratic growth condition (4.12) we obtain

$$j_\varepsilon(\bar q) \ge j_\varepsilon(\bar q_\varepsilon) = j(\bar q_\varepsilon) + F_\varepsilon(\bar q_\varepsilon) - F(\bar q_\varepsilon) \ge j(\bar q) + \frac{\delta}{2}\|\bar q_\varepsilon - \bar q\|^2_{L^2(\Lambda,\varrho)} + F_\varepsilon(\bar q_\varepsilon) - F(\bar q_\varepsilon).$$

Thus, using the definition of j and the Cauchy-Schwarz inequality we arrive at

$$\begin{aligned}\frac{\delta}{2}\|\bar q_\varepsilon - \bar q\|^2_{L^2(\Lambda,\varrho)} &\le j_\varepsilon(\bar q) - j(\bar q) + F(\bar q_\varepsilon) - F_\varepsilon(\bar q_\varepsilon) = F_\varepsilon(\bar q) - F(\bar q) + F(\bar q_\varepsilon) - F_\varepsilon(\bar q_\varepsilon)\\ &\le \frac{1}{2}\|S_\varepsilon(\bar q) + S(\bar q) - 2\hat u\|_{L^2(I\times\Omega)}\|S_\varepsilon(\bar q) - S(\bar q)\|_{L^2(I\times\Omega)}\\ &\quad + \frac{1}{2}\|S(\bar q_\varepsilon) + S_\varepsilon(\bar q_\varepsilon) - 2\hat u\|_{L^2(I\times\Omega)}\|S(\bar q_\varepsilon) - S_\varepsilon(\bar q_\varepsilon)\|_{L^2(I\times\Omega)}.\end{aligned}$$

Now, applying Lemma 5.1 and (5.1), we obtain

$$\frac{\delta}{2}\|\bar q_\varepsilon - \bar q\|^2_{L^2(\Lambda,\varrho)} \le c\|\xi - \xi_\varepsilon\|_\infty \le c\varepsilon,$$

where we have used that $S(q)$, respectively $S_\varepsilon(q)$, can be estimated independently of q due to Theorem 3.20 and boundedness of Q_{ad}. For ε small enough it is clear that $\bar q_\varepsilon$ is in the interior of $\overline{\mathcal{B}_\rho(\bar q)}$ and hence a local solution of (P_ε). ☐

Assuming differentiability of the nonlinearity ξ_ε, we are able to improve the estimate of Theorem 5.2. Precisely suppose that

$$\|\xi - \xi_\varepsilon\|_\infty + \|\xi' - \xi_\varepsilon'\|_\infty \le c\varepsilon, \quad \varepsilon > 0. \tag{5.5}$$

From the Lipschitz stability result of Lemma 5.1 we immediately infer

Corollary 5.3. *There is $c > 0$ such that for all $q \in Q_{ad}$ it holds*

$$|[j'(q) - j_\varepsilon'(q)]\,\eta| \le c\varepsilon\|\eta\|_{L^2(\Lambda,\varrho)}, \quad \eta \in L^2(\Lambda, \varrho).$$

Theorem 5.4. *Let $\bar q \in Q_{ad}$ be a locally optimal control of (P) that satisfies the second order sufficient optimality conditions (4.11). There exist a sequence $(\bar q_\varepsilon)_\varepsilon$ of local solutions to (P_ε) and constants $\varepsilon_0 > 0$ and $c > 0$ such that*

$$\|\bar q - \bar q_\varepsilon\|_{L^2(\Lambda,\varrho)} \le c\varepsilon, \quad \forall\, 0 < \varepsilon \le \varepsilon_0. \tag{5.6}$$

Proof. We proceed similarly to [16, Theorem 2.14] and argue by contradiction. Let $(\bar q_\varepsilon)_\varepsilon$ be a sequence of local solutions of (P_ε) from Theorem 5.2. Assume (5.6) is false, i.e. there exist sequences $(\varepsilon_n)_n$ with $\varepsilon_n \to 0$ and $(\bar q_{\varepsilon_n})_n$ with $\bar q_{\varepsilon_n} \in Q_{ad}$ and $\bar q_{\varepsilon_n} \to \bar q$, such that

$$\|\bar q_{\varepsilon_n} - \bar q\|_{L^2(\Lambda,\varrho)} > n\varepsilon_n, \quad n \in \mathbb{N}. \tag{5.7}$$

We define $\rho_n \coloneqq \|\bar q_{\varepsilon_n} - \bar q\|_{L^2(\Lambda,\varrho)}$ and $\eta_n \coloneqq \frac{1}{\rho_n}(\bar q_{\varepsilon_n} - \bar q)$, and may assume without restriction that $\eta_n \rightharpoonup \eta$ in $L^2(\Lambda, \varrho)$.

Step 1: $\eta \in C_{\bar q}$. Since $\bar q$ is a locally optimal control of (P), it holds

$$j'(\bar q)(q - \bar q) \ge 0 \quad \forall q \in Q_{ad}. \tag{5.8}$$

Whence, weak convergence of η_n implies $j'(\bar q)\eta = \lim_{n\to\infty} j'(\bar q)\eta_n \ge 0$. For the converse inequality, optimality of $\bar q_{\varepsilon_n}$ for (P_{ε_n}) implies

$$j_\varepsilon'(\bar q_{\varepsilon_n})(q - \bar q_{\varepsilon_n}) \ge 0 \quad \forall q \in Q_{ad}. \tag{5.9}$$

Therefore, we find

$$\begin{aligned} j'(\bar q)\eta &\le \limsup_{n\to\infty} j_{\varepsilon_n}'(\bar q_{\varepsilon_n})\eta_n + \limsup_{n\to\infty}[j'(\bar q_{\varepsilon_n}) - j_{\varepsilon_n}'(\bar q_{\varepsilon_n})]\eta_n + \limsup_{n\to\infty}[j'(\bar q) - j'(\bar q_{\varepsilon_n})]\eta_n \\ &\le c\limsup_{n\to\infty} \varepsilon_n\|\eta_n\|_{L^2(\Lambda,\varrho)} + c\limsup_{n\to\infty}\|\bar q_{\varepsilon_n} - \bar q\|_{L^2(\Lambda,\varrho)}\|\eta_n\|_{L^2(\Lambda,\varrho)}, \end{aligned}$$

where we have used Corollary 5.3 and Lipschitz continuity of j' in the last inequality. As $\bar q_{\varepsilon_n} \to \bar q$ in $L^2(\Lambda, \varrho)$ we conclude $j'(\bar q)\eta \le 0$. Whence, $j'(\bar q)\eta = 0$.

Since $\bar q_{\varepsilon_n} \in Q_{ad}$, we have $\eta_n \ge 0$ if $\bar q = q_a$ and $\eta_n \le 0$ if $\bar q = u_b$. Because $C_{\bar q}$ is convex and closed in $L^2(\Lambda, \varrho)$, it is weakly closed and the weak limit satisfies the sign condition (4.10) as well. Hence, $\eta \in C_{\bar q}$.

Step 2: $\eta = 0$. Using the optimality conditions (5.8) and (5.9) we obtain

$$[j'(\bar q_{\varepsilon_n}) - j'(\bar q)](\bar q_{\varepsilon_n} - \bar q) \le [j'(\bar q_{\varepsilon_n}) - j_\varepsilon'(\bar q_{\varepsilon_n})](\bar q_{\varepsilon_n} - \bar q) \le c\varepsilon_n\|\bar q_{\varepsilon_n} - \bar q\|_{L^2(\Lambda,\varrho)},$$

where we have used Corollary 5.3 in the last inequality. Taylor expansion yields

$$j''(\tilde q_n)(\bar q_{\varepsilon_n} - \bar q)^2 = [j'(\bar q_{\varepsilon_n}) - j'(\bar q)](\bar q_{\varepsilon_n} - \bar q) \le c\varepsilon_n\|\bar q_{\varepsilon_n} - \bar q\|_{L^2(\Lambda,\varrho)} \tag{5.10}$$

for some appropriate $\tilde q_n$. Employing weak lower semicontinuity of $\eta \mapsto j''(\bar q)\eta^2$, continuity of $q \mapsto j''(q)$ and (5.10) we conclude

$$\begin{aligned} j''(\bar q)\eta^2 &\le \liminf_{n\to\infty} j''(\bar q)\eta_n^2 = \liminf_{n\to\infty} j''(\tilde q_n)\eta_n^2 \\ &\le \limsup_{n\to\infty} j''(\tilde q_n)\eta_n^2 \le \limsup_{n\to\infty} \frac{c\varepsilon_n}{\rho_n} \le \limsup_{n\to\infty} \frac{c}{n} = 0 \end{aligned} \tag{5.11}$$

due to (5.7). Hence, the second order sufficient optimality conditions imply $\eta = 0$.

Step 3: Final contradiction. Since $\eta_n \rightharpoonup 0$ in $L^2(\Lambda, \varrho)$, the sequence $S'(\bar{q})\eta_n$ converges weakly to zero in $W^{1,2}(I; W_D^{-1,p}) \cap L^2(I; W_D^{1,p})$ and strongly to zero in $L^{r_1}(I; L^\infty)$ for $r_1 \in (1, 2p/d)$, due to the compact embedding (3.3). Therefore, the concrete expression (4.8) for the second derivative of j yields

$$\lim_{n\to\infty} j''(\bar{q})\eta_n^2 = \lambda \lim_{n\to\infty} \|\eta_n\|^2_{L^2(\Lambda,\varrho)}.$$

Because of $\|\eta_n\|^2_{L^2(\Lambda,\varrho)} = 1$, we find

$$0 < \lambda = \lambda \liminf_{n\to\infty} \|\eta_n\|^2_{L^2(\Lambda,\varrho)} = \liminf_{n\to\infty} j''(\bar{q})\eta_n^2 = 0,$$

where the last conclusion follows from (5.11). This completes the proof. □

Appendix A. Regularity of domains

For the geometric setting, we introduce:

$$K := \{x \in \mathbb{R}^d : |x| < 1\}, \qquad K_- := K \cap \{x \colon x_d < 0\},$$
$$\Sigma := K \cap \{x \colon x_d = 0\}, \qquad \Sigma_0 := \Sigma \cap \{x \colon x_{d-1} < 0\},$$

where x_i denotes the i-th component of $x \in \mathbb{R}^d$, $i \in \{1, \ldots, d\}$.

Definition A.1. Let $\Omega \subset \mathbb{R}^d$ and Γ_N a relatively open subset of $\partial\Omega$. The set $\Omega \cup \Gamma_N$ is called *Gröger regular* [34] if for any point $x \in \partial\Omega$ there exist an open neighborhood $U_x \subset \mathbb{R}^d$ of x and a bi-Lipschitz mapping ϕ_x from U_x onto αK such that $\phi_x(x) = 0$ and

$$\phi_x\left((\Omega \cup \Gamma_N) \cap U_x\right) \in \{\alpha K_-, \alpha(K_- \cup \Sigma), \alpha(K_- \cup \Sigma_0)\}.$$

In the paper, we require further regularity properties of the domain Ω. We give short proofs or references of these well-known results for convenience. Here $\mathcal{B}_\rho^{d-1}(y)$ denotes the open ball in $\mathbb{R}^{d-1}$ with radius $r > 0$ and center y, the symbol $\mathcal{H}_{d-1}$ stands for the $(d-1)$-dimensional Hausdorff measure and $\operatorname{dist}(z, M)$ denotes the distance of z to $M \subset \mathbb{R}^d$.

Proposition A.2. *If Ω is a Lipschitz domain, then Ω is a d-set* [40, Chapter II] *and $\partial\Omega$ is of class (A_α)* [42, Definition II.C.1].

Proof. Since K_- is a d-set and $\{x_d = 0\} \cap K$ is of class (A_α), this follows from the definition of Lipschitz domains, because bi-Lipschitz mappings preserve either properties [29, Chapter 2.4.1] and any finite union of d-sets (class (A_α)) is again a d-set (of class (A_α)). □

Proposition A.3. *If $\Omega \cup \Gamma_N$ is Gröger regular, then:*

(i) For all $x \in \overline{\Gamma_N}$, there is an open neighborhood U_x and a bi-Lipschitz mapping ϕ_x from a neighborhood of $\overline{U}_x$ onto an open subset of $\mathbb{R}^d$, such that $\phi_x(U_x) = K$, $\phi_x(\Omega \cap U_x) = K_-$, $\phi_x(\partial\Omega \cap U_x) = \Sigma$ and $\phi_x(0) = 0$,

(ii) For all $x \in \partial\Gamma_N$, there are $c_0 \in (0,1)$ and $c_1 > 0$ such that

$$\mathcal{H}_{d-1}\left(\left\{z \in \mathcal{B}_r^{d-1}(y) \colon \operatorname{dist}(z, \phi_x(\Gamma_N \cap U_x)) > c_0 r\right\}\right) \geq c_1 r^{d-1}$$

for all $r \in (0,1]$ and $y \in \mathbb{R}^{d-1}$ such that $(y, 0) \in \phi_x(\partial\Gamma_N \cap U_x)$ with ϕ_x and U_x as in (i),

(iii) Γ_D is a $(d-1)$-set.

Proof. Set $\hat{\phi}_x = 2\phi_x/\alpha$ and $\hat{U}_x = \hat{\phi}_x^{-1}(K)$. Then (i) follows from the fact, that bi-Lipschitz mappings pass inner points to inner points and boundary points to boundary points. Since $\hat{\phi}_x(\Gamma_N \cap \hat{U}_x) = \Sigma_0$ and $\hat{\phi}_x(\partial\Gamma_N \cap \hat{U}_x) = \Sigma \cap \{x\colon x_{d-1} = 0\}$ we infer (ii) by a direct calculation. (iii) is proved in [48, Theorem 4.3]. □

Appendix B. Exponential stability of the semigroups on $L^p(\Omega)$

Proposition B.1. *Let $0 < \mu_\bullet < \mu^\bullet$. For each $\omega \in [0, \mu_\bullet)$ it holds*

$$\|S_{-\nabla\cdot\mu\nabla+1}(z)\|_{\mathcal{L}(L^2)} \le \mathrm{e}^{-\omega\Re z}, \quad z \in \Sigma_\theta,$$

for all coefficient functions $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$, where $S_{-\nabla\cdot\mu\nabla+1}$ stands for the semigroup generated by $\nabla\cdot\mu\nabla - 1$ and

$$\theta = \arctan\left(\frac{\mu^\bullet + \omega}{\mu_\bullet - \omega}\right) \in (\pi/4, \pi/2).$$

Proof. Let a be the form associated with $-\nabla\cdot\mu\nabla + 1$, i.e.

$$a(u, v) = \int_\Omega \mu\nabla u \cdot \nabla v + \int_\Omega uv, \quad u, v \in W^{1,2}_D.$$

Since $\omega \in [0, \mu_\bullet)$ the form $b(u, v) \coloneqq a(u, v) - \omega(u, v)_{L^2}$ is coercive. Let B denote the operator associated with the form b. According to [10, Theorem 4.2] we conclude that $-B$ generates a sectorially contractive holomorphic C_0-semigroup S_B. Moreover, the angle of sectoriality θ is determined by the quotient of the continuity constant and coercivity constant of b, i.e. $\tan\theta = (\mu^\bullet + \omega)/(\mu_\bullet - \omega)$. Note that [10, Theorem 4.2] also holds in real spaces. Contractivity of S_B yields

$$\|S_{-\nabla\cdot\mu\nabla+1}(z)\|_{\mathcal{L}(L^2)} = \mathrm{e}^{-\omega\Re z}\|S_B(z)\|_{\mathcal{L}(L^2)} \le \mathrm{e}^{-\omega\Re z}, \quad z \in \Sigma_\theta,$$

due to the representation $-(-\nabla\cdot\mu\nabla + 1) = -B - \omega I$. □

Proposition B.2. *Let $p \in (1, \infty)$ and $0 < \mu_\bullet < \mu^\bullet$. There is $\omega > 0$ such that*

$$\|S_{-\nabla\cdot\mu\nabla+1}(z)\|_{\mathcal{L}(L^p)} \le \mathrm{e}^{-\omega z}, \quad z > 0,$$

for all coefficient functions $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$.

Proof. Due to Proposition B.1 there is $\omega > 0$ such that

$$\|S_{-\nabla\cdot\mu\nabla+1}(z)\|_{\mathcal{L}(L^2)} \le \mathrm{e}^{-\omega z}, \quad z > 0,$$

uniformly in $\mu \in \mathcal{M}_d(\mu_\bullet, \mu^\bullet)$. Furthermore, it holds $(1 \wedge |u|)\operatorname{sign} u = u - (u-1)^+ + (-u-1)^+ \in W^{1,2}_D$ for all $u \in W^{1,2}_D$; see proof of Theorem 3.1 in [59]. Thus, $\nabla\cdot\mu\nabla - 1$ generates a semigroup $S_{-\nabla\cdot\mu\nabla+1}$ of contractions on L^∞ according to [51, Theorem 4.9]. Due to [8, Lemma 2.1 (i)], the semigroup $S_{-\nabla\cdot\mu\nabla+1}$ interpolates on L^p for all $p \in [1, \infty]$. Hence, the assertion follows by interpolation. □

Appendix C. Sectorial operators and interpolation spaces

Definition C.1. A linear operator B on a Banach space X is called *sectorial* of angle $\theta \in (0, \pi)$, if there exists $C > 0$ such that

$$\sigma(B) \subset \overline{\Sigma_\theta}, \quad \|zR(z, B)\|_{\mathcal{L}(X)} \le C, \quad \forall z \in \mathbb{C} \setminus \overline{\Sigma_\theta}.$$

The following characterization of the real interpolation space is well-known. In our analysis we are particularly interested in the constants of equivalence of norms.

Proposition C.2. *Let B be a sectorial operator on X of angle $\theta \in (0, \pi)$. Then*

$$(X, D(B))_{\beta,\infty} = \Big\{ \varphi \in X \colon \|\varphi\|_{1,\beta,\infty} := \|\varphi\| + \sup_{z \in \mathbb{C} \setminus \overline{\Sigma_\theta}} |z|^\beta \|BR(z,B)\varphi\| < \infty \Big\}$$

for all $\beta \in (0,1)$ with equivalence of norms. Precisely,

$$\|\cdot\|_{\beta,\infty} \leq (C+2)\|\cdot\|_{1,\beta,\infty}, \quad \|\cdot\|_{1,\beta,\infty} \leq (C+1)(2C+1)\|\cdot\|_{\beta,\infty},$$

where C denotes the constant from Definition C.1 of sectoriality.

Proof. This follows from [47, Proposition 3.1.1] and the resolvent identity. □

Appendix D. Maximal parabolic regularity of nonautonomous operators

We remind the reader of the following definitions. Let $\{ A(t) \colon t \in [0,T] \}$ be a family of operators on a Banach space X and $\theta \in (0, \pi/2)$. We say that $A(\cdot)$ satisfies a resolvent estimate if there is $c > 0$ such that

$$\|R(z, A(t))\|_{\mathcal{L}(X)} \leq \frac{c}{1+|z|}, \quad z \in \mathbb{C} \setminus \overline{\Sigma_\theta},\ t \in [0,T]. \tag{D.1}$$

The family $A(\cdot)$ satisfies the *Acquistapace-Terreni condition* if there are constants $0 \leq \beta < \alpha < 1$ and $c > 0$ such that

$$\|A(t)R(z, A(t))[A(t)^{-1} - A(s)^{-1}]\|_{\mathcal{L}(X)} \leq c|t-s|^\alpha |z|^{\beta-1} \tag{D.2}$$

for all $t, s \in [0,T]$ and $z \in \mathbb{C} \setminus \overline{\Sigma_\theta}$. Moreover, $A(\cdot)$ is uniformly $\mathcal{R}$-sectorial if

$$\mathcal{R}\left(\left\{ zR(z, A(t)) \colon z \in \mathbb{C} \setminus \overline{\Sigma_\theta} \right\}\right) \leq c, \quad t \in [0,T]. \tag{D.3}$$

Lemma D.1. *Let X be an UMD space, $s \in (1,\infty)$ and $\mathcal{I}$ be a set. Consider families of operators $\{ A_\iota(t) \colon t \in [0,T] \}_{\iota \in \mathcal{I}}$ on X that satisfy (D.1), (D.2) and (D.3) with uniform constants. Then for each $\iota \in \mathcal{I}$ the operator A_ι possesses maximal parabolic regularity on $L^s((0,T);X)$ with norm independent of $\iota \in \mathcal{I}$, i.e. there exists $c > 0$ such that for all $f \in L^s((0,T);X)$, $u_0 \in (X, \mathcal{D}_X(A_\iota(0)))_{1-1/s,s}$ and all $\iota \in \mathcal{I}$ the solution u to*

$$\partial_t u + A_\iota u = f, \quad u(0) = u_0,$$

satisfies

$$\|\partial_t u\|_{L^s((0,T);X)} + \|A_\iota u\|_{L^s((0,T);X)} \leq c\|f\|_{L^s((0,T);X)}.$$

Proof. This is essentially the result of [53, Corollary 14] (cf. also [57, Satz 4.2.6]) except for the uniformity. We take a step back and consider the problem

$$\partial_t u + A_\iota u = f, \quad u(t_0) = 0, \tag{D.4}$$

on the interval $(t_0, t_1) \subseteq (0,T)$. If u is a solution to (D.4), then it holds

$$A_\iota(t)u(t) = (Q_\iota u)(t) + (S_\iota f)(t),$$

see, e.g., the heuristic derivation [2, p. 56f], where

$$(Q_\iota u)(t) := \int_{t_0}^t q_\iota(t,s)u(s)\,\mathrm{d}s := \int_{t_0}^t A_\iota(t)^2 S_{A_\iota(t)}(t-s)\left(A_\iota(t)^{-1} - A_\iota(s)^{-1}\right)u(s)\,\mathrm{d}s,$$

$$(S_\iota f)(t) := \int_{t_0}^t A_\iota(t) S_{A_\iota(t)}(t-s) f(s)\,\mathrm{d}s.$$

Step 1: $\|Q_\iota\| \leq 1/2$. According to [2, Lemma 2.3 (i)] the operator-valued kernel q_ι satisfies $\|q_\iota(t,s)\|_{\mathcal{L}(X)} \leq c(t-s)^{\alpha-\beta-1}$, where the constant depends proportionally on the constant c in (3.19). Whence, using Fubini's theorem we estimate

$$\|Q_\iota f\|_{L^1((t_0,t_1);X)} \leq c\int_{t_0}^{t_1}\int_{t_0}^{t}(t-s)^{\alpha-\beta-1}\|f(s)\|_X \leq c(t_1-t_0)^{\alpha-\beta}\|f\|_{L^1((t_0,t_1);X)}$$

for all $0 \leq t_0 < t_1 \leq T$. Similarly, we find

$$\|Q_\iota f\|_{L^\infty((t_0,t_1);X)} \leq c(t_1-t_0)^{\alpha-\beta}\|f\|_{L^\infty((t_0,t_1);X)}.$$

Interpolation yields $\|Q_\iota\|_{L^s((t_0,t_1);X)} \leq c(t_1-t_0)^{\alpha-\beta}$; see [61, Theorem 1.18.6.1]. Thus, there is $\tau > 0$ such that for all $I_0 = (t_0, t_0+\tau)$ with $t_0 \in [0,T)$ it holds $\|Q_\iota\|_{L^s(I_0;X)} \leq 1/2$ and $1-Q_\iota$ is invertible. Therefore, $A_\iota(\cdot)u = (1-Q_\iota)^{-1}S_\iota f$. For an alternative argument see [39, Section 3].

Step 2: Boundedness of S_ι. For maximal parabolic regularity we have to show that S_ι is bounded on $L^s(I_0;X)$ which is done in [53, Corollary 14] based on the operator-valued symbol associated with the resolvent $R(z,-A_\iota(t))$. Note that due to the supposition [53, Conditions (4),(5)] are uniform with respect to $\iota \in \mathcal{I}$.

We first consider the regular version [53, Theorem 6]. Its proof is based on [53, Proposition 11] stating that every symbol a has a Coifman-Meyer type decomposition. Concerning the constants, using Remark 3.10, we infer that C on page 813 [53] depends on the properties of the symbol [53, Definition 3], only. The definition $b_k(x,\xi) = a(x,2^k\xi)\phi_k(2^k\xi)$ and Remark 3.10 immediately yield C' with the same dependence. Employing [9, Lemma 2.3] we see that the estimate of [53, Proposition 11, (ii)] exclusively depends on C'. The remaining estimate with C'_β essentially uses [53, Condition (3)]. The decomposition is then used to define a bounded operator on $L^p(\mathbb{R}^n;X)$ by means of [53, Proposition 10]. In its proof we first use [53, Theorem 7] yielding constants that are independent of T_j. Then we apply Kahane's inequality (exclusively depending on p and X) and the $\mathcal{R}$-bound of D_k, but D_k depends on the dyadic partition of unity, only. Thereafter we use the definition of $T_jf(x)$. The $\mathcal{R}$-bound of a_j justifies the next inequality and we are left with terms that are independent of the symbol a. For the second part of the proof, the only point where the symbol enters is in the middle of page 811. There we use the estimate $|(I-\Delta_z)^m a_j(z)| \leq C2^{2j\delta m}$ due to [53, Proposition 10, (ii)].

Second, we consider the general version [53, Theorem 5] using Nagase's reduction to the smooth case. The symbol a is decomposed into $a = b+c$ [53, Proposition 13], where b is regular and c is treated by [53, Lemma 12]. In the second last estimates on pages 815 and 816 we use [53, Condition (4)], the third last estimate on page 817 uses [53, Condition (5)]. The remaining estimates are independent of a. Last, in the proof of [53, Lemma 12] the constant C exclusively depends on C_0 and C_α of the supposition and χ from the proof.

Step 3: Inhomogeneous initial data. Let $u_0 \in (X, \mathcal{D}_X(A_\iota(t_0)))_{1-1/s,s}$ and consider v_1 the solution to $\partial v_1 + A_\iota v_1 = 0$, $v_1(t_0) = u_0$. Then

$$\|v_2\|_{W^{1,s}(I_0;X)} \leq c\|u_0\|_{(X,\mathcal{D}_X(A_\iota(t_0)))_{1-1/s,s}}$$

according to [21, Lemma 2.1] that is based on [1, Lemma 2.2]. Carefully inspecting its proof we see that $c > 0$ depends on the resolvent estimate and the Acquistapace-Terreni condition and not on ι. Then, we have $(\partial_t + A_\iota, \gamma_0)^{-1}(f,u_0) = v_1 + v_2$, where v_2 is the solution to $(\partial_t + A_\iota)\,v_2 = f - (\partial_t + A_\iota)\,v_1$ with $v_2(t_0) = 0$.

Step 3: Maximal regularity on $(0, T)$. Consider a finite partition of $[0, T]$ into intervals $[t_i, t_{i+1}]$ each of length at most τ. We iterate the procedure above on each interval, where we use the terminal value $u(t_{i+1}) \in (X, \mathcal{D}_X(A_\iota(t_{i+1})))_{1-1/s,s}$ due to the embedding [21, Corollary 2.3] as the initial value for the next interval. $\square$

Acknowledgments

The authors are grateful to Joachim Rehberg and Hannes Meinlschmidt for discussions on maximal parabolic regularity for nonautonomous equations, to Moritz Egert for providing reference [26] used for uniform boundedness of the square root operators in Proposition 3.15, and to Pierre Portal for confirming the constant dependencies in [53]. Moreover, the second author acknowledges the support of her former host institution Technische Universität München.

References

[1] P. Acquistapace, Evolution operators and strong solutions of abstract linear parabolic equations, *Differential Integral Equations*, **1** (1988), 433–457.

[2] P. Acquistapace and B. Terreni, A unified approach to abstract linear nonautonomous parabolic equations, *Rend. Sem. Mat. Univ. Padova*, **78** (1987), 47–107, URL http://www.numdam.org/item?id=RSMUP_1987__78__47_0.

[3] N. U. Ahmed, Optimal control of a class of strongly nonlinear parabolic systems, *J. Math. Anal. Appl.*, **61** (1977), 188–207.

[4] H. Amann, Linear parabolic problems involving measures, *RACSAM. Rev. R. Acad. Cienc. Exactas Fs. Nat. Ser. A Mat.*, **95** (2001), 85–119.

[5] H. Amann, Maximal regularity for nonautonomous evolution equations, *Adv. Nonlinear Stud.*, **4** (2004), 417–430.

[6] H. Amann, Nonautonomous parabolic equations involving measures, *Journal of Mathematical Sciences*, **130** (2005), 4780–4802.

[7] H. Amann, *Linear and quasilinear parabolic problems. Vol. I*, vol. 89 of Monographs in Mathematics, Birkhäuser Boston, Inc., Boston, MA, 1995, Abstract linear theory.

[8] W. Arendt and A. F. M. ter Elst, Gaussian estimates for second order elliptic operators with boundary conditions, *J. Operator Theory*, **38** (1997), 87–130.

[9] W. Arendt and S. Bu, Tools for maximal regularity, *Math. Proc. Cambridge Philos. Soc.*, **134** (2003), 317–336.

[10] W. Arendt and A. F. M. ter Elst, From forms to semigroups, in *Spectral theory, mathematical system theory, evolution equations, differential and difference equations*, vol. 221 of Oper. Theory Adv. Appl., Birkhäuser/Springer Basel AG, Basel, 2012, 47–69.

[11] P. Auscher, N. Badr, R. Haller-Dintelmann and J. Rehberg, The square root problem for second-order, divergence form operators with mixed boundary conditions on L^p, *J. Evol. Equ.*, **15** (2015), 165–208.

[12] E. Casas and V. Dhamo, Optimality conditions for a class of optimal boundary control problems with quasilinear elliptic equations, *Control Cybernet.*, **40** (2011), 457–490.

[13] E. Casas, L. A. Fernndez and J. M. Yong, Optimal control of quasilinear parabolic equations, *Proc. Roy. Soc. Edinburgh Sect. A*, **125** (1995), 545–565.

[14] E. Casas and F. Tröltzsch, Error estimates for the finite-element approximation of a semilinear elliptic control problem, *Control Cybernet.*, **31** (2002), 695–712, Well-posedness in optimization and related topics (Warsaw, 2001).

[15] E. Casas and F. Tröltzsch, First- and second-order optimality conditions for a class of optimal control problems with quasilinear elliptic equations, *SIAM J. Control Optim.*, **48** (2009), 688–718.

[16] E. Casas and F. Tröltzsch, A general theorem on error estimates with application to a quasilinear elliptic optimal control problem, *Comput. Optim. Appl.*, **53** (2012), 173–206.

[17] E. Casas and F. Tröltzsch, Second order analysis for optimal control problems: improving results expected from abstract theory, *SIAM J. Optim.*, **22** (2012), 261–279.

[18] E. Casas and F. Tröltzsch, Second Order Optimality Conditions and Their Role in PDE Control, *Jahresber. Dtsch. Math.-Ver.*, **117** (2015), 3–44.

[19] J. C. de Los Reyes, P. Merino, J. Rehberg and F. Tröltzsch, Optimality conditions for state-constrained PDE control problems with time-dependent controls, *Control Cybernet.*, **37** (2008), 5–38.

[20] R. Denk, M. Hieber and J. Prüss, $\mathcal{R}$-boundedness, Fourier multipliers and problems of elliptic and parabolic type, *Mem. Amer. Math. Soc.*, **166** (2003), viii+114.

[21] D. Di Giorgio, A. Lunardi and R. Schnaubelt, Optimal regularity and Fredholm properties of abstract parabolic operators in L^p spaces on the real line, *Proc. London Math. Soc. (3)*, **91** (2005), 703–737.

[22] K. Disser, A. F. M. ter Elst and J. Rehberg, Hölder estimates for parabolic operators on domains with rough boundary, *Ann. Sc. Norm. Sup. Pisa*.

[23] K. Disser, H.-C. Kaiser and J. Rehberg, Optimal Sobolev regularity for linear second-order divergence elliptic operators occurring in real-world problems, *SIAM J. Math. Anal.*, **47** (2015), 1719–1746.

[24] K. Disser, A. F. M. ter Elst and J. Rehberg, On maximal parabolic regularity for non-autonomous parabolic operators, *J. Differential Equations*, **262** (2017), 2039–2072.

[25] X. T. Duong and D. W. Robinson, Semigroup kernels, Poisson bounds, and holomorphic functional calculus, *J. Funct. Anal.*, **142** (1996), 89–128.

[26] M. Egert, Elliptic systems with mixed boundary conditions: H^∞-calculus and the L^p Kato problem, Work in progress.

[27] M. Egert, R. Haller-Dintelmann and P. Tolksdorf, The Kato square root problem for mixed boundary conditions, *J. Funct. Anal.*, **267** (2014), 1419–1461.

[28] J. Elschner, J. Rehberg and G. Schmidt, Optimal regularity for elliptic transmission problems including C^1 interfaces, *Interfaces Free Bound.*, **9** (2007), 233–252.

[29] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[30] L. A. Fernández, Integral state-constrained optimal control problems for some quasilinear parabolic equations, *Nonlinear Anal.*, **39** (2000), 977–996.

[31] H. Gajewski, K. Gröger and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974, Mathematische Lehrbücher und Monographien, II. Abteilung, Mathematische Monographien, Band 38.

[32] J. A. Griepentrog, K. Gröger, H.-C. Kaiser and J. Rehberg, Interpolation for function spaces related to mixed boundary value problems, *Math. Nachr.*, **241** (2002), 110–120.

[33] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.

[34] K. Gröger, A $W^{1,p}$-estimate for solutions to mixed boundary value problems for second order elliptic differential equations, *Math. Ann.*, **283** (1989), 679–687.

[35] M. Haase, *The functional calculus for sectorial operators*, vol. 169 of Operator Theory: Advances and Applications, Birkhäuser Verlag, Basel, 2006.

[36] R. Haller-Dintelmann, C. Meyer, J. Rehberg and A. Schiela, Hölder continuity and optimal control for nonsmooth elliptic problems, *Appl. Math. Optim.*, **60** (2009), 397–428.

[37] R. Haller-Dintelmann, A. Jonsson, D. Knees and J. Rehberg, Elliptic and parabolic regularity for second- order divergence operators with mixed boundary conditions, *Mathematical Methods in the Applied Sciences*, **39** (2016), 5007–5026, Mma.3484.

[38] R. Haller-Dintelmann and J. Rehberg, Maximal parabolic regularity for divergence operators including mixed boundary conditions, *J. Differential Equations*, **247** (2009), 1354–1396.

[39] M. Hieber and S. Monniaux, Pseudo-differential operators and maximal regularity results for non-autonomous parabolic equations, *Proc. Amer. Math. Soc.*, **128** (2000), 1047–1053.

[40] A. Jonsson and H. Wallin, *Function spaces on subsets of* $\mathbb{R}^n$, vol. 2, 1984.

[41] T. Kato, *Perturbation theory for linear operators*, 2nd edition, Springer-Verlag, Berlin-New York, 1976, Grundlehren der Mathematischen Wissenschaften, Band 132.

[42] D. Kinderlehrer and G. Stampacchia, *An introduction to variational inequalities and their applications*, vol. 31 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000, Reprint of the 1980 original.

[43] K. Krumbiegel and J. Rehberg, Second order sufficient optimality conditions for parabolic optimal control problems with pointwise state constraints, *SIAM J. Control Optim.*, **51** (2013), 304–331.

[44] M. Krzek and P. Neittaanmäki, *Mathematical and numerical modelling in electrical engineering*, vol. 1 of Mathematical Modelling: Theory and Applications, Kluwer Academic Publishers, Dordrecht, 1996, Theory and applications, With a foreword by Ivo Babuška.

[45] P. C. Kunstmann and L. Weis, Maximal L_p-regularity for parabolic equations, Fourier multiplier theorems and H^∞-functional calculus, in *Functional analytic methods for evolution equations*, vol. 1855 of Lecture Notes in Math., Springer, Berlin, 2004, 65–311.

[46] J.-L. Lions, Optimisation pour certaines classes d'équations d'évolution non linéaires, *Ann. Mat. Pura Appl. (4)*, **72** (1966), 275–293.

[47] A. Lunardi, *Interpolation theory*, 2nd edition, Appunti. Scuola Normale Superiore di Pisa (Nuova Serie). [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], Edizioni della Normale, Pisa, 2009.

[48] H. Meinlschmidt and J. Rehberg, Hölder-estimates for non-autonomous parabolic problems with rough data, *Evolution Equations and Control Theory*, **5** (2016), 147–184.

[49] S. Nababan and E. S. Noussair, Existence theorems for optimal control of quasilinear parabolic partial differential equations, *J. Austral. Math. Soc. Ser. B*, **21** (1979/80), 90–101.

[50] E. S. Noussair, S. Nababan and K. L. Teo, On the existence of optimal controls for quasilinear parabolic partial differential equations, *J. Optim. Theory Appl.*, **34** (1981), 99–115.

[51] E. M. Ouhabaz, *Analysis of heat equations on domains*, vol. 31 of London Mathematical Society Monographs Series, Princeton University Press, Princeton, NJ, 2005.

[52] N. S. Papageorgiou, On the optimal control of strongly nonlinear evolution equations, *J. Math. Anal. Appl.*, **164** (1992), 83–103.

[53] P. Portal and Ž. Štrkalj, Pseudodifferential operators on Bochner spaces and an application, *Math. Z.*, **253** (2006), 805–819.

[54] J.-P. Raymond and F. Tröltzsch, Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints, *Discrete Contin. Dynam. Systems*, **6** (2000), 431–450.

[55] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer-Verlag, 1984.

[56] E. Shamir, Regularization of mixed second-order elliptic problems, *Israel J. Math.*, **6** (1968), 150–168.

[57] Ž. Štrkalj, *R-Beschränktheit, Summensätze abgeschlossener Operatoren und operatorwertige Pseudodifferentialoperatoren*, PhD thesis, Karlsruher Institut für Technologie, 2000.

[58] A. F. M. ter Elst and J. Rehberg, Hölder estimates for second-order operators on domains with rough boundary, *Adv. Differential Equations*, **20** (2015), 299–360, URL http://projecteuclid.org/euclid.ade/1423055203.

[59] A. F. M. ter Elst and J. Rehberg, L^∞-estimates for divergence operators on bad domains, *Anal. Appl. (Singap.)*, **10** (2012), 207–214.

[60] D. Tiba, Optimality conditions for distributed control problems with nonlinear state equation, *SIAM J. Control Optim.*, **23** (1985), 85–110.

[61] H. Triebel, *Interpolation theory, function spaces, differential operators*, vol. 18 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam-New York, 1978.

[62] F. Tröltzsch, *Optimal control of partial differential equations*, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010, Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[63] L. Weis, A new approach to maximal L_p-regularity, in *Evolution equations and their applications in physical and life sciences (Bad Herrenalb, 1998)*, vol. 215 of Lecture Notes in Pure and Appl. Math., Dekker, New York, 2001, 195–214.

[64] L. Weis, Operator-valued Fourier multiplier theorems and maximal L_p-regularity, *Math. Ann.*, **319** (2001), 735–758.

E-mail address: lucas.bonifacius@tum.de

E-mail address: neitzel@ins.uni-bonn.de