AVOIDING DEGENERACY IN THE WESTERVELT EQUATION BY STATE CONSTRAINED OPTIMAL CONTROL

Christian Clason* Barbara Kaltenbacher†

September 5, 2012

The Westervelt equation, which describes nonlinear acoustic wave propagation in high intensity ultrasound applications, exhibits potential degeneracy for large acoustic pressure values. While well-posedness results on this PDE have so far been based on smallness of the solution in a higher order spatial norm, non-degeneracy can be enforced explicitly by a pointwise state constraint in a minimization problem, thus allowing for pressures with large gradients and higher-order derivatives, as is required in the mentioned applications. Using regularity results on the linearized state equation, well-posedness and necessary optimality conditions for the PDE constrained optimization problem can be shown via a relaxation approach by Alibert and Raymond [Alibert and Raymond 1998].

1 INTRODUCTION

The propagation of high intensity focused ultrasound (HIFU) is often modeled by a nonlinear acoustic wave equation of the form

$$
\begin{cases}
(1-ky)y_{tt} - c^2\Delta y - b\Delta y_t + dy_t - k(y_t)^2 = 0 & \text{in } Q = (0,T)\times\Omega,\\
\partial_\nu y = u & \text{on } (0,T)\times\Gamma\\
y_t + c\partial_\nu y = 0 & \text{on } (0,T)\times\hat{\Gamma}\\
y = y^0,\ y_t = y^1 & \text{in } \{0\}\times\Omega
\end{cases}
\tag{1.1}
$$

*Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria (christian.clason@uni-graz.at).

†Institute of Mathematics, Alpen-Adria Universität Klagenfurt, Universitätsstr. 65–67 9020 Klagenfurt, Austria (barbara.kaltenbacher@aau.at).

for $\Omega \subseteq \mathbb{R}^n$, $n \in \{1,2,3\}$ (typically $n = 3$). Here, y is the acoustic pressure fluctuation, $c > 0$ is the speed of sound, $b > 0$, $d > 0$ are coefficients for strong and weak damping, respectively (related to the diffusivity of sound), $k = 2\beta_a/(rc^2)$, $r > 0$ is the mass density, and $\beta_a > 1$ is the parameter of nonlinearity. The control u acting on part of the boundary Γ models excitation of the normal derivative of the acoustic pressure by the normal acceleration of the piezoelectric transducers, which are typically arranged in a two dimensional array. The absorbing boundary conditions on the rest $\hat{\Gamma} := \partial\Omega \setminus \Gamma$ of the boundary are used to avoid reflections on the artificial boundary of the computational domain. For details we refer, e.g., to the original article [Westervelt 1963], as well as to [Clason, Kaltenbacher, and Veljovic 2009; Kaltenbacher and Lasiecka 2009] and the references therein.

The Westervelt equation (1.1) is not only nonlinear but in particular exhibits potential degeneracy due to the coefficient $(1-2ku)$ of u_{tt}. Therefore any well-posedness proof requires some estimate on $\|u\|_{L^\infty((0,T)\times\Omega}$ so that $1-2ku$ can be guaranteed to stay bounded away from zero. So far this has been achieved by deriving $C(0,T;H^2(\Omega))$ bounds on u (by means of energy estimates) and using Sobolev's embedding $H^2(\Omega) \to L^\infty(\Omega)$. The drawback of this approach is that it requires u to be small in $C(0,T;H^2(\Omega))$ (enforcing strong smoothness of the pressure distribution), whereas it should be sufficient to only have pointwise boundedness by a typically relatively large constant $\overline{m} < \frac{1}{2k}$. This is especially relevant in HIFU applications. Hence a major issue we wish to address is to improve the existing theory of $C(0,T;H^2(\Omega))$-small solutions and establish existence of large solutions (up to $\|y\|_{L^\infty((0,T)\times\Omega)} < \frac{1}{2k}$). The idea used here to achieve these goals is to explicitly impose the L^∞ bound as a state constraint in a minimization problem, in which the state is driven to possibly large sound pressure levels by boundary control (i.e., excitation) of a tracking type cost functional. Of course, minimization is not only a means for obtaining existence of large solutions, but of practical relevance on its own.

This paper is organized as follows. After introducing the precise problem formulation and some necessary notation in the remainder of this section, we discuss well-posedness of the state equation in Section 2. Section 3 is concerned with the existence of and first order optimality conditions for solutions of the state constrained control problem.

We consider the optimal control problem

$$\begin{cases} \min\limits_{y,u} \dfrac{1}{2}\displaystyle\int_0^T \|y(t)-y_d(t)\|^2_{L^2(\Omega)}\,dt + \dfrac{\alpha}{2}\displaystyle\int_0^T \|u(t)\|^2_{L^2(\Gamma)}\,dt =: J(u,y) \\ \text{s.t. (1.1).} \end{cases}$$

To avoid the degeneracy at $ky \geqslant 1$ and control large negative pressure values, we use pointwise state constraints:

$$-\underline{M}_y \leqslant y(x) \leqslant \overline{M}_y$$

for

$$0 < \underline{M}_y \quad \text{and} \quad 0 < \overline{M}_y < \frac{1}{k}. \tag{1.2}$$

We will assume from here on that the bounds (1.2) are satisfied. Moreover, large values of the quadratic term are prevented via the gradient constraint

$$\|y_t\|_{L^Q(0,T;L^P(\Omega))} \leqslant m_y,$$

where $m_y > 0$, $P, Q \in [1, \infty]$ are supposed to satisfy

$$m_y < \frac{\min\{b,d\}(1-k\overline{M}_y)Q}{20C_{H^1,L^6}(Q-1)^{1-1/Q}} \quad \text{and} \quad \frac{1}{3} = \frac{1}{2P} + \frac{1}{3Q}, \tag{1.3}$$

(e.g., $P = 2$, $Q = 4$), and C_{H^1,L^6} is the norm of the embedding $H^1(\Omega) \to L^6(\Omega)$.

In the following, for given initial data y^0, y^1, we denote the control space by

$$\mathcal{U} = \{u \in H^2(0,T;H^{-1/2}(\Gamma)) \,:\, u(0) = \partial_\nu y^0\}$$

and the unconstrained state space by

$$\mathcal{Y} = \{y \in \tilde{\mathcal{Y}} \,:\, y(0) = y^0,\ y_t(0) = y^1\}$$

where

$$\begin{aligned}\tilde{\mathcal{Y}} &= C(0,T;H^2(\Omega)) \cap C^1(0,T;H^1(\Omega)) \cap H^1(0,T;H^2(\Omega)) \\ &\qquad\qquad \cap C^2(0,T;L^2(\Omega)) \cap H^2(0,T;H^1(\Omega)) \\ &\hookrightarrow C(0,T;C(\Omega)).\end{aligned}$$

The constrained state space is

$$\mathcal{Y}_M = \left\{y \in \mathcal{Y} : -\underline{M}_y \leqslant y(t,x) \leqslant \overline{M}_y,\ \|y_t\|_{L^Q(0,T;L^P(\Omega))} \leqslant m_y\right\}$$

For future reference, we also introduce here the spaces

$$\begin{aligned}\mathcal{V} = \mathcal{W} &= L^2(0,T;L^2(\Omega)) \cap H^1(0,T;(H^1(\Omega))^*), \\ \mathcal{Z} &= C(0,T;C(\Omega)) \cap C^1(0,T;L^2(\Omega)), \\ \mathcal{Z}_M &= \left\{z \in \mathcal{Z} : 1 + k\underline{M}_y \geqslant z \geqslant 1 - k\overline{M}_y\right\}, \\ \check{\mathcal{Y}}^r &= C(0,T;H^2(\Omega)) \cap W^{1,r}(0,T;H^2(\Omega)) \cap C^2(0,T;L^2(\Omega)) \cap H^2(0,T;H^1(\Omega)), \\ \hat{\mathcal{Y}}^s &= L^s(0,T;H^2(\Omega)) \cap W^{1,s}(0,T;H^1(\Omega)), \\ \mathcal{W}^{r,p} &= L^r(0,T;L^2(\Omega)) \cap W^{1,p}(0,T;H^{-2\frac{p-1}{p}}(\Omega)), \\ \check{\mathcal{P}}^{r,p} &= (\mathcal{W}^{r,p})^{-*}, \\ \mathcal{P}^{\sigma,\theta} &= (C^\sigma(0,T;L^2(\Omega)) \cap H^\sigma(0,T;H^\theta(\Omega))) + (L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H^{\frac{1}{3}}(\Omega))),\end{aligned} \tag{1.4}$$

where V^{-*} denotes the predual of the normed vector space V, i.e., $(V^{-*})^* = V$.

2 STATE EQUATION

Similarly to Proposition 4 in [Clason, Kaltenbacher, and Veljovic 2009] or Theorem 2.1 in [Kaltenbacher and Lasiecka 2009], under the compatibility and smoothness conditions

$$
\begin{cases} (y^0, y^1) \in H^2(\Omega) \times H^2(\Omega), \\ c^2\Delta y^0 + b\Delta y^1 - dy^1 + k(y^1)^2 \in L^2(\Omega), \\ \partial_\nu y^0|_{\partial\Omega} = u(0) \end{cases} \tag{2.1}
$$

(the last condition being already incorporated in the definition of the space $\mathcal{U}$) one can show a well-posedness results for the following linear problem related to (1.1) for given z, w, u:

$$
\begin{cases} zy_{tt} - c^2\Delta y - b\Delta y_t + dy_t + w = 0 & \text{in } Q, \\ \partial_\nu y = u & \text{on } (0,T)\times\Gamma, \\ y_t + c\partial_\nu y = 0 & \text{on } (0,T)\times\hat{\Gamma}, \\ y = y^0,\ y_t = y^1 & \text{in } \{0\}\times\Omega, \end{cases} \tag{2.2}
$$

whose weak formulation is

$$
\begin{cases} \int_0^T \left\{ \int_\Omega \left(zy_{tt}v + c^2\nabla y\nabla v + b\nabla y_t\nabla v + dy_tv + wv \right) dx + \int_{\hat{\Gamma}} (cy_t + \frac{b}{c}y_{tt})v\, d\Gamma \right\} dt \\ \qquad = \int_0^T \int_\Gamma (c^2u + bu_t)v\, d\Gamma\, dt, \\ y(0) = y^0, \quad y_t(0) = y^1, \end{cases} \tag{2.3}
$$

for all $v \in L^2(0,T;H^1(\Omega))$.

Lemma 2.1. *For any*

- $r \in [1,\infty]$, $p \in [1,2]$,
- $z \in \mathcal{Z}_M$,
- $w \in \mathcal{W}^{r,p}$,
- $u \in \mathcal{U}$, *and*
- y^0, y^1 *satisfying* (2.1),

there exists a unique solution $y \in \check{\mathcal{Y}}^r$ *to* (2.2).

Furthermore, there exists a $C > 0$ *depending only on* $\|z\|_{C^1(0,T;L^2(\Omega))}$, $\underline{M}_y$, $\overline{M}_y$, *such that for all such* w, u, *and* y^0, y^1 *satisfying in addition* $c^2\Delta y^0 + b(\Delta y^1 - y^1) - w(0) \in L^2(\Omega)$,

$$
\|y\|_{\check{\mathcal{Y}}^r} \leqslant C\Big(\|w\|_{L^r(0,T;L^2(\Omega))} + \|w\|_{H^1(0,T;(H^1(\Omega))^*)} + \|y^0\|_{H^2(\Omega)} + \|y^1\|_{H^2(\Omega)} + \|u\|_{H^2(0,T;H^{-1/2}(\Gamma))} \Big). \tag{2.4}
$$

Proof. Existence of a solution follows analogously to the proof of Proposition 4 in [Clason, Kaltenbacher, and Veljovic 2009]. To see the energy estimate (2.4), we set $v = y_t$ in (2.3) and use the fact that

$$zy_{tt}y_t = \frac{1}{2}\frac{d}{dt}[z(y_t)^2] - \frac{1}{2}z_t(y_t)^2$$

to arrive at

$$\begin{aligned}(2.5)\quad \frac{1}{2}\left[\left\|\sqrt{z}y_t\right\|^2_{L^2(\Omega)} + c^2\left\|\nabla y\right\|^2_{L^2(\Omega)} + \frac{b}{c}\left\|y_t\right\|^2_{L^2(\hat{\Gamma})}\right]_0^t \\ + \int_0^t \left(b\left\|\nabla y_t\right\|^2_{L^2(\Omega)} + d\left\|y_t\right\|^2_{L^2(\Omega)} + c\left\|y_t\right\|^2_{L^2(\hat{\Gamma})}\right) d\tau \\ = \int_0^t \left\{\int_\Omega \left(\frac{1}{2}z_t(y_t)^2 - wy_t\right) dx + \int_\Gamma (c^2u + bu_t)y_t\, d\Gamma\right\} d\tau,\end{aligned}$$

where we can estimate

$$\int_\Omega z_t(y_t)^2\, dx \leqslant 2\epsilon\left\|y_t\right\|^2_{H^1(\Omega)} + C_\epsilon\left\|y_t\right\|^2_{L^2(\Omega)}\left\|z_t\right\|^4_{L^2(\Omega)}$$

with $C_\epsilon = \frac{27}{32}\frac{C^6_{H^1,L^6}}{\epsilon^3}$, $\epsilon = \frac{\underline{b}}{12}$, $\underline{b} = \min\{b, d\}$, and C_{H^1,L^6} again the embedding norm (see the proof of Theorem 2.1 in [Kaltenbacher and Lasiecka 2009]), as well as

$$\begin{aligned}\int_0^t\int_\Omega wy_t\, dx\, dt &\leqslant \|w\|_{L^p(0,T;H^{-2(p-1)/p}(\Omega))}\|y_t\|_{L^{p/(p-1)}(0,T;H^{2(p-1)/p}(\Omega))} \\ &\leqslant \|w\|_{L^p(0,T;H^{-2(p-1)/p}(\Omega))}\|y_t\|^{(2-p)/p}_{C(0,T;L^2(\Omega))}\|y_t\|^{2(p-1)/p}_{L^2(0,T;H^1(\Omega))} \\ &\leqslant \tilde{C}\|w\|^2_{L^p(0,T;H^{-2(p-1)/p}(\Omega))} + \frac{1-k\overline{M}_y}{8}\|y_t\|^2_{C(0,T;L^2(\Omega))} \\ &\quad + \frac{\underline{b}}{6}\|y_t\|^2_{L^2(0,T;H^1(\Omega))}\end{aligned}$$

for an appropriate constant $\tilde{C} > 0$, and

$$\int_0^t\int_\Gamma (c^2u + bu_t)y_t\, d\Gamma\, d\tau \leqslant \frac{3C^2_{tr}}{2\underline{b}}\|c^2u + bu_t\|^2_{L^2(0,T;H^{-1/2}(\Gamma))} + \frac{\underline{b}}{6}\|y_t\|^2_{L^2(0,T;H^1(\Omega))},$$

where C_{tr} is the norm of the trace operator $H^1(\Omega) \to H^{1/2}(\Gamma)$. This yields

$$\begin{aligned}(2.6)\quad \frac{1-k\overline{M}_y}{8}\|y_t\|^2_{C(0,T;L^2(\Omega))} + \frac{c^2}{8}\|\nabla y\|^2_{C(0,T;L^2(\Omega))} + \frac{\underline{b}}{4}\|y_t\|^2_{L^2(0,T;H^1(\Omega))} \\ \leqslant \frac{1}{2}\left[\left\|\sqrt{z(0)}y^1\right\|^2_{L^2(\Omega)} + c^2\left\|\nabla y^0\right\|^2_{L^2(\Omega)} + \frac{b}{c}\left\|y^1\right\|^2_{L^2(\hat{\Gamma})}\right] \\ + \tilde{C}\|w\|^2_{L^p(0,T;H^{-2(p-1)/p}(\Omega))} + \frac{3C^2_{tr}}{2\underline{b}}\|c^2u + bu_t\|^2_{L^2(0,T;H^{-1/2}(\Gamma))}\end{aligned}$$

$$+ C_{\frac{b}{12}} \|z_t\|^4_{C(0,T;L^2(\Omega))} \int_0^t \|y_t(t)\|^2_{L^2(\Omega)}$$

Moreover, we differentiate the PDE with respect to time and multiply with y_{tt} to obtain

$$\tag{2.7} \frac{1}{2}\left[\left\|\sqrt{z}y_{tt}\right\|^2_{L^2(\Omega)} + c^2\left\|\nabla y_t\right\|^2_{L^2(\Omega)} + \frac{b}{c}\left\|y_{tt}\right\|^2_{L^2(\hat{\Gamma})}\right]_0^t + \int_0^t \Big(b\left\|\nabla y_{tt}\right\|^2_{L^2(\Omega)} + d\left\|y_{tt}\right\|^2_{L^2(\Omega)} + c\left\|y_{tt}\right\|^2_{L^2(\Gamma)}\Big)\, d\tau = \int_0^t \Big\{\int_\Omega \Big(-\frac{1}{2}z_t(y_{tt})^2 - w_t y_{tt}\Big)\, dx + \int_\Gamma (c^2u_t + bu_{tt})y_{tt}\, d\Gamma\Big\}\, d\tau,$$

and proceed analogously to above to obtain

$$\tag{2.8} \begin{aligned} &\frac{1-k\overline{M}_y}{8}\|y_{tt}\|^2_{C(0,T;L^2(\Omega))} + \frac{c^2}{8}\|\nabla y_t\|^2_{C(0,T;L^2(\Omega))} + \frac{b}{4}\|y_{tt}\|^2_{L^2(0,T;H^1(\Omega))} \\ &\quad\leqslant \frac{1}{2}\left[\left\|\sqrt{z(0)}y_{tt}(0)\right\|^2_{L^2(\Omega)} + c^2\left\|\nabla y^1\right\|^2_{L^2(\Omega)} + \frac{b}{c}\left\|y_{tt}(0)\right\|^2_{L^2(\hat{\Gamma})}\right] \\ &\quad + \tilde{C}\|w_t\|^2_{L^p(0,T;H^{-2(p-1)/p}(\Omega))} + \frac{3C_{tr}^2}{2\underline{b}}\|c^2u_t + bu_{tt}\|^2_{L^2(0,T;H^{-1/2}(\Gamma))} \\ &\qquad + C_{\frac{b}{12}}\|z_t\|^4_{C(0,T;L^2(\Omega))}\int_0^t \|y_{tt}(t)\|^2_{L^2(\Omega)}, \end{aligned}$$

which by $z \geqslant 1-k\overline{M}_y > 0$ and Gronwall's inequality gives the $C^1(0,T;H^1(\Omega))\cap C^2(0,T;L^2(\Omega))\cap H^2(0,T;H^1(\Omega))$ part of the estimate.

To get regularity in space, we multiply the PDE with $-\Delta y$ and obtain

$$\tag{2.9} c^2\int_0^t \|\Delta y\|_{L^2(\Omega)}\, d\tau + \frac{1}{2}\left[b\|\Delta y\|_{L^2(\Omega)} + d\|\nabla y\|^2_{L^2(\Omega)}\right]_0^t + \frac{d}{c}\int_0^t \|y_t\|^2_{L^2(\hat{\Gamma})}\, d\tau = \int_0^t \Big\{\int_\Omega \Big(zy_{tt}\Delta y + w\Delta y\, dx\Big) + d\int_\Gamma uy_t\, d\Gamma\Big\}\, d\tau,$$

hence

$$\begin{aligned} &\frac{c^2}{8}\|\Delta y\|^2_{L^2(0,T;L^2(\Omega))} + \frac{b}{8}\left[\|\Delta y\|^2_{C(0,T;L^2(\Omega))} + \|\nabla y\|^2_{C(0,T;L^2(\Omega))}\right] \\ &\quad\leqslant \left[\left\|\Delta y^0\right\|_{L^2(\Omega)} + \left\|\nabla y^0\right\|^2_{L^2(\Omega)}\right] + \frac{1}{2c^2}(1 + k\max\{\underline{M}_y, \overline{M}_y\})\|y_{tt}\|^2_{L^2(0,T;L^2(\Omega))} \\ &\qquad + \frac{1}{\underline{b}}\|w\|^2_{\mathcal{M}(0,T;L^2(\Omega))} + \frac{C_{tr}^2}{2}\int_0^t \|y_t(\tau)\|^2_{H^1(\Omega)} + \frac{d^2}{2}\|u\|^2_{L^2(0,T;H^{-1/2}(\Gamma))}. \end{aligned}$$

Finally, resolving the PDE with respect to the strong damping term $b\Delta y_t = zy_{tt} - c^2\Delta y + dy_t + w \in L^r(0,T;L^2(\Omega))$, we arrive at $W^{1,r}(0,T;H^2(\Omega))$ regularity. □

In the following we will derive some regularity results for the solution of the adjoint PDE, which will be part of the optimality system and whose right hand side is composed of a measure-valued part and a part in the dual of $W^{1,Q}(0,T;L^P(\Omega))$.

By duality we first of all conclude a regularity result for the adjoint PDE with the right hand side being measure-valued or a derivative (via the differential operator D) of a measure.

Corollary 2.2. *Fix* $r \in [1,\infty]$, $p \in [1,2]$, *and let* D *be an arbitrary bounded linear operator* $D : \check{\mathcal{Y}}^r \to C(0,T;C(\Omega))$ *(e.g.,* $(-\Delta_{t,x})^\alpha$ *with* $\alpha \in (0,\frac{1}{4})$*). For any*

- $z \in \mathcal{Z}_M$,
- $\mu \in \mathcal{M}(0,T;\mathcal{M}(\Omega))$,

any solution p *to*

$$\left\{\begin{aligned} &\int_0^T \Big\{ \int_\Omega \Big((zv)_{tt} p + c^2 \nabla v \nabla p + b \nabla v_t \nabla p + d v_t p \Big)\, dx \\ &\quad + \int_{\hat{\Gamma}} (c v_t + \frac{b}{c} v_{tt}) p \, d\Gamma \Big\}\, dt = \langle \mu, Dv \rangle_{\mathcal{M},C} = \langle D^* \mu, v \rangle_{\check{\mathcal{Y}}^{r*},\check{\mathcal{Y}}^r} \end{aligned}\right. \tag{2.10}$$

for all $v \in \check{\mathcal{Y}}^r$ *is contained in the space* $\check{\mathfrak{P}}^{r,p}$ *defined as the predual of* $\mathcal{W}^{r,p}$*, cf. (1.4).*

Furthermore, there exists a $C > 0$ *depending only on* $\|z\|_{C^1(0,T;L^2(\Omega))}$, $\underline{M}_y$, $\overline{M}_y$*, such that for all such* μ,

$$\|p\|_{\check{\mathfrak{P}}^{r,p}} \leqslant C \|\mu\|_{\mathcal{M}(0,T;\mathcal{M}(\Omega))}.$$

If some subspace H *of* $\mathcal{W}^{r,p}$ *is dense in* $D^*\mathcal{M}(0,T;\mathcal{M}(\Omega))$ *with respect to the topology of* $\check{\mathcal{Y}}^{r*}$*, then for all* $\mu \in \mathcal{M}(0,T;\mathcal{M}(\Omega))$ *a solution* $p \in \check{\mathfrak{P}}^{r,p}$ *exists.*

Proof. To obtain an estimate of p in $\check{\mathfrak{P}}^{r,p}$, we insert the solution v^* of

$$\left\{\begin{aligned} &\int_0^T \Big\{ \int_\Omega \Big((zv^*)_{tt} \varphi + c^2 \nabla v \nabla \varphi + b \nabla v_t^* \nabla \varphi + d v_t^* \varphi \Big)\, dx \\ &\qquad + \int_{\hat{\Gamma}} (c v_t^* + \frac{b}{c} v_{tt}^*) \varphi \, d\Gamma \Big\}\, dt = \langle J_2(p), \varphi \rangle_{\check{\mathfrak{P}}^{r,p*},\check{\mathfrak{P}}^{r,p}}\,, \\ &v^*(0) = 0, \quad v_t^*(0) = 0, \end{aligned}\right. \tag{2.11}$$

for all $\varphi \in L^2(0,T;H^1(\Omega))$ as a test function into (2.10), where $J_2 : \check{\mathfrak{P}}^{r,p} \to \check{\mathfrak{P}}^{r,p*} = \mathcal{W}^{r,p}$ is the duality mapping with gauge function $t \to \frac{1}{2}t^2$, i.e., such that

$$\|J_2(p)\|^2_{\check{\mathfrak{P}}^{r,p*}} = \|p\|^2_{\check{\mathfrak{P}}^{r,p}} = \langle J_2(p), p \rangle_{\check{\mathfrak{P}}^{r,p*},\check{\mathfrak{P}}^{r,p}} \tag{2.12}$$

Indeed, analogously to Lemma 2.1, $v^* \in \check{\mathcal{Y}}^r$ and

$$\|v^*\|_{\check{\mathcal{Y}}^r} \leqslant C \|J_2(p)\|_{\check{\mathfrak{P}}^{r,p*}} = C \|p\|_{\check{\mathfrak{P}}^{r,p}}. \tag{2.13}$$

Therefore, by (2.10), (2.11), (2.12) and (2.13) we get

$$\begin{aligned} \|p\|^2_{\check{\mathcal{P}}^{r,p}} &= \langle J_2(p), p\rangle_{\check{\mathcal{P}}^{r,p*},\check{\mathcal{P}}^{r,p}} = \langle \mu, Dv^*\rangle_{\mathcal{M},C} \\ &\leqslant \|D\|_{\check{\mathcal{Y}}^r\to C(0,T;C(\Omega))}\|\mu\|_{\mathcal{M}(0,T;\mathcal{M}(\Omega))}\|v^*\|_{\check{\mathcal{Y}}^r} \\ &\leqslant C\|D\|_{\check{\mathcal{Y}}^r\to C(0,T;C(\Omega))}\|\mu\|_{\mathcal{M}(0,T;\mathcal{M}(\Omega))}\|p\|_{\check{\mathcal{P}}^{r,p}}. \end{aligned} \tag{2.14}$$

Existence of p can be obtained by considering an approximating sequence

$$(D^*\mu_k)_{k\in\mathbb{N}} \subseteq H \subseteq \mathcal{W}^{r,p}$$

converging to $D^*\mu$ in $\check{\mathcal{Y}}^{r*}$. Similarly to Lemma 2.1 one sees that for all $k \in \mathbb{N}$ a solution p_k of (2.10) with $\mu := \mu_k$ exists, and by (2.14) the sequence $(p_k)_{k\in\mathbb{N}}$ is bounded in $\check{\mathcal{P}}^{r,p}$, which is reflexive for $p, r \in (1,\infty)$. Thus, taking limits along a weakly convergent subsequence one arrives at a solution to (2.10). □

Similarly to Corollary 2.2 we get

Corollary 2.3. *Fix* $P, Q, r \in [1,\infty]$, $p \in [1,2]$, *and let* D *be an arbitrary bounded linear operator* $D : \check{\mathcal{Y}}^r \to W^{1,Q}(0,T; L^P(\Omega))$. *For any*

- $z \in \mathcal{Z}_M$,
- $\mu \in W^{-1,\frac{Q}{Q-1}}(0,T; L^{\frac{P}{P-1}}(\Omega))$,

any solution p *to*

$$\left\{ \begin{aligned} &\int_0^T \Big\{ \int_\Omega \Big((zv)_{tt}p + c^2\nabla v\nabla p + b\nabla v_t \nabla p + dv_t p\Big)\, dx + \int_{\hat{\Gamma}} (cv_t + \frac{b}{c}v_{tt})p\, d\Gamma \Big\}\, dt \\ &\qquad = \langle \mu, Dv\rangle_{W^{-1,\frac{Q}{Q-1}}(0,T;L^{\frac{P}{P-1}}(\Omega)),W^{1,Q}(0,T;L^P(\Omega))} = \langle D^*\mu, v\rangle_{\check{\mathcal{Y}}^{r*},\check{\mathcal{Y}}^r} \end{aligned}\right.$$

for all $v \in \check{\mathcal{Y}}$ *is contained in the space* $\check{\mathcal{P}}^{r,p}$.

Furthermore, there exists a $C > 0$ *depending only on* $\|z\|_{C^1(0,T;L^2(\Omega))}$, $\underline{M}_y$, $\overline{M}_y$, *such that for all such* μ,

$$\|p\|_{\check{\mathcal{P}}^{r,p}} \leqslant C\|\mu\|_{W^{-1,\frac{Q}{Q-1}}(0,T;L^{\frac{P}{P-1}}(\Omega))}.$$

If some subspace H *of* $\mathcal{W}^{r,p}$ *is dense in* $D^*W^{-1,\frac{Q}{Q-1}}(0,T; L^{\frac{P}{P-1}}(\Omega))$ *with respect to the topology of* $\check{\mathcal{Y}}^{r,*}$, *then for all* $\mu \in W^{-1,\frac{Q}{Q-1}}(0,T; L^{\frac{P}{P-1}}(\Omega))$ *a solution* $p \in \check{\mathcal{P}}^{r,p}$ *exists.*

By interpolation we obtain the following intermediate result between Lemma 2.1 and Corollary 2.2 which will be useful for establishing regularity of the adjoint state later on.

Corollary 2.4. *Fix* $0 < \sigma < \frac{3}{\sqrt{8}} - 1$, $0 < \theta < \frac{1}{\sqrt{8}}$, *and* $P \in [1,2]$, $Q \in [1,\infty]$ *satisfying*

$$\frac{n}{2} - \frac{n}{P} - \frac{2}{Q} \leqslant 0. \tag{2.15}$$

For any

- $z \in \mathcal{Z}_M$,
- $\mu \in \mathcal{M}(0,T;\mathcal{M}(\Omega)) + W^{-1,\frac{Q}{Q-1}}(0,T;L^{\frac{P}{P-1}}(\Omega))$,

a solution p *to*

$$\left\{ \int_0^T \left\{ \int_\Omega \Big((zv)_{tt}p + c^2\nabla v\nabla p + b\nabla v_t\nabla p + dv_t p \Big)\,dx + \int_{\hat\Gamma} (cv_t + \frac{b}{c}v_{tt})p\,d\Gamma \right\} dt \right. \\ = \langle \mu, v\rangle_{\mathcal{M},C}$$

for all $v \in \check{\mathcal{Y}}$ *exists and is contained in* $\mathcal{P}^{\sigma,\theta}$ *as in* (1.4).

Furthermore, there exists a $C > 0$ *depending only on* $\|z\|_{C^1(0,T;L^2(\Omega))}$, $\underline{M}_y$, $\overline{M}_y$, *such that for all such* μ,

$$\|p\|_{\mathcal{P}^{\sigma,heta}} \leqslant C\|\mu\|_{\mathcal{M}(0,T;\mathcal{M}(\Omega))+W^{-1,\frac{Q}{Q-1}}(0,T;L^{\frac{P}{P-1}}(\Omega))}.$$

Proof. The proof is based on the exact interpolation theorem [Adams and Fournier 2003, Theorem 7.23], which states that boundedness of

$$\mathcal{S} : X_0 \to Y_0 \quad \text{and} \quad \mathcal{S} : X_1 \to Y_1$$

implies boundedness of

$$\mathcal{S} : (X_0, X_1)_{\theta,q} \to (Y_0, Y_1)_{\theta,q},$$

where $\theta \in (0,1)$, $q \in (1,\infty)$ (these restrictions can be relaxed to include $q \in \{1,\infty\}$ under certain conditions, but this will not be needed here). Here X_0, X_1 are continuously embedded in a common Hausdorff topological space, likewise for Y_0, Y_1, and the interpolation spaces are defined either by the J- or by the K- method of real interpolation (see, e.g., [Adams and Fournier 2003, Chapter 7]).

Consider now the solution operator $\mathcal{S}$ that maps the right hand side f to a (weak) solution p of the adjoint equation

$$\begin{cases} zp_{tt} - c^2\Delta p + b\Delta p_t - dp_t = f & \text{in } Q, \\ \partial_\nu p = 0 & \text{on } (0,T)\times\Gamma \\ -p_t + c\partial_\nu p = 0 & \text{on } (0,T)\times\hat\Gamma \\ p = 0,\ p_t = 0 & \text{in } \{T\}\times\Omega \end{cases}$$

(Here we take an arbitrary but fixed selection from the solution set.)

Consider first of all the part of μ that lies in $\mathcal{M}(0,T;\mathcal{M}(\Omega))$. Corollary 2.2 with $D = D_t^\alpha D_x^\beta$, $r = \infty$, and $L^1(0,T;L^1(\Omega)) \subseteq \mathcal{M}(0,T;\mathcal{M}(\Omega))$ yields boundedness of

$$\mathcal{S} : W^{-\alpha,1}(0,T;W^{-\beta,1}(\Omega)) \to \check{\mathcal{P}}^{\infty,p} = (W^{\infty,p})^{-*} \\ \subseteq (W^{1,p}(0,T;L^2(\Omega)))^{-*} = W^{-1,\frac{p}{p-1}}(0,T;L^2(\Omega)), \tag{2.16}$$

for any $p \in (1,2]$ (so that the space $W^{1,p}(0,T;L^2(\Omega))$ on the right hand side is reflexive) and for any

$$\alpha \in [0,1), \quad \beta \in \left[0, 2-\frac{n}{2}\right). \tag{2.17}$$

Here, we have used Sobolev's embedding

$$D\check{\mathcal{Y}}^r \subseteq W^{1-\alpha,\infty}(0,T;H^{2-\beta}) \subseteq C(0,T;C(\Omega)).$$

Similarly to the lower and higher order in time energy estimates (2.6), (2.8) in the proof of Lemma 2.1, we have boundedness of

$$\mathcal{S} : W^{1,\tilde{p}}(0,T;H^{-2\frac{\tilde{p}-1}{\tilde{p}}}(\Omega)) \to C^2(0,T;L^2(\Omega)) \cap H^2(0,T;H^1(\Omega)) \tag{2.18}$$

for any $\tilde{p} \in [1,2]$.

For applying interpolation to (2.16) and (2.18), we consider appropriate sub- and superspaces in order to

- guarantee $\mathcal{M}(0,T;\mathcal{M}(\Omega)) \subseteq X_1$,
- avoid the (pre)dual of an intersection in the definition of Y_0,
- work with Sobolev spaces instead of spaces of continuous functions, and
- match the Lebesgue space indices between spaces to be interpolated (which is required if they have different smoothness index).

Specifically, instead of (2.16) and (2.18) we use boundedness of

$$\mathcal{S} : W^{-\alpha,\tilde{p}}(0,T;W^{-\beta,\hat{p}}(\Omega)) \to W^{-1,\frac{p}{p-1}}(0,T;L^2(\Omega)) \tag{2.19}$$

and of

$$\mathcal{S} : W^{1,\tilde{p}}(0,T;W^{-2\frac{\tilde{p}-1}{\tilde{p}}-\frac{n}{2}+\frac{n}{\hat{p}},\hat{p}}(\Omega)) \to \underbrace{W^{2,\tilde{r}}(0,T;L^2(\Omega))}_{a)} \cap \underbrace{H^2(0,T;H^1(\Omega))}_{b)} \tag{2.20}$$

for any $p, \tilde{p}, \hat{p} \in (1,2]$, $\tilde{r} = \frac{p}{p-1} \in [2,\infty)$, $\tilde{q} \in (1,\infty)$. The exact interpolation theorem applied to (2.19) and (2.20) then yields boundedness of

$$\begin{aligned} \mathcal{S} : B^{-\alpha+\theta(1+\alpha),\tilde{p},\tilde{q}}&(0,T;B^{-\beta+\theta(\beta+\frac{n}{\hat{p}}-\frac{n}{2}-2\frac{\tilde{p}-1}{\tilde{p}}),\hat{p},\tilde{q}}(\Omega)) \\ &\to B^{-1+3\theta,\tilde{r},\tilde{q}}(0,T;B^{0,2,\tilde{q}}(\Omega)) \cap B^{-1+3\theta,2,\tilde{q}}(0,T;B^{\theta,2,\tilde{q}}(\Omega)) \end{aligned} \tag{2.21}$$

(where we have interpolated separately for a) and b) in (2.20)). Here $B^{s,\overline{p},\tilde{q}}$ denotes the Besov space of order s over $L^{\overline{p}}$ with interpolation index $\tilde{q}$ in the J-method of interpolation; cf. [Adams and Fournier 2003, Chapter 7]. The choice of θ in (2.21) is dictated by the need for embedding $\mathcal{M}(0,T;\mathcal{M}(\Omega))$ into the preimage space of S using the continuous and dense embeddings

$$
\begin{aligned}
&B^{s,\frac{P}{P-1},\frac{\tilde{q}}{\tilde{q}-1}}(0,T) \hookrightarrow C(0,T) && \text{for} \quad s\frac{P}{P-1} > 1, \\
&B^{t,\frac{Q}{Q-1},\frac{\tilde{q}}{\tilde{q}-1}}(\Omega) \hookrightarrow C(\Omega) && \text{for} \quad t\frac{Q}{Q-1} > n
\end{aligned}
$$

which imply

$$
B^{-s,P,\tilde{q}}(0,T) \hookrightarrow \mathcal{M}(0,T) \qquad \text{and} \qquad B^{-t,Q,\tilde{q}}(\Omega) \hookrightarrow \mathcal{M}(\Omega)
$$

(cf. [Adams and Fournier 2003, Theorem 7.34] for the embedding result), and the restrictions (2.17) on α, β.

Since we can let $\frac{\hat{p}}{\hat{p}-1} \to \infty$, and set $\tilde{p} = \sqrt{2}$, we may choose $\theta \in [1, \frac{1}{\sqrt{8}})$ close to $\frac{1}{\sqrt{8}}$. Indeed, fixing $\epsilon > 0$ arbitrarily small, setting $\alpha = 1-\epsilon$, $\beta = 2 - \frac{n}{2} - \epsilon$, $\frac{\hat{p}}{\hat{p}-1} = \frac{1}{\epsilon}$, $\theta = \theta_\epsilon = \min\left\{\frac{1-(\sqrt{2}+1)\epsilon}{\sqrt{2}(2-\epsilon)}, \frac{2-\frac{n}{2}-(n+2)\epsilon}{\frac{1}{\sqrt{2}}-(n+1)\epsilon}\right\}$, we have

$$
(\alpha - \theta(1+\alpha))\frac{\tilde{p}}{\tilde{p}-1} = (1-\epsilon-\theta(2-\epsilon))\frac{\sqrt{2}}{\sqrt{2}-1} \geqslant \left(\frac{\sqrt{2}-1+\epsilon}{\sqrt{2}}\right)\frac{\sqrt{2}}{\sqrt{2}-1} > 1
$$

and

$$
\begin{aligned}
\left(\beta - \theta(\beta + \frac{n}{\hat{p}} - \frac{n}{2} - 2\frac{\tilde{p}-1}{\tilde{p}})\right)\frac{\hat{p}}{\hat{p}-1} &= \left(2 - \frac{n}{2} - \epsilon - \theta(\frac{1}{\sqrt{2}} - (n+1)\epsilon)\right)\frac{1}{\epsilon} \\
&= n+1 > n.
\end{aligned}
$$

Putting this together, we get from (2.21) that for all $\epsilon > 0$, $\tilde{r} \in [2,\infty)$, $\tilde{q} \in (1,\infty)$

$$
\mathcal{S} : \mathcal{M}(0,T;\mathcal{M}(\Omega)) \to B^{-1+3\theta_\epsilon,\tilde{r},\tilde{q}}(0,T;B^{0,2,\tilde{q}}(\Omega)) \cap B^{-1+3\theta_\epsilon,2,\tilde{q}}(0,T;B^{\theta_\epsilon,2,\tilde{q}}(\Omega))
$$

is bounded.

Similarly as for the part of μ in $W^{-1,\frac{Q}{Q-1}}(0,T;L^{\frac{P}{P-1}}(\Omega))$, Corollary 2.3 with $D = D_t^\alpha D_x^\beta$, $r = \infty$, yields

$$
\mathcal{S} : W^{-(1+\alpha),\frac{Q}{Q-1}}(0,T;W^{-\beta,\frac{P}{P-1}}(\Omega)) \to \check{\mathfrak{P}}^{\infty,p} \subseteq W^{-1,\frac{p}{p-1}}(0,T;L^2(\Omega))
$$

for any $p \in (1,2]$ and for any

$$
\alpha \in [0,1], \qquad \beta = 0,
$$

using $D\check{\mathcal{Y}}^r \subseteq DC^2(0,T;L^2(\Omega)) \subseteq W^{1,Q}(0,T;L^P(\Omega))$. Hence with $\hat{p} = \frac{P}{P-1}$, $\alpha = 1$, $\tilde{p} = \frac{Q}{Q-1}$ we get, in place of (2.21), boundedness of

$$
\begin{aligned}
\mathcal{S} : &B^{-2+3\theta,\frac{Q}{Q-1},\tilde{q}}(0,T;B^{\theta(\frac{n(P-1)}{P}-\frac{n}{2}-\frac{2}{Q}),\hat{p},\tilde{q}}(\Omega)) \\
&\to B^{-1+3\theta,\tilde{r},\tilde{q}}(0,T;B^{0,2,\tilde{q}}(\Omega)) \cap B^{-1+3\theta,2,\tilde{q}}(0,T;B^{\theta,2,\tilde{q}}(\Omega)).
\end{aligned}
$$

Thus we have with $\theta = \frac{1}{3}$, for any $\tilde{r} \in [2, \infty]$, $\tilde{q} \in (1, \infty]$, that

$$\mathcal{S} : B^{-1, \frac{Q}{Q-1}, \tilde{q}}(0, T; B^{0, \frac{P}{P-1}, \tilde{q}}(\Omega)) \to B^{0, \tilde{r}, \tilde{q}}(0, T; B^{0,2,\tilde{q}}(\Omega)) \cap B^{0,2,\tilde{q}}(0, T; B^{\frac{1}{3},2,\tilde{q}}(\Omega))$$

is bounded provided (2.15) holds. □

For the nonlinear model we will not show existence of a solution, since this would require a smallness condition on u, which – as opposed to [Clason, Kaltenbacher, and Veljovic 2009; Kaltenbacher and Lasiecka 2009] – we do not want to impose here. We rather prove a regularity result provided that a solution exists and satisfies the state constraints.

Lemma 2.5. *Fix m_y, P, Q satisfying (1.3). Then for any weak solution y to (1.1) with $-\underline{M}_y \leqslant y \leqslant \overline{M}_y$, $\|y_t\|_{L^Q(0,T;L^P(\Omega))} \leqslant m_y$, $u \in \mathcal{U}$, and y^0, y^1 satisfying (2.1), we have $y \in \mathcal{Y}$.*

Furthermore, there exists a $C > 0$ depending only on $\underline{M}_y$, $\overline{M}_y$, m_y, such that for all $u \in \mathcal{U}$, y^0, y^1 with (2.1), any such weak solution y to (1.1) satisfies

$$\|y\|_{\mathcal{Y}} \leqslant C\Big(\|y^0\|_{H^2(\Omega)} + \|y^1\|_{H^2(\Omega)} + \|u\|_{H^2(0,T;H^{-1/2}(\Gamma))}\Big). \tag{2.22}$$

Proof. Setting

$$z = (1 - ky), \quad w = -k(y_t)^2 \tag{2.23}$$

in (2.5), we get

$$\begin{aligned}
\frac{1}{2}\left[\left\|\sqrt{1-ky}\,y_t\right\|^2_{L^2(\Omega)} + c^2\|\nabla y\|^2_{L^2(\Omega)} + \frac{b}{c}\|y_t\|^2_{L^2(\hat{\Gamma})}\right]_0^t \\
+ \int_0^t \Big(b\|\nabla y_t\|^2_{L^2(\Omega)} + d\|y_t\|^2_{L^2(\Omega)} + c\|y_t\|^2_{L^2(\Gamma)}\Big)\,d\tau \\
= k\int_0^t \left\{\int_\Omega \frac{1}{2}(y_t)^3 + \int_\Gamma (c^2u + bu_t)y_t\,d\Gamma\right\} d\tau \\
\leqslant \frac{1}{2}km_y\|y_t\|^2_{L^{\frac{2Q}{Q-1}}(0,T;L^{\frac{2P}{P-1}}(\Omega))} \\
+ \frac{3C^2_{tr}}{2b}\|c^2u + bu_t\|^2_{L^2(0,T;H^{-1/2}(\Gamma))} + \int_0^t \frac{b}{2}\|y_t\|^2_{H^1(\Omega)}\,d\tau.
\end{aligned}$$

To control the first term on the right hand side by the first and the fourth term on the left hand side, we use interpolation

$$\|f\|_{L^r} \leqslant \|f\|^\theta_{L^p}\|f\|^{1-\theta}_{L^q} \quad \text{with} \quad \frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q} \tag{2.24}$$

between $L^\infty(0, T; L^2(\Omega))$ and $L^2(0, T; L^6(\Omega))$, i.e., (2.24) with $\theta = \frac{1}{Q} = 1 - \frac{3}{2P}$, cf. (1.3), and the embedding $H^1(\Omega) \to L^6(\Omega)$, which yields

$$
\sup_{t\in(0,T)} \Big(\frac{1}{2} \left\| \sqrt{1-ky}v \right\|^2_{L^2(\Omega)} + \underline{b} \int_0^t \|v\|^2_{H^1(\Omega)} \, d\tau \Big)
\geqslant \frac{(1-k\overline{M}_y)\underline{b}Q}{8C_{H^1,L^6}(Q-1)^{1-1/Q}} \|v\|^{2\theta}_{L^\infty(0,T;L^2(\Omega))} \|v\|^{2(1-\theta)}_{L^2(0,T;L^6(\Omega))}
\geqslant \frac{(1-k\overline{M}_y)\underline{b}Q}{8C_{H^1,L^6}(Q-1)^{1-1/Q}} \|v\|^2_{L^{\frac{2Q}{Q-1}}(0,T;L^{\frac{2P}{P-1}}(\Omega))} \tag{2.25}
$$

for $P, Q \in [1,\infty]$ such that (1.3) holds, and apply this to $v = y_t$.

Similarly, inserting (2.23) in (2.7) yields

$$
\frac{1}{2} \Big[\left\| \sqrt{1-ky}y_{tt} \right\|^2_{L^2(\Omega)} + c^2 \|\nabla y_t\|^2_{L^2(\Omega)} + \frac{b}{c} \|y_{tt}\|^2_{L^2(\hat\Gamma)} \Big]_0^t
+ \int_0^t \Big(b \|\nabla y_{tt}\|^2_{L^2(\Omega)} + d \|y_{tt}\|^2_{L^2(\Omega)} + c \|y_{tt}\|^2_{L^2(\Gamma)} \, d\tau
= k \int_0^t \Big\{ \int_\Omega \frac{5}{2} y_t (y_{tt})^2 \, dx + \int_\Gamma (c^2 u_t + b u_{tt}) y_{tt} \, d\Gamma \Big\} \, d\tau
\leqslant \frac{5}{2} k m_y \|y_{tt}\|^2_{L^{\frac{2Q}{Q-1}}(0,T;L^{\frac{2P}{P-1}}(\Omega))}
+ \frac{3C^2_{tr}}{2b} \|c^2 u_t + b u_{tt}\|^2_{L^2(0,T;H^{-1/2}(\Gamma))} + \int_0^t \frac{b}{2} \|y_{tt}\|^2_{H^1(\Omega)} \, d\tau,
$$

where we estimate the first term on the right hand side by means of (2.25) with $v = y_{tt}$. This yields

$$
\|y\|_{C^1(0,T;H^1(\Omega))} + \|y\|_{C^2(0,T;L^2(\Omega))} + \|y\|_{H^2(0,T;H^1(\Omega))} \leqslant C \|u\|_{H^2(0,T;H^{-1/2}(\Gamma))}.
$$

Finally, with (2.23) in (2.9) we get

$$
c^2 \int_0^t \|\Delta y\|_{L^2(\Omega)} \, d\tau + \frac{1}{2} \Big[b \|\Delta y\|_{L^2(\Omega)} + d \|\nabla y\|^2_{L^2(\Omega)} \Big]_0^t + \frac{d}{c} \int_0^t \|y_t\|^2_{L^2(\hat\Gamma)} \, d\tau
= \int_0^t \Big\{ \int_\Omega \Big((1-ky) y_{tt} \Delta y - k (y_t)^2 \Delta y \, dx \Big) + d \int_\Gamma u y_t \, d\Gamma \Big\} \, d\tau,
$$

hence

$$
\frac{1}{2} \Big[b \|\Delta y\|_{L^2(\Omega)} + d \|\nabla y\|^2_{L^2(\Omega)} \Big]_0^t + \frac{d}{c} \int_0^t \|y_t\|^2_{L^2(\hat\Gamma)} \, d\tau
\leqslant \frac{1}{2c^2} \Big((1 + k\underline{M}_y) \|y_{tt}\|^2_{L^2(0,T;L^2(\Omega))} + k^2 \|y_t\|^2_{L^2(0,T;L^4(\Omega))} \Big)
+ \frac{C^2_{tr}}{2} \int_0^t \|y_t(\tau)\|^2_{H^1(\Omega)} + \frac{d^2}{2} \|u\|^2_{L^2(0,T;H^{-1/2}(\Gamma))}.
$$

Resolving the PDE with respect to the strong damping term

$$b\Delta y_t = (1-ky)y_{tt} - c^2\Delta y + dy_t - 2k(y_t)^2 \in C(0,T;L^2(\Omega))$$

this time even provides $C^1(0,T;H^2(\Omega))$ regularity. □

Remark 2.6. In the following we will work with homogeneous initial conditions $y^0 = 0$, $y^1 = 0$, $u(0) = 0$, and hence linear control and state spaces $\mathcal{U}$, $\mathcal{Y}$. Note that this choice automatically satisfies the compatibility conditions (2.1).

Non-vanishing (and potentially even large) initial data can be tackled by considering the extension $y^{ext} = y^0 + ty^1$ of the initial data and the PDE

$$(1-k(\tilde{y}+y^{ext}))\tilde{y}_{tt} - c^2\Delta\tilde{y} - b\Delta\tilde{y}_t + d\tilde{y}_t - k(\tilde{y}_t + y_t^{ext})^2 = f$$

for $\tilde{y} = y - y^{ext}$, with homogeneous initial data and inhomogeneous right hand side $f(t,x) = c^2(\Delta y^0(x) + t\Delta y^1(x)) + b\Delta y^1(x) - dy^1(x)$. Lemma 2.5 easily extends to this PDE with an additional term $C\Big(\|f\|_{L^r(0,T;L^2(\Omega))} + \|f\|_{H^1(0,T;(H^1(\Omega))^*)}\Big)$ (cf. (2.4)) on the right hand side of the estimate (2.22). Since no smallness on f has to be imposed, this allows for large initial data up to the L^∞ constraint $-\underline{M}_y \leqslant y^0(x) \leqslant \overline{M}_y$.

3 STATE CONSTRAINED OPTIMAL CONTROL

We turn to the state constrained optimal control problem

$$(\mathcal{P}_{sc}) \qquad \begin{cases} \min\limits_{y\in\mathcal{Y},u\in\mathcal{U}} J(u,y) \\ (1-ky)y_{tt} - c^2\Delta y - b\Delta y_t + dy_t - k(y_t)^2 = 0 \text{ in } Q, \\ y_t + c\partial_\nu y = 0 \text{ on } (0,T)\times\Gamma \\ \partial_\nu y = u \text{ on } (0,T)\times\hat{\Gamma}, \quad y = y_t = 0 \text{ in } \{0\}\times\Omega \\ -\underline{M}_y \leqslant y(t,x) \leqslant \overline{M}_y \quad \text{for all } t,x \in Q, \quad \|y_t\|_{L^Q(0,T;L^P(\Omega))} \leqslant m_y \end{cases}$$

for m_y, P, Q satisfying (1.3), and define

$$G : \mathcal{U} \times \mathcal{Y}_M \to L^2(0,T;(H^1(\Omega))^*)$$

by the weak form of the PDE with boundary and initial conditions on y in ($\mathcal{P}_{sc}$). Here $\mathcal{U}$ and $\mathcal{Y}$ are defined as in (1.4) with $y^0 = y^1 = 0$ and are thus linear spaces. The case of inhomogeneous initial data can be treated as described in Remark 2.6.

We then have the following existence result.

Theorem 3.1. *There exists a minimizer* $(u^*, y^*) \in \mathcal{U} \times \mathcal{Y}_M$ *of* ($\mathcal{P}_{sc}$).

Proof. By non-emptiness of the feasible set (u, y) (take $(u, y) = (0, 0)$), boundedness of J from below, and the coercivity of the functional in u, we obtain the existence of a minimizing sequence whose control part is bounded in $\mathcal{U}$. The equality and state constraints together with Lemma 2.5 imply that the y components of the minimizing sequence are uniformly bounded in $\mathcal{Y}$. Hence, there exists a subsequence, denoted by $\{(u_n, y_n)\}_{n\in\mathbb{N}}$, that weakly converges in $\mathcal{U} \times \hat{\mathcal{Y}}^s$ to $(u^*, y^*) \in \mathcal{U} \times \hat{\mathcal{Y}}^s$ for $s \in (1, \infty)$. Due to the compact embedding of $\mathcal{Y}$ in $\mathcal{Z}$ and the weak continuity of the mapping $y \mapsto k(y_t)^2$ from $\hat{\mathcal{Y}}^s$ to $\mathcal{V}$, we have that along some subsequence $(y_{n_k})_{k\in\mathbb{N}}$,

$$
\begin{aligned}
1 - ky_{n_k} &\to 1 - ky^* && \text{in } \mathcal{Z} \\
k(y_{n_k t})^2 &\rightharpoonup k(y_t^*)^2 && \text{in } \mathcal{V},
\end{aligned}
$$

and y^* satisfies the inequality constraints. Thus, we can pass to the limit in (the weak formulation of) $G(u_{n_k}, y_{n_k}) = 0$ to obtain $G(u^*, y^*) = 0$ and therefore by Lemma 2.5, $y^* \in \mathcal{Y}$. □

We mention in passing that the result extends to cost functions $J(u, y)$ that are bounded from below, weakly lower semi-continuous in $\mathcal{U} \times \mathcal{Y}$, and $\mathcal{U}$-coercive with respect to u.

Due to the fact that we deal with a nonlinear PDE and the control only acts on the boundary, a regular point condition according to, e.g., [Alibert and Raymond 1998], would require existence of $(u_0, y_0) \in \mathcal{U} \times \mathcal{Y}_M$ such that $(y^* + y_0) \in \operatorname{int} \mathcal{Y}_M$ and

$$
G_y(u^*, y^*)y_0 + G_u(u^*, y^*)(u_0 - u^*) = 0,
$$

where G_y denotes the Fréchet derivative of G with respect to y. This seems hard or even impossible to satisfy due to the relative low dimensionality of the control as compared to the state. Therefore, as in [Clason and Kaltenbacher 2012], we introduce a relaxation according to [Bonnans and Casas 1989], combined with a localization technique as in [Casas and Tröltzsch 2002]. Specifically, we introduce new independent variables w and z in place of the nonlinearities $k(y_t)^2$ and $(1 - ky)$, respectively, and penalize the deviation from the original minimizers. Taking the limit with respect to the penalty parameter in the corresponding optimality conditions yields the optimality system for the original problem. We thus consider

$$
(\mathcal{P}_{sc,\varepsilon}) \quad
\begin{cases}
\displaystyle\min_{u\in\mathcal{U}, w\in\mathcal{W}, (z,y)\in\mathcal{Z}\mathcal{Y}} J_\varepsilon(u, w, z, y) \\
\qquad \text{s. t.} \quad zy_{tt} - c^2\Delta y - b\Delta y_t + dy_t + w = 0 \text{ in } Q, \\
\qquad\qquad \partial_\nu y = u \text{ on } (0, T) \times \Gamma, \\
\qquad\qquad y_t + c\partial_\nu y = 0 \text{ on } (0, T) \times \hat{\Gamma}, \quad y = 0,\ y_t = 0 \text{ in } \{0\} \times \Omega, \\
\qquad\qquad 1 + k\underline{M}_y \geqslant z \geqslant 1 - k\overline{M}_y > 0 \quad \text{for all } t, x \in Q, \\
\qquad\qquad \|y_t\|_{L^Q(0,T;L^P(\Omega))} \leqslant m_y
\end{cases}
$$

where

$$
\mathcal{Z}\mathcal{Y} = \{(z, y) \in \mathcal{Z}_M \times \mathcal{Y} \;:\; zy_{tt} \in \mathcal{V}\},
$$

and $J_\varepsilon : \mathcal{U} \times \mathcal{W} \times \mathcal{ZY} \to \mathbb{R}$ is given by

$$\begin{aligned}
J_\varepsilon(u, w, z, y) &= J(u, y) + \frac{1}{2\varepsilon} \left\|w + k(y_t)^2\right\|^2_{\hat{\mathcal{W}}} + \frac{1}{2\varepsilon} \|z + ky - 1\|^2_{\hat{\mathcal{Z}}} \\
&\quad + \frac{1}{2\delta} \|u - u^*\|^2_{\hat{\mathcal{U}}} + \frac{1}{2\delta} \left\|w + k(y_t^*)^2\right\|^2_{\hat{\mathcal{W}}} + \frac{1}{2\delta} \|z + ky^* - 1\|^2_{\hat{\mathcal{Z}}} \\
&= \frac{1}{2} \int_0^T \Bigg\{ \int_\Omega \Bigg((y - y_d)^2 \\
&\quad + \frac{1}{\varepsilon} \sum_{i=1}^{2} (\mathcal{A}_i(w + k(y_t)^2))^2 + \frac{1}{\varepsilon} \sum_{i=1}^{3} (\mathcal{B}_i(z + ky - 1))^2 \\
&\quad + \frac{1}{\delta} \sum_{i=1}^{2} (\mathcal{A}_i(w + k(y_t^*)^2))^2 + \frac{1}{\delta} \sum_{i=1}^{3} (\mathcal{B}_i(z + ky^* - 1))^2 \Bigg) \, dx \\
&\quad + \int_\Gamma \left(\alpha u^2 + \frac{1}{\delta} \sum_{i=1}^{3} (\mathcal{C}_i(u - u^*))^2 \right) d\Gamma \Bigg\} \, dt,
\end{aligned}$$

with

$$\delta < \frac{\rho}{2 \max\{1, C\} J(u^*, y^*)},$$

and C as in Lemma 2.5, where

- $\hat{\mathcal{W}} = H^1(0, T; L^2(\Omega))$ with

$$\|w\|^2_{\hat{\mathcal{W}}} = \int_0^T \int_\Omega \left((w_t)^2 + w^2 \right) dx\, dt =: \sum_{i=1}^{2} \int_0^T \int_\Omega (\mathcal{A}_i w)^2 dx\, dt,$$

- $\hat{\mathcal{Z}} = H^1(0, T; H^2(\Omega)) \cap H^2(0, T; L^2(\Omega))$ with

$$\|z\|^2_{\hat{\mathcal{Z}}} = \int_0^T \int_\Omega \left((z_{tt})^2 + (\Delta z)^2 + z^2 \right) dx\, dt =: \sum_{i=1}^{3} \int_0^T \int_\Omega (\mathcal{B}_i z)^2 dx\, dt,$$

- $\hat{\mathcal{U}} = H^2(0; T; L^2(\Gamma))$ with

$$\|u\|_{\hat{\mathcal{U}}} = \int_0^T \int_\Gamma \left((u_{tt})^2 + (u_t)^2 + u^2 \right) dx\, dt =: \sum_{i=1}^{3} \int_0^T \int_\Omega (\mathcal{C}_i u)^2 dx\, dt,$$

so that boundedness of w, z, u in the spaces $\hat{\mathcal{W}}, \hat{\mathcal{Z}}, \hat{\mathcal{U}}$ respectively, will imply their boundedness in the spaces as needed according to Lemma 2.1.

Existence of a global minimizer to $(\mathcal{P}_{sc,\varepsilon})$ follows analogously to the proof of Theorem 3.1.

Lemma 3.2. *There exists a minimizer* $(u_\varepsilon^*, w_\varepsilon^*, (z_\varepsilon^*, y_\varepsilon^*)) \in \mathcal{U} \times \mathcal{W} \times \mathcal{ZY}$ *of* $(\mathcal{P}_{sc,\varepsilon})$.

We now derive optimality conditions for ($\mathcal{P}_{sc,\varepsilon}$), which will yield optimality conditions for ($\mathcal{P}_{sc}$) by taking the limit $\varepsilon \to 0$. To this end, we define

$$G_{\text{bil}} : \mathcal{U} \times \mathcal{W} \times \mathcal{Z}\mathcal{Y} \to \mathcal{V}$$

by the weak form (2.3) of the PDE with boundary and initial conditions on y in ($\mathcal{P}_{sc,\varepsilon}$). (The mapping properties of G_{bil} can be verified by inspection of (2.3).)

Lemma 3.3. *Let* $(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*) \in \mathcal{U} \times \mathcal{W} \times \mathcal{Z}\mathcal{Y}$ *be a local minimizer of* ($\mathcal{P}_{sc,\varepsilon}$). *Then there exist* $\underline{\mu}_\varepsilon^*, \overline{\mu}_\varepsilon^* \in \mathcal{M}(0,T;\mathcal{M}(\Omega))$ *satisfying*

$$\left\langle \underline{\mu}_\varepsilon^*, \varphi \right\rangle_{\mathcal{M},C} \geqslant 0, \quad \langle \overline{\mu}_\varepsilon^*, \varphi \rangle_{\mathcal{M},C} \geqslant 0 \quad \text{for all } \varphi \in C_0(\Omega) \text{ with } \varphi \geqslant 0, \tag{3.1}$$

$\lambda_\varepsilon^* \geqslant 0$ *and* $p_\varepsilon^* \in \mathcal{V}^*$ *such that*

$$\begin{aligned}
\int_0^T \Big\{ \int_\Omega \Big(z_\varepsilon^* v_{tt} p_\varepsilon^* + c^2 \nabla v \nabla p_\varepsilon^* + b \nabla v_t \nabla p_\varepsilon^* + d v_t p_\varepsilon^* \Big)\, dx + \int_{\hat\Gamma} (c v_t + \frac{b}{c} v_{tt}) p_\varepsilon^* \, d\Gamma \Big\}\, dt \\
= - \int_0^T \int_\Omega \Big((y_\varepsilon^* - y_d) v + \frac{1}{\varepsilon} \sum_{i=1}^{2} (\mathcal{A}_i (w_\varepsilon^* + k (y_{\varepsilon,t}^*)^2)) (\mathcal{A}_i (2k y_{\varepsilon,t}^* v_t)) \\
+ \frac{1}{\varepsilon} \sum_{i=1}^{3} (\mathcal{B}_i (z_\varepsilon^* + k y_\varepsilon^* - 1)) (\mathcal{B}_i (k v)) \Big)\, dx\, dt \\
- \lambda_\varepsilon^* Q \int_0^T \|y_{\varepsilon,t}^*(t)\|_{L^P(\Omega)}^{Q-P} \int_\Omega |y_{\varepsilon,t}^*|^{P-1} \operatorname{sign}(y_{\varepsilon,t}^*) v_t \, dx\, dt
\end{aligned}$$

for all $v \in \mathcal{Y}$,

$$\begin{aligned}
\int_0^T \Big\{ \int_\Omega \Big(\frac{1}{\varepsilon} \sum_{i=1}^{3} (\mathcal{B}_i (z_\varepsilon^* + k y_\varepsilon^* - 1)) (\mathcal{B}_i v) + \frac{1}{\delta} \sum_{i=1}^{3} (\mathcal{B}_i (z_\varepsilon^* + k y^* - 1)) (\mathcal{B}_i v) \Big)\, dx \\
+ v y_{\varepsilon,tt}^* p_\varepsilon^* \Big)\, dx\, dt = - \left\langle \underline{\mu}_\varepsilon^*, v \right\rangle_{\mathcal{M},C} + \langle \overline{\mu}_\varepsilon^*, v \rangle_{\mathcal{M},C}
\end{aligned}$$

for all $v \in \mathcal{Z}$,

$$\begin{aligned}
\int_0^T \Big\{ \int_\Omega \Big(\frac{1}{\varepsilon} \sum_{i=1}^{2} (\mathcal{A}_i (w_\varepsilon^* + k (y_{\varepsilon,t}^*)^2)) (\mathcal{A}_i v) + \frac{1}{\delta} \sum_{i=1}^{2} (\mathcal{A}_i (w_\varepsilon^* + k (y_t^*)^2)) (\mathcal{A}_i v) \\
+ p_\varepsilon^* v \Big)\, dx\, dt = 0
\end{aligned}$$

for all $v \in \mathcal{W}$,

$$\int_0^T \int_\Gamma \Big(\alpha u_\varepsilon^* v + \frac{1}{\delta} \sum_{i=1}^{3} (\mathcal{C}_i (u_\varepsilon^* - u^*)) (\mathcal{C}_i v) - (c^2 v + b v_t) p_\varepsilon^* \Big)\, d\Gamma\, dt = 0$$

for all $v \in \mathcal{U}$,

$$\begin{aligned}
\left\langle \underline{\mu}_\varepsilon^*, z_\varepsilon^* - k \underline{M}_Y - 1 \right\rangle_{\mathcal{M},C} = 0, \quad \langle \overline{\mu}_\varepsilon^*, z_\varepsilon^* + k \overline{M}_Y - 1 \rangle_{\mathcal{M},C} = 0, \\
\lambda_\varepsilon^* (\|y_{\varepsilon,t}^*\|_{L^Q(0,T;L^P(\Omega))}^Q - m_y^Q) = 0.
\end{aligned}$$

Proof. For the existence of Lagrange multipliers, we appeal to the regular point condition from [Alibert and Raymond 1998]: For any $\mathbf{x} = (u_\varepsilon^*, w_\varepsilon^*, (z_\varepsilon^*, y_\varepsilon^*))$, there exist $(u_0, w_0, (z_0, y_0)) \in \mathcal{U} \times \mathcal{W} \times \mathcal{ZY}$ such that

$$G_{\mathrm{bil},y}(\mathbf{x})y_0 + G_{\mathrm{bil},z}(\mathbf{x})z_0 + G_{\mathrm{bil},w}(\mathbf{x})(w_0 - w_\varepsilon^*) + G_{\mathrm{bil},u}(\mathbf{x})(u_0 - u_\varepsilon^*) = 0$$

holds, and $(z_\varepsilon^* + z_0, y_\varepsilon^* + y_0)$ satisfies the inequality constraints in ($\mathcal{P}_{sc,\varepsilon}$) with strict inequality. Although the mapping G_{bil} is not linear but bilinear, this can be satisfied by setting

$$z_0 = 1 - z_\varepsilon^*, \quad y_0 = -y_\varepsilon^*, \quad u_0 = 0, \quad w_0 = -(1 - z_\varepsilon^*)y_{\varepsilon\,tt}^*.$$

Indeed, obviously

$$w_0 = z_0 y_{0,tt} = z_\varepsilon^* y_{\varepsilon\,tt}^* - y_{\varepsilon\,tt}^* \in \mathcal{V} + L^2(0,T;H^1(\Omega)) \subseteq L^2(0,T;L^2(\Omega)).$$

Moreover, differentiating the PDE once with respect to time and using $y_\varepsilon^* \in \mathcal{Y}$ and $z_\varepsilon^* \in \mathcal{Z}$, we see that

$$\begin{aligned} w_{0t} &= (z_0 y_{0,tt})_t \\ &= z_{\varepsilon\,t}^* y_{\varepsilon\,tt}^* - \frac{(1-z_\varepsilon^*)}{z_\varepsilon^*}\Big(-z_{\varepsilon\,t}^* y_{\varepsilon\,tt}^* + c^2 \Delta y_{\varepsilon\,t}^* + b\Delta y_{\varepsilon\,tt}^* - d y_{\varepsilon\,tt}^* - w_{\varepsilon\,t}^*\Big) \\ &= \frac{1}{z_\varepsilon^*} \underbrace{z_{\varepsilon\,t}^*}_{\in C(0,T;L^2(\Omega))} \underbrace{y_{\varepsilon\,tt}^*}_{\in C(0,T;L^2(\Omega))} - \frac{(1-z_\varepsilon^*)}{z_\varepsilon^*}\Big(c^2 \underbrace{\Delta y_{\varepsilon\,t}^*}_{\in C(0,T;(H^1(\Omega))^*)} + b \underbrace{\Delta y_{\varepsilon\,tt}^*}_{\in L^2(0,T;(H^1(\Omega))^*)} \\ &\qquad - d \underbrace{y_{\varepsilon\,tt}^*}_{\in C(0,T;L^2(\Omega))} - \underbrace{w_{\varepsilon\,t}^*}_{\in L^2(0,T;(H^1(\Omega))^*)}\Big) \\ &\in L^2(0,T;(H^1(\Omega))^*). \end{aligned}$$

Together, we have $(u_0, w_0, (z_0, y_0)) \in \mathcal{U} \times \mathcal{W} \times \mathcal{ZY}$.

Furthermore, surjectivity of $G_{\mathrm{bil},(z,y)}(u_\varepsilon^*, w_\varepsilon^*, (z_\varepsilon^*, y_\varepsilon^*)) : \mathcal{ZY} \to \mathcal{V}$, i.e., existence of a solution $(z, y) \in \mathcal{ZY}$ to the weak form of

$$\begin{cases} z y_{\varepsilon,tt}^* + z_\varepsilon^* y_{tt} - c^2 \Delta y - b\Delta y_t + d y_t = f \text{ in } Q, \\ \partial_\nu y = 0 \text{ on } (0,T) \times \Gamma, \\ y_t + c\partial_\nu y = 0 \text{ on } (0,T) \times \hat{\Gamma}, \quad y = 0,\ y_t = 0 \text{ in } \{0\} \times \Omega \end{cases} \tag{3.2}$$

for any $f \in \mathcal{V}$ can be deduced from Lemma 2.1 (we simply set $z = 0$ in (3.2)).

Existence of Lagrange multipliers

$$p_\varepsilon \in \mathcal{V}^* = L^2(0,T;L^2(\Omega)) + H^{-1}(0,T;H^1(\Omega))$$

and $\underline{\mu}_\varepsilon, \overline{\mu}_\varepsilon \in C(0,T;C(\Omega))^*$ now follows from [Alibert and Raymond 1998, Theorem 2.1].

Finally, the explicit form of the optimality conditions can be obtained by formal differentiation of the Lagrangian

$$\begin{aligned}
\mathcal{L}(u, w, z, y, p, \underline{\mu}, \overline{\mu}, \lambda) = \frac{1}{2}\int_0^T \Big\{ \int_\Omega \Big(& (y - y_d)^2 \\
& + \frac{1}{\varepsilon}\sum_{i=1}^{2}(\mathcal{A}_i(w + k(y_t)^2))^2 + \frac{1}{\varepsilon}\sum_{i=1}^{3}(\mathcal{B}_i(z + ky - 1))^2 \\
& + \frac{1}{\delta}\sum_{i=1}^{2}(\mathcal{A}_i(w + k(y_t^*)^2))^2 + \frac{1}{\delta}\sum_{i=1}^{3}(\mathcal{B}_i(z + ky^* - 1))^2\Big)\, dx \\
& + \int_\Gamma \Big(\alpha u^2 + \frac{1}{\delta}\sum_{i=1}^{3}(\mathcal{C}_i(u - u^*))^2 \Big)\, d\Gamma \Big\}\, dt \\
& - \int_0^T \Big\{ \int_\Omega \Big(z y_{tt} p + c^2 \nabla y \nabla p + b \nabla y_t \nabla p + d y_t p + w p \Big)\, dx \\
& - \int_{\hat{\Gamma}} (c y_t + \frac{b}{c} y_{tt}) p\, d\Gamma + \int_\Gamma (c^2 u + b u_t) p\, d\Gamma \Big\}\, dt \\
& + \langle \underline{\mu}, z - k\underline{M}_y - 1 \rangle_{\mathcal{M},C} - \langle \overline{\mu}, z + k\overline{M}_y - 1 \rangle_{\mathcal{M},C} \\
& + \lambda(\|y_t\|^Q_{L^Q(0,T;L^P(\Omega))} - m_y^Q)
\end{aligned}$$

Note that the variational inequalities

$$\Big\langle \underline{\mu}_\varepsilon^*, z_\varepsilon^* + ky - 1 \Big\rangle_{\mathcal{M},C} \geqslant 0, \qquad \langle \overline{\mu}_\varepsilon^*, z_\varepsilon^* + ky - 1 \rangle_{\mathcal{M},C} \geqslant 0$$

for all $y \in \mathcal{Y}_M$, which can be concluded from [Alibert and Raymond 1998, Theorem 2.1], imply non-negativity (3.1), as well as the complementarity conditions for $\overline{\mu}_\varepsilon$ and $\underline{\mu}_\varepsilon$. □

Similarly to [Bonnans and Casas 1989, Lemma 3] we obtain

Lemma 3.4. *Let (u^*, y^*) be a local minimizer of $(\mathcal{P}_{sc})$, and let $\{(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*)\}_{\varepsilon>0}$ be a family of global minimizers of $(\mathcal{P}_{sc,\varepsilon})$. Then we have convergence*

$$\begin{aligned}
& u_\varepsilon^* \to u^* \text{ in } \hat{\mathcal{U}}, \quad w_\varepsilon^* \to k(y_t^*)^2 \text{ in } \hat{\mathcal{W}}, \quad z_\varepsilon^* \to 1 - ky^* \text{ in } \hat{\mathcal{Z}}, \\
& y_\varepsilon^* \rightharpoonup y^* \text{ in } \mathcal{Y}, \quad y_\varepsilon^* \to y^* \text{ in } \mathcal{Z}
\end{aligned}$$

as $\varepsilon \to 0$.

Proof. By minimality of $(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*)$ and admissibility of $(u^*, -k(y_t^*)^2, 1 - ky^*, y^*)$ for $(\mathcal{P}_{sc,\varepsilon})$ (note that in particular $(1 - ky^*, y^*) \in \mathcal{Z}\mathcal{Y}$ since $(1 - ky^*)y_{tt}^* = c^2\Delta y^* + b\Delta y_t^* - dy_t^* + k(y_t^*)^2 \in \mathcal{V}$) we have

$$J_\varepsilon(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*) \leqslant J(u^*, -k(y_t^*)^2, 1 - ky^*, y^*) = J(u^*, y^*) \tag{3.3}$$

hence, $\{(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*)\}_{\varepsilon>0}$ is bounded in $\hat{\mathcal{U}} \times \hat{\mathcal{W}} \times \hat{\mathcal{Z}}$ and by $G_{bil}(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*) = 0$ and Lemma 2.1, $\{(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*)\}_{\varepsilon>0}$ bounded in $\hat{\mathcal{U}} \times \hat{\mathcal{W}} \times \hat{\mathcal{Z}} \times \mathcal{Y}$. By compactness of the embeddings $\hat{\mathcal{Z}} \hookrightarrow$

$C(0,T;C(\Omega)), \mathcal{Y} \hookrightarrow \mathcal{Z}$, we can therefore extract from any subsequence of $\{(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*)\}_{\varepsilon>0}$ a subsequence (for simplicity denoted by $\{(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*)\}_{\varepsilon>0}$ again) such that

$$
\begin{aligned}
&u_\varepsilon^* \rightharpoonup \hat{u} \text{ in } \hat{\mathcal{U}}, \quad w_\varepsilon^* \rightharpoonup \hat{w} \text{ in } \hat{\mathcal{W}}, \quad z_\varepsilon^* \rightharpoonup \hat{z} \text{ in } \hat{\mathcal{Z}}, \quad y_\varepsilon^* \rightharpoonup \hat{y} \text{ in } \mathcal{Y}, \\
&z_\varepsilon^* \to \hat{z} \text{ in } C(0,T;C(\Omega)), \quad y_\varepsilon^* \to \hat{y} \text{ in } \mathcal{Z}
\end{aligned} \tag{3.4}
$$

as $\varepsilon \to 0$, for some $(\hat{u}, \hat{w}, \hat{z}, \hat{y}) \in \hat{\mathcal{U}} \times \hat{\mathcal{W}} \times \hat{\mathcal{Z}} \times \mathcal{Y}$ satisfying the inequality constraints. Moreover, taking the weak limit in the PDE, we see that $G_{\text{bil}}(\hat{u}, \hat{w}, \hat{z}, \hat{y}) = 0$ (here we use $z_\varepsilon^* \to \hat{z}$ in $C(0,T;C(\Omega))$). The uniform boundedness (3.3) and the penalty terms with factor $\frac{1}{\varepsilon}$ in J_ε imply

$$
\left\|w_\varepsilon^* + k(y_{\varepsilon,t}^*)^2\right\|_{\hat{\mathcal{W}}} \to 0, \quad \|z_\varepsilon^* + k y_\varepsilon^* - 1\|_{\hat{\mathcal{Z}}} \to 0
$$

hence by (3.4) and weak lower semicontinuity of the norm we obtain

$$
\hat{w} + k(\hat{y}_t)^2 = 0, \quad \hat{z} + k\hat{y} - 1 = 0, \tag{3.5}
$$

which together with $G_{\text{bil}}(\hat{u}, \hat{w}, \hat{z}, \hat{y}) = 0$ implies that $G(\hat{u}, \hat{y}) = 0$. Thus $(\hat{u}, \hat{y})$ is admissible for $(\mathcal{P}_{sc})$ and we can use minimality of (u^*, y^*) for this problem, relation (3.3), and weak lower semicontinuity of J and the norms to conclude

(3.6)

$$
\begin{aligned}
J(\hat{u}, \hat{y}) &\geqslant J(u^*, y^*) \geqslant \limsup_{\varepsilon\to 0} J_\varepsilon(u_\varepsilon^*, w_\varepsilon^*, z_\varepsilon^*, y_\varepsilon^*) \\
&\geqslant \limsup_{\varepsilon\to 0}\Big(J(u_\varepsilon^*, y_\varepsilon^*) + \frac{1}{2}\|u_\varepsilon^* - u^*\|_{\hat{\mathcal{U}}}^2 + \frac{1}{2}\left\|w_\varepsilon^* + k(y_{\varepsilon,t}^*)^2\right\|_{\hat{\mathcal{W}}}^2 + \frac{1}{2}\|z_\varepsilon^* + k y_\varepsilon^* - 1\|_{\hat{\mathcal{Z}}}^2\Big) \\
&\geqslant \liminf_{\varepsilon\to 0} J(u_\varepsilon^*, y_\varepsilon^*) \\
&\quad + \liminf_{\varepsilon\to 0}\Big(\frac{1}{2}\|u_\varepsilon^* - u^*\|_{\hat{\mathcal{U}}}^2 + \frac{1}{2}\left\|w_\varepsilon^* + k(y^*)^2\right\|_{\hat{\mathcal{W}}}^2 + \frac{1}{2}\|z_\varepsilon^* + k y^* - 1\|_{\hat{\mathcal{Z}}}^2\Big) \\
&\geqslant J(\hat{u}, \hat{y}) + \frac{1}{2}\|\hat{u} - u^*\|_{\hat{\mathcal{U}}}^2 + \frac{1}{2}\left\|\hat{w} + k(y_t^*)^2\right\|_{\hat{\mathcal{W}}}^2 + \frac{1}{2}\|\hat{z} + k y^* - 1\|_{\hat{\mathcal{Z}}}^2
\end{aligned}
$$

which implies $\hat{u} = u^*$, $\hat{w} = -k(y_t^*)^2$, $\hat{z} = 1 - k y^*$, and by (3.5) therefore $k(\hat{y}_t)^2 = k(y_t^*)^2$, $k\hat{y} - 1 = k y^* - 1$ and hence $\hat{y} = y^*$. Strong convergence of the u, w, and z components along a subsequence immediately follows from (3.6). Using a subsequence-subsequence argument, we arrive at the assertion. □

Taking the limit as $\varepsilon \to 0$ in Lemma 3.3, we arrive at the following necessary optimality conditions for solutions to $(\mathcal{P}_{sc})$.

Theorem 3.5. *Let $(u^*, y^*) \in \mathcal{U} \times \mathcal{Y}_M$ be a local minimizer of $(\mathcal{P}_{sc})$ with $P \in [1,2]$, $Q \in [1,\infty]$ such that (1.3), (2.15) hold (e.g., $P = 2$, $Q = 4$). Then there exist $\underline{\mu}^*, \overline{\mu}^* \in \mathcal{M}(0,T;\mathcal{M}(\Omega))$ satisfying*

$$
\langle \underline{\mu}^*, \varphi \rangle_{\mathcal{M},C} \geqslant 0, \quad \langle \overline{\mu}^*, \varphi \rangle_{\mathcal{M},C} \geqslant 0 \quad \textit{for all } \varphi \in C_0(\Omega) \textit{ with } \varphi \geqslant 0,
$$

$\lambda^* \geqslant 0$, $p^* \in \mathcal{P}^{\sigma,heta}$ *as in* (1.4), *with* $0 < \sigma < \frac{3}{\sqrt{8}} - 1$, $0 < \theta < \frac{1}{\sqrt{8}}$ *such that*

$$\begin{aligned}\int_0^T \Big\{ \int_\Omega \Big(((1-ky^*)v)_{tt} p^* + c^2 \nabla v \nabla p^* + b \nabla v_t \nabla p^* + d v_t p^* \Big)\, dx \\ + \int_{\hat{\Gamma}} (c v_t + \frac{b}{c} v_{tt}) p^*\, d\Gamma \Big\}\, dt \\ = -\gamma \int_0^T \int_\Omega (y^* - y_d) v \, dx\, dt - \langle \underline{\mu}^* - \overline{\mu}^*, kv \rangle_{\mathcal{M},C} \\ - \lambda^* Q \int_0^T \|y_t^*(t)\|_{L^P(\Omega)}^{Q-P} \int_\Omega |y_t^*|^{P-1} \operatorname{sign}(y_t^*) v_t \, dx\, dt\end{aligned}$$

for all $v \in \mathcal{Y}$,

$$\int_0^T \int_\Gamma \Big(\gamma \alpha u^* v - (c^2 v + b v_t) p^* \Big)\, d\Gamma\, dt = 0$$

for all $v \in \mathcal{U}$,

$$\begin{aligned}&\langle \underline{\mu}^*, y^* + \underline{M}_Y \rangle_{\mathcal{M},C} = 0, \quad \langle \overline{\mu}^*, y^* - \overline{M} \rangle_{\mathcal{M},C} = 0, \\ &\lambda^* (\|y_t^*\|_{L^Q(0,T;L^P(\Omega))}^Q - m_y^Q) = 0\end{aligned}$$

where $\gamma \in \{0,1\}$, $\gamma + \|p^*\|_{\mathcal{P}} + \lambda^* > 0$.

Proof. Eliminating the $\frac{1}{\varepsilon}$ terms in Lemma 3.3 (using the fact that for any $v \in \mathcal{Y}$ we have $2ky_{\varepsilon,t}^* v_t \in \mathcal{W}$ and $kv \in \mathcal{Z}$), yields

$$\begin{aligned}&\int_0^T \Big\{ \int_\Omega \Big(z_\varepsilon^* v_{tt} p_\varepsilon^* + c^2 \nabla v \nabla p_\varepsilon^* + b \nabla v_t \nabla p_\varepsilon^* + d v_t p_\varepsilon^* \Big)\, dx + \int_{\hat{\Gamma}} (c v_t + \frac{b}{c} v_{tt}) p_\varepsilon^*\, d\Gamma \Big\}\, dt \\ &\quad = - \int_0^T \int_\Omega \Big((y_\varepsilon^* - y_d) v - \frac{1}{\delta} \sum_{i=1}^2 (\mathcal{A}_i (w_\varepsilon^* + k(y_t^*)^2))(\mathcal{A}_i (2ky_{\varepsilon,t}^* v_t)) - p_\varepsilon^* 2ky_{\varepsilon,t}^* v_t \\ &\quad\quad - \frac{1}{\delta} \sum_{i=1}^3 (\mathcal{B}_i (z_\varepsilon^* + ky^* - 1))(\mathcal{B}_i kv) - kv y_{\varepsilon,tt}^* p_\varepsilon^* \Big)\, dx\, dt - \Big\langle \underline{\mu}_\varepsilon^* - \overline{\mu}_\varepsilon^*, kv \Big\rangle_{\mathcal{M},C} \\ &\quad\quad - \lambda_\varepsilon^* Q \int_0^T \|y_{\varepsilon,t}^*(t)\|_{L^P(\Omega)}^{Q-P} \int_\Omega |y_{\varepsilon,t}^*|^{P-1} \operatorname{sign}(y_{\varepsilon,t}^*) v_t \, dx\, dt\end{aligned}$$

for all $v \in \mathcal{Y}$,

$$\int_0^T \int_\Gamma \Big(\alpha u_\varepsilon^* v + \frac{1}{\delta} \sum_{i=1}^3 (\mathcal{C}_i (u_\varepsilon^* - u^*))(\mathcal{C}_i v) - (c^2 v + b v_t) p_\varepsilon^* \Big)\, d\Gamma\, dt = 0$$

for all $v \in \mathcal{U}$,

$$\Big\langle \underline{\mu}_\varepsilon^*, z_\varepsilon^* - k\underline{M}_Y - 1 \Big\rangle_{\mathcal{M},C} = 0, \quad \langle \overline{\mu}_\varepsilon^*, z_\varepsilon^* + k\overline{M}_Y - 1 \rangle_{\mathcal{M},C} = 0,$$

$\lambda_\varepsilon^*(\|y_{\varepsilon,t}^*\|^Q_{L^Q(0,T;L^P(\Omega))} - m_y^Q) = 0.$

The derivation of the optimality conditions now goes analogously to the proof of Theorem 2 in [Bonnans and Casas 1989]. Indeed, for p_ε^* (or $\frac{p_\varepsilon^*}{r_\varepsilon}$ with $r_\varepsilon = \|p_\varepsilon^*\|_{\mathcal{V}} + \|\underline{\mu}_\varepsilon^* + \overline{\mu}_\varepsilon^*\|_{\mathcal{M}(0,T;\mathcal{M}(\Omega))} + \lambda_\varepsilon^*$) weakly converging to p^* in $\mathcal{V}$, we get for $v \in \mathcal{Y}$:

$$\int_0^T \int_\Omega z_\varepsilon^* v_{tt} p_\varepsilon^* \,dx\,dt \to \int_0^T \int_\Omega (1 - ky^*) v_{tt} p^* \,dx\,dt$$

by strong convergence of z_ε^* to $1-ky^*$ in $C(0,T;C(\Omega)) \cap C(0,T;H^{1+\frac{n}{4}}(\Omega))$, $v_{tt} \in L^2(0,T;H^1(\Omega))$, (hence $z_\varepsilon^* v_{tt} \to (1-ky^*)v_{tt}$ in $L^2(0,T;H^1(\Omega))$), and weak convergence of p_ε^* to p^* in $L^2(0,T;(H^1(\Omega))^*)$;

$$\int_0^T \int_\Omega p_\varepsilon^* 2ky_{\varepsilon,t}^* v_t \,dx\,dt \to \int_0^T \int_\Omega p^* 2ky_t^* v_t \,dx\,dt$$

by strong convergence of $y_{\varepsilon,t}^*$ to y_t^* in $L^2(0,T;C(\Omega)) \cap L^2(0,T;W^{1,3}(\Omega))$ (using boundedness of $y_{\varepsilon,t}^*$ in $H^1(0,T;H^1(\Omega)) \cap L^2(0,T;H^2(\Omega)) \subseteq H^{1-\theta}(0,T;H^{1+\theta}(\Omega)$ for $\frac{n}{6} < \theta < 1$ and compact embedding), $v_t \in C(0,T;H^1(\Omega)) \cap C(0,T;L^6(\Omega))$ (hence $y_{\varepsilon,t}^* v_t \to y_t^* v_t$ in $L^2(0,T;H^1(\Omega))$), and weak convergence of p_ε^* to p^* in $L^2(0,T;(H^1(\Omega))^*)$;

$$\int_0^T \int_\Omega (\mathcal{B}_i(z_\varepsilon^* + ky^* - 1))(\mathcal{B}_i kv) \,dx\,dt \to 0$$

by strong convergence of $\mathcal{B}_i(z_\varepsilon^* + ky^* - 1)$ to zero in $L^2(0,T;L^2(\Omega))$ and $\mathcal{B}_i kv \in L^2(0,T;L^2(\Omega))$; and

$$\int_0^T \int_\Omega (\mathcal{A}_i(w_\varepsilon^* + k(y_t^*)^2))(\mathcal{A}_i(2ky_{\varepsilon,t}^* v_t)) \,dx\,dt \to 0$$

by strong convergence of $\mathcal{A}_i(w_\varepsilon^* + k(y_t^*)^2)$ to zero in $L^2(0,T;L^2(\Omega))$ and uniform boundedness of $\mathcal{A}_i(2ky_{\varepsilon,t}^* v_t)$ in $L^2(0,T;L^2(\Omega))$. For the rest of the terms convergence is straightforward.

Observing that the right hand side term

$$v \mapsto \int_0^T \|y_t^*(t)\|^{Q-P}_{L^P(\Omega)} \int_\Omega |y_t^*|^{P-1} \mathrm{sign}(y_t^*) v_t \,dx\,dt \in (W^{1,Q}(0,T;L^P(\Omega)))^*,$$

we have by Corollary 2.4 that $p^* \in \mathcal{P}^{\sigma,heta}$ defined as in (1.4). □

4 CONCLUSION

Pointwise state constraints appear to be a powerful tool not only for incorporating physically or otherwise imposed a priori bounds on the state, but also for avoiding singularities in nonlinear PDEs without having to assume smallness in some higher order norm.

Further research in this direction will be concerned with different equations of nonlinear acoustics and their coupling to other physical fields. In particular, the state constraint approach seems promising for problems lacking $H^2(\Omega)$ smoothness due to interface coupling of nonconvex domains.

ACKNOWLEDGMENTS

The work of the first-named author was supported by the Austrian Science Fund (FWF) under grant SFB F32 (SFB "Mathematical Optimization and Applications in Biomedical Sciences"). The second-named author would like to thank the Austrian Science Fund (FWF) for financial support within the grant P24970 ("Mathematics of Nonlinear Acoustics: Analysis, Numerics, and Optimization").

The authors wish to thank Ira Neitzel, Christian Meyer, and Daniel Wachsmuth for pointing them to the localization technique from [Casas and Tröltzsch 2002].

REFERENCES

Adams, R. and Fournier, J. (2003). *Sobolev Spaces*. Elsevier, Oxford.

Alibert, J.-J. and Raymond, J.-P. (1998). *A Lagrange multiplier theorem for control problems with state constraints*. Numer. Funct. Anal. Optim. 19.7-8, pp. 697–704. DOI: 10.1080/01630569808816854.

Bonnans, J. F. and Casas, E. (1989). *Optimal control of semilinear multistate systems with state constraints*. SIAM J. Contr. Opt. 27.2, pp. 446–455. DOI: 10.1137/0327023.

Casas, E. and Tröltzsch, F. (2002). *Error estimates for the finite-element approximation of a semilinear elliptic control problem*. Control and Cybernetics 31, pp. 695–712.

Clason, C. and Kaltenbacher, B. (2012). *Optimal control of a singular PDE modeling transient MEMS with control or state constraints*. Tech. rep. 2012–017. SFB "Mathematical Optimization and Applications in Biomedical Sciences". URL: http://math.uni-graz.at/mobis/publications/SFB-Report-2012-017.pdf.

Clason, C., Kaltenbacher, B., and Veljovic, S. (2009). *Boundary optimal control of the Westervelt and the Kuznetsov equation*. Journal of Mathematical Analysis and Applications 356, pp. 738–751. DOI: 10.1016/j.jmaa.2009.03.043.

Kaltenbacher, B. and Lasiecka, I. (2009). *Global existence and exponential decay rates for the Westervelt equation*. Discrete and Continuous Dynamical Systems (DCDS) 2, pp. 503–525. DOI: 10.3934/dcdss.2009.2.503.

Westervelt, P. (1963). *Parametric acoustic array*. The Journal of the Acoustic Society of America 35, pp. 535–537. DOI: 10.1121/1.1918525.