Time optimal control of the wave equation, its regularization and numerical realization

K. Kunisch, D. Wachsmuth

RICAM-Report 2011-16

Time optimal control of the wave equation, its regularization and numerical realization

Karl Kunisch[1], Daniel Wachsmuth[2]

July 14, 2011

[1]Institute for Mathematics and Scientific Computing, Heinrichstraße 36, A–8010 Graz, Austria, karl.kunisch@uni-graz.at, http://www.kfunigraz.ac.at/imawww/kunisch, partially supported by 'Fonds zur Förderung der Wissenschaftlichen Forschung' under SFB 32, Mathematical Optimization and Applications in the Biomedical Sciences.

[2]Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstraße 69, A–4040 Linz, Austria, daniel.wachsmuth@oeaw.ac.at, http://www.ricam.oeaw.ac.at/people/page/wachsmuth

Abstract. An approximation procedure for time optimal control problems for the linear wave equation is analyzed. Its asymptotic behavior is investigated and an optimality system including a transversality conditions for the regularized and unregularized problems are derived. The regularized problem also serves as a starting point for numerical techniques. For this purpose a family of parameterized optimal control problems, depending on the time horizon τ as parameter, is introduced. Each member of this family can be solved by a locally superlinearly convergent semi-smooth Newton algorithm. It is verified that the derivative of its minimal value functional of the parameterized problems coincides with the transversality condition of the time optimal problems. Selected numerical examples are given.

Keywords. Time optimal control, wave equation, optimality condition, transversality condition, minimal value functional, semi-smooth Newton method.

1 Introduction

This paper is devoted to the time optimal control problem

$$
(\tilde{P}) \qquad \begin{cases} \min \int_0^\tau dt \\ \text{subject to } \tau \geq 0, \\ y_{tt} - \Delta y = \chi_\omega u \text{ in } (0,\tau)\times\Omega, \\ y(0) = y_1,\ y_t(0) = y_2,\ y(\tau) = z_1,\ y_t(\tau) = z_2 \text{ in } \Omega, \\ \|u(t)\|_{L^2(\omega)} \leq \gamma, \text{ for a.e. } t \in (0,\tau). \end{cases}
$$

Here, $\gamma > 0$ is a fixed positive constant and $\Omega \subset \mathbb{R}^n$, $n = 2$ or 3 is a fixed bounded domain with smooth boundary Γ. Further $\omega \subset \Omega$ is a measurable subset and $\chi_\omega u$ denotes the extension-by-zero operator from ω to Ω. The initial and terminal states are fixed and - unless specified otherwise - are assumed to satisfy

$$
y_1 \in H_0^1(\Omega),\ z_1 \in H_0^1(\Omega),\ y_2 \in L^2(\Omega),\ z_2 \in L^2(\Omega).
$$

We shall analyze an approximation scheme for $(\tilde{P})$ that can be used as the basis for numerical approximation schemes and also to derive first order necessary optimality systems. These conditions are complete in the sense that they contain as many equations as unknowns. The optimality system consists of the primal and adjoint equations, the maximum principle and the transversality condition. Time optimal control for the wave equation was investigated in the work of Fattorini [6, 7], but there the transversality condition is not addressed. In [2] the geometric form of the transversality conditions in infinite dimensions is derived along the well-known lines of Lee and Markus [14] in finite dimensions. It states that the terminal state of the adjoined equation is normal to a supporting hyperplane to the reachable cone at the optimal time and the state (z_1, z_2). This form, however, does not appear to be applicable computationally. Rather we aim for an analytical form of the transversality condition as utilized e.g. in [8], pg. 88, in the case of time optimal control for ordinary differential equations. In the case of time optimal control for parabolic problems such a form was obtained by Barbu in [1]. The technique used there does not appear to be applicable for $(\tilde{P})$. It could be used for the case where $\omega = \Omega$ and the terminal constraint $y_t(\tau) = z_2$ is not enforced.

The development of numerical techniques for time optimal control problems, has received much attention in the context of ordinary differential equations. They are frequently categorized into direct and indirect methods. Indirect methods based on multiple shooting techniques [4, 12] solve the two point boundary value problem describing first order necessary conditions. In [10] a semi-smooth Newton method was recently proposed for solving the non-smooth optimality systems. Direct methods on the other hand, consider time optimal problems as a genuine nonlinear programming problems. They are used in several variants, which frequently involve re-parametrization of the controls as the unknowns. The new unknowns can be the switching times as in [17] or the arc durations as in [11].

While the focus of this work is put on time-optimal control of the wave equation, it contains also two results that are of interest in a broader context: It will be proved that when considering a regularized form of $(\tilde{P})$ as minimization problem in the control u with τ as parameter, then the derivative of the value functional with respect to τ is given by the associated transversality condition. Secondly we analyze the semi-smooth Newton method for mixed integral (in space)-pointwise (in time) bounds. This widens the scope of the use of semi-smooth Newton methods which are so-far well analyzed for point-wise bounds.

The paper is organized as follows. In the remainder of this section we recall regularity results on the controlled wave equation and also provide its abstract form. Section 2 contains results on $(\tilde{P})$ and a discussion of controllability results as far as they are relevant for the present paper. In Section 3 we introduce a family of approximating problems and derive their optimality systems. These lend themselves to efficient numerical realizations. Then convergence of the primal variables of the approximating problems to a solution of $(\tilde{P})$ is shown. Convergence of the adjoint variables is addressed in Section 4. There the maximum principle and the transversality condition for $(\tilde{P})$ are obtained as well. As we shall see, the transversality condition requires us to impose additional assumptions. If these assumptions are not met then the justification of our approximation rests on the convergence of the primal variables only. Section 5 is devoted to the study of the parametric optimization problems that arise from the regularized version of $(\tilde{P})$ with u as the only optimization variable, and τ as parameter. We show that the derivative of the value functional with respect to τ is given by the transversality condition. These parametric optimization problems are the innermost building block of our numerical approach. Essentially they are optimal control problems for the wave equation with spatial L^2-norm bounds, pointwise in time, on the controls. We show that semi-smooth Newton methods are applicable for this class of problems and we verify their local superlinear convergence in Section 6. Here we show that the radial projections in $L^2(\Omega)$ are Newton differentiable. The final Section 7 is devoted to the explanation of the numerical approach and selected numerical examples.

Let us turn to

$$\begin{cases} y_{tt} - \Delta y = \chi_\omega u \text{ in } (0,T] \times \Omega, \\ y(0) = y_1,\, y_t(0) = y_2 \end{cases} \tag{1.1}$$

where $T > 0$ is fixed. For existence and uniqueness of weak and very weak solutions, we have the following well-known result, [5, 15].

Theorem 1.1. *Let $y_1 \in L^2(\Omega)$, $y_2 \in H^{-1}(\Omega)$, and $u \in L^2(0,T;L^2(\omega))$ be given. Then there exists a uniquely determined very weak solution y of* (1.1), *which satisfies*

$$y \in C(0,T;L^2(\Omega)),\ y_t \in C(0,T;H^{-1}(\Omega)).$$

If in addition $y_1 \in H_0^1(\Omega)$ and $y_2 \in L^2(\Omega)$ holds, then there exists a uniquely determined weak solution y of (1.1), *which satisfies*

$$y \in C(0,T;H_0^1(\Omega)),\ y_t \in C(0,T;L^2(\Omega)),\ y_{tt} \in L^2(0,T;H^{-1}(\Omega)).$$

If moreover $y_1 \in H^2(\Omega)$, $y_2 \in H_0^1(\Omega)$, and $u_t \in L^2(0,T;L^2(\omega))$ holds, then the solution y satisfies

$$y \in C(0,T;H^2(\Omega)),\ y_t \in C(0,T;H_0^1(\Omega)),\ y_{tt} \in C(0,T;L^2(\Omega)).$$

In all cases, the mapping from (u, y_1, y_2) to (y, y_t) is linear and continuous between the indicated spaces.

To express (1.1) in abstract form we introduce the operators

$$\mathbf{A} := \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}, \quad \mathbf{B} := \begin{pmatrix} 0 \\ \chi_\omega \end{pmatrix},$$

and vectors

$$\mathbf{y}_0 := \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \ \mathbf{z} := \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

$$\mathbf{y}(t) := \begin{pmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{pmatrix}.$$

Then the wave equation (1.1) is equivalent to the first-order evolution equation

$$\begin{aligned} \mathbf{y}_t &= \mathbf{A}\mathbf{y} + \mathbf{B}u, \\ \mathbf{y}(0) &= \mathbf{y}_0. \end{aligned} \tag{1.2}$$

The components of the solution of this equation fulfill $\mathbf{y}_2 = (\mathbf{y}_1)_t$. For convenience of notation we introduce the function spaces

$$Y^s = \begin{cases} H^s(\Omega) & 0 \le s < 1/2 \\ H^s(\Omega) \cap \{y :\ y|_\Gamma = 0\} & s \ge 1/2 \\ (H^{-s}(\Omega))^* & s < 0. \end{cases}$$

Let us define the following spaces, which take account of the regularity of the components of solutions $\mathbf{y}$ of (1.2)

$$\mathbf{Y}^s := Y^s \times Y^{s-1}.$$

Here, the index s means that the first component of the vector function $\mathbf{y} \in \mathbf{Y}^s$ is in $H^s(\Omega)$, whereas the second component, which is its time derivative, is in $H^{s-1}(\Omega)$. Given this notation, we have that the operator $\mathbf{A}$ is linear and continuous in the following sense

$$\mathbf{A} \in \mathcal{L}(\mathbf{Y}^s,\ \mathbf{Y}^{s-1}).$$

Moreover, the operator $\mathbf{B}$ has the property

$$\mathbf{B} \in \mathcal{L}(L^2(\omega), \mathbf{Y}^1).$$

With this notation, the previous results of Theorems 1.1 can be expressed as

Corollary 1.2. *Let $\mathbf{y}_0 \in \mathbf{Y}^0$, $u \in L^2(0,T;L^2(\omega))$ be given. Then the first-order equation* (1.2) *admits a unique very weak solution $\mathbf{y}$ that satisfies*

$$\mathbf{y} \in C(0,T;\mathbf{Y}^0).$$

If in addition $\mathbf{y}_0 \in \mathbf{Y}^1$ holds, then the first-order equation (1.2) *admits a unique weak solution $\mathbf{y}$ that satisfies*

$$\mathbf{y} \in C(0,T;\mathbf{Y}^1),\ \mathbf{y}_t \in L^2(0,T;\mathbf{Y}^0)$$

If moreover $\mathbf{y}_0 \in \mathbf{Y}^2$, $u_t \in L^2(0,T;L^2(\omega))$ then

$$\mathbf{y} \in C(0,T;\mathbf{Y}^2),\ \mathbf{y}_t \in C(0,T;\mathbf{Y}^1).$$

2 Problem $(\tilde{P})$, its transformation and controllability

Concerning existence of a solution to $(\tilde{P})$ we have:

Theorem 2.1. *If there exists an admissible control u driving $\mathbf{y}_0$ to $\mathbf{z}$, then there exists a solution (τ^*, y^*, u^*) of $(\tilde{P})$. Moreover τ^* is unique.*

Proof. A subsequential limit argument based on Theorem 1.1 readily implies the assertion. □

In all that follows, it will be convenient to transform problem $(\tilde{P})$ to the fixed time interval $I = (0,1)$. The transformed time optimal control problem reads:

$$
\text{(P)} \qquad \begin{cases} \min \tau \\ \text{subject to } \tau \geq 0 \text{ and} \\ y_{tt} - \tau^2 \Delta y = \tau^2 \chi_\omega u \\ y(0) = y_1, \ \partial_t y(0) = \tau y_2 \\ y(1) = z_1, \ \partial_t y(1) = \tau z_2 \\ \|u(t)\|_{L^2(\omega)} \leq \gamma, \ \text{for a.e. } t \in (0, \tau). \end{cases}
$$

Obviously, the original problem $(\tilde{P})$ and the transformed problem (P) are equivalent.

Let us define the set of admissible controls by

$$
U_{ad} := \{u \in L^\infty(I; L^2(\omega)) : \ u(t) \in U \text{ a.e. on } I\},
$$

where U is given by

$$
U = \{u \in L^2(\omega) : \ \|u\|_{L^2(\omega)} \leq \gamma\}.
$$

Using the abstract operators introduced in Section 1, problem (P) is further equivalent to

$$
\text{(P)} \qquad \begin{cases} \min \tau \\ \text{subject to } \tau \geq 0 \text{ and} \\ y_t = \tau(\mathbf{A}\mathbf{y} + \mathbf{B}u) \text{ on } (0, 1], \\ \mathbf{y}(0) = \Theta_\tau \mathbf{y}_0, \ \mathbf{y}(1) = \Theta_\tau z, \\ u \in U_{ad}, \end{cases}
$$

where $\mathbf{y}_0 = (y_1, y_2)$, $\mathbf{z} = (z_1, z_2) \in \mathbf{Y}^1$, and $\Theta_\tau = \begin{pmatrix} I & 0 \\ 0 & \tau \end{pmatrix}$. We turn to a brief discussion of controllability. System (1.1) is called controllable in time $T > 0$, if for any $(y_1, y_2) \in \mathbf{Y}^1$, there exists $u \in L^2(0, T; \omega)$ such that this control drives the system to rest, in the sense that $y(T) = y_t(T) = 0$. Controllability is equivalent to the following observability inequality

$$
|(p_1, p_2)|_{L^2(\Omega) \times H^{-1}(\Omega)} \leq C \int_0^T \int_\omega |p(t, x)|^2 \, dx dt, \tag{2.1}
$$

where

$$\begin{cases} p_{tt} - \Delta p = 0 \text{ in } (0,T) \times \omega \\ p\,(T) = p_1,\, p_t(T) = p_2, \end{cases}$$

for a constant C independent of $(p_1, p_2) \in L^2(\omega) \times H^{-1}(\omega) = \mathbf{Y}^0$, see e.g. [19]. Due to finite speed of propagation T and ω have to satisfy appropriate conditions for controllability to hold. For $\omega \subset \Omega$, it is required that T is sufficiently large and that ω satisfies certain geometric conditions as for instance the geometric control condition, see for instance [3] and the references given in [19].

3 A family of regularized problems

In this section we discuss a family of regularized problems that is obtained by penalization. Their optimality condition is derived and convergence of the primal variables is proven. For $\varepsilon > 0$ we consider

$$(P_\varepsilon) \qquad \begin{cases} \min J_\varepsilon(\tau, u) = \tau \left(1 + \frac{\varepsilon}{2} \|u\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon} \|\mathbf{y}(1) - \Theta_\tau \mathbf{z}\|^2_{\mathbf{Y}^0}, \\ \text{subject to } \tau \geq 0 \text{ and} \\ \mathbf{y}_t = \tau \mathbf{A} \mathbf{y} + \tau \mathbf{B} u, \text{ on } (0,1], \\ \mathbf{y}(0) = \Theta_\tau \mathbf{y}_0, \\ u \in U_{ad}. \end{cases}$$

Here and below the norm on $\mathbf{Y}^0 = L^2(\Omega) \times H^{-1}(\Omega)$ is chosen to be

$$\|\mathbf{v}\|^2_{\mathbf{Y}^0} = \|v_1\|^2_{L^2(\Omega)} + ((-\Delta)^{-1} v_2, v_2)_{L^2(\Omega)},$$

where $w = (-\Delta)^{-1}\, v_2$ is the solution of

$$-\Delta w = v_2 \text{ in } \Omega,\ w = 0 \text{ on } \Gamma.$$

Proposition 3.1. *Problem (P_ε) admits a solution.*

Proof. Let (τ_n, y_n, u_n) denote a minimizing sequence. Since J_ε is bounded from below, the sequence τ_n is bounded, and therefore admits an accumulation point $\tilde{\tau}$. Since $u_n \in U_{ad}$ for all n, the sequence $\{u_n\}$ is bounded in

$L^\infty(I;L^2(\Omega))$. By Theorem 1.1, the sequence $\mathbf{y}_n$ is bounded in $C(I;\mathbf{Y}^1)\cap C^1(I;\mathbf{Y}^0)$. Choosing a weakly converging subsequence of $\{u_n\}$ in $L^2(I;L^2(\omega))$ and of y_n in $L^2(I;\mathbf{Y}^1)\cap H^1(I;\mathbf{Y}^0)$ there exists $(\tilde\tau,\tilde y,\tilde u)$ such that we can pass to the weak subsequential limit in

$$\begin{cases}(\mathbf{y}_n)_t=\tau_n(A\mathbf{y}_n+Bu_n)\\ \mathbf{y}_n(0)=\Theta_{\tau_n}\mathbf{y}_0\end{cases}$$

to find that

$$\begin{aligned}&\tilde{\mathbf{y}}_t=\tilde\tau(A\tilde{\mathbf{y}}+B\tilde u)\\ &\tilde{\mathbf{y}}(0)=\Theta_{\tilde\tau}\mathbf{y}_0.\end{aligned}$$

Weak lower semi-continuity of J_ε implies that $(\tilde\tau,\tilde u)$ is a solution to (P_ε). □

In the sequel, $(\tau_\varepsilon,\mathbf{y}_\varepsilon,u_\varepsilon)$ denotes a solution of the penalized problem (P_ε) for $\varepsilon>0$.

Theorem 3.2. *For $\varepsilon\to0^+$ let $\{(\tau_\varepsilon,\mathbf{y}_\varepsilon,u_\varepsilon)\}_{\varepsilon>0}$ denote a family of solutions of (P_ε). Then we have that*

$$\tau_\varepsilon\to\tau^*,\ \text{for } \varepsilon\to0^+,$$

and $(\mathbf{y}_\varepsilon,u_\varepsilon)$ is uniformly bounded in $(L^\infty(I;\mathbf{Y}^1)\cap H^1(I;\mathbf{Y}^0))\times L^\infty(T;L^2(\omega))$. Moreover, for each weakly-star converging subsequence $\{(\mathbf{y}_{\varepsilon_n},u_{\varepsilon_n})\}$ with

$$\mathbf{y}_{\varepsilon_n}\rightharpoonup^*\tilde{\mathbf{y}}\ \text{in } L^\infty(I;\mathbf{Y}^1)\cap H^1(I;\mathbf{Y}^0),\quad u_{\varepsilon_n}\rightharpoonup^*\tilde u\ \text{in } L^\infty(I;L^2(\omega)),$$

the limit $(\tilde{\mathbf{y}},\tilde u)$ is a solution of the original time-optimal control problem (P).

If $\tilde u$ is bang-bang, i.e. $\|\tilde u(t)\|_{L^2(\omega)}=\gamma$ for a.e. $t\in I$, then the convergence $(\mathbf{y}_{\varepsilon_n},u_{\varepsilon_n})\to(\tilde{\mathbf{y}},\tilde u)$ is strong in $(C(I;\mathbf{Y}^1)\cap H^1(I;\mathbf{Y}^0))\times L^2(I;L^2(\omega))$.

Proof. Let $(\tau^*,\mathbf{y}^*,u^*)$ denote a solution of $(\tilde P)$. Since it is feasible for the penalized problem we have

(3.1)
$$\tau_\varepsilon\left(1+\frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))}\right)+\frac{1}{2\varepsilon}\|\mathbf{y}_\varepsilon(1)-\Theta_{\tau_\varepsilon}\mathbf{z}\|^2_{\mathbf{Y}^0}\le\tau^*\left(1+\frac{\varepsilon}{2}\|u^*\|^2_{L^2(I;L^2(\omega))}\right).$$

This implies that $\limsup\tau_\varepsilon\le\tau^*$ and hence $\{\tau_\varepsilon\}_{\varepsilon>0}$ is bounded.

Moreover due to the control constraints and Corollary 1.2, the set $\{\mathbf{y}_\varepsilon,\ u_\varepsilon\}_{\varepsilon\in(0,1]}$ is bounded in

$$(C(I;\mathbf{Y}^1)\cap H^1(I;\mathbf{Y}^0))\ \times\ L^\infty(I;L^2(\omega)).$$

Let us choose a subsequence $\varepsilon_n \to 0$ such that $\tau_{\varepsilon_n} \to \tilde{\tau}$, $u_{\varepsilon_n} \rightharpoonup^* \tilde{u}$ in $L^\infty(I; L^2(\omega))$, $\mathbf{y}_{\varepsilon_n} \rightharpoonup^* \tilde{\mathbf{y}}$ in $L^\infty(I; \mathbf{Y}^1) \cap H^1(I; \mathbf{Y}^0)$ as $n \to \infty$. Arzela-Ascoli's theorem and the compact embedding $\mathbf{Y}^1$ in $\mathbf{Y}^0$ imply that we can choose ε_n such that $\mathbf{y}_{\varepsilon_n} \to \tilde{\mathbf{y}}$ in $C(I; \mathbf{Y}^0)$ as $\varepsilon_n \to 0$. By (3.1)

$$\mathbf{y}_{\varepsilon,1}(1) \to z_1 \text{ in } L^2(\Omega), \quad \tau_\varepsilon \mathbf{y}_{\varepsilon,2}(1) \to \tilde{\tau} z_2 \text{ in } H^{-1}(\Omega).$$

It further follows that $\tilde{\mathbf{y}}$ solves

$$\begin{aligned} \tilde{\mathbf{y}}_t &= \tilde{\tau}\mathbf{A}\tilde{\mathbf{y}} + \tilde{\tau}\mathbf{B}\tilde{u}, \\ \tilde{\mathbf{y}}(0) &= \Theta_{\tilde{\tau}}\mathbf{y}_0,\, \tilde{\mathbf{y}}(1) = \Theta_{\tilde{\tau}}\mathbf{z}. \end{aligned}$$

Since U_{ad} is weakly closed, we have that $\tilde{u} \in U_{ad}$. Hence, $(\tilde{\tau}, \tilde{\mathbf{y}}, \tilde{u})$ is feasible for the time-optimal control problem with $\tilde{\tau} \leq \tau^*$. Since τ^* was the minimal time, $\tilde{\tau} = \tau^*$ and $(\tilde{\tau}, \tilde{\mathbf{y}}, \tilde{u})$ is a solution of the time-optimal control problem. The minimal time τ^* is unique, and consequently the whole family τ_ε converges to τ^* as $\varepsilon \to 0^+$.

If $\tilde{u}$ is bang-bang, then we have by feasibility of u_{ε_n} and weakly lower-semicontinuity of norms

$$\begin{aligned} \|\tilde{u}\|_{L^2(I;L^2(\omega))} = \gamma \geq \limsup \|u_{\varepsilon_n}\|_{L^2(I;L^2(\omega))} \\ \geq \liminf \|u_{\varepsilon_n}\|_{L^2(I;L^2(\omega))} \geq \|\tilde{u}\|_{L^2(I;L^2(\omega))}. \end{aligned}$$

This implies norm convergence $\lim_{n\to\infty} \|u_{\varepsilon_n}\|_{L^2(I;L^2(\omega))} = \|\tilde{u}\|_{L^2(I;L^2(\omega))}$ and the strong convergence, as stated, is obtained. □

The assumption on the bang-bang nature of all time-optimal controls holds for the case $\omega = \Omega$, see Fattorini [7, Thm. 6.12.3, p. 304]. The verification utilizes a controllability property.

Corollary 3.3. *Under the assumptions of the previous theorem, we have*

$$\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon}\mathbf{z}\|^2_{\mathbf{Y}^0} \leq \varepsilon(|\tau_\varepsilon - \tau^*| + O(\varepsilon))$$

for $\varepsilon \to 0^+$.

Proof. This is a direct consequence of (3.1) and the fact that the set of admissible controls is bounded. □

Lemma 3.4. *Let $\varepsilon > \varepsilon' > 0$ be given. Then the solutions $(\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon)$ and $(\tau_{\varepsilon'}, \mathbf{y}_{\varepsilon'}, u_{\varepsilon'})$ satisfy*

$$\tau_\varepsilon \|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} - \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon}\mathbf{z}\|^2_{Y^0} \le \tau_{\varepsilon'}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))} - \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0}.$$

Proof. By optimality of $(\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon)$ we have

$$\begin{aligned}
&\tau_\varepsilon \left(1 + \frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon}\|\mathbf{y}_\varepsilon(1) - \Theta_\varepsilon \mathbf{z}\|^2_{Y^0} \\
&\quad \le \tau_{\varepsilon'} \left(1 + \frac{\varepsilon}{2}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0} \\
&\quad \le \tau_{\varepsilon'} \left(1 + \frac{\varepsilon'}{2}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon'}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0} \\
&\qquad + \frac{\varepsilon - \varepsilon'}{2}\left(\tau_{\varepsilon'}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))} - \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0}\right).
\end{aligned} \tag{3.2}$$

This can be further estimated by optimality of $(\tau_{\varepsilon'}, \mathbf{y}_{\varepsilon'}, u_{\varepsilon'})$ by

$$\begin{aligned}
&\tau_\varepsilon \left(1 + \frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon}\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon} \mathbf{z}\|^2_{Y^0} \\
&\quad \le \tau_\varepsilon \left(1 + \frac{\varepsilon'}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon'}\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon}\mathbf{z}\|^2_{Y^0} \\
&\qquad + \frac{\varepsilon - \varepsilon'}{2}\left(\tau_{\varepsilon'}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))} - \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0}\right).
\end{aligned}$$

Hence

$$\begin{aligned}
&\frac{\varepsilon - \varepsilon'}{2}\left(\tau_\varepsilon\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} - \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0}\right) \\
&\qquad \le \frac{\varepsilon - \varepsilon'}{2}\left(\tau_{\varepsilon'}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))} - \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0}\right),
\end{aligned}$$

which proves the claim for $\varepsilon > \varepsilon' > 0$. □

Let $V(\varepsilon)$ denote the optimal value functional of the penalized problem, i.e.

$$V(\varepsilon) := J_\varepsilon(\mathbf{y}_\varepsilon, u_\varepsilon)$$

subject to the constraints in (P_ε).

Proposition 3.5. *The optimal value function V is locally Lipschitz continuous from $(0,\infty)$ to $\mathbb{R}$.*

Proof. For $\varepsilon, \varepsilon' > 0$, we obtain by (3.2), which does not utilize $\varepsilon > \varepsilon'$

$$V(\varepsilon) \le V(\varepsilon') + \frac{\varepsilon - \varepsilon'}{2}\left(\tau_{\varepsilon'}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))} - \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0}\right),$$

which gives

$$\begin{aligned}|V(\varepsilon) - V(\varepsilon')| \le \frac{|\varepsilon - \varepsilon'|}{2}\Big(&\tau_\varepsilon\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} + \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon}\mathbf{z}\|^2_{Y^0}\\ &+ \tau_{\varepsilon'}\|u_{\varepsilon'}\|^2_{L^2(I;L^2(\omega))} + \frac{1}{\varepsilon\varepsilon'}\|\mathbf{y}_{\varepsilon'}(1) - \Theta_{\tau_{\varepsilon'}}\mathbf{z}\|^2_{Y^0}\Big).\end{aligned}$$

The term in parentheses on the right-hand side is bounded for $\varepsilon, \varepsilon'$ in bounded subsets of $(0,\infty)$ and hence the claim follows. ☐

First order necessary optimality conditions for (P_ε) are derived next. To employ the method of transposition the adjoint state $\mathbf{p}_\varepsilon$ is defined as the solution to

$$-\mathbf{p}_t = \tau\mathbf{A}^*\mathbf{p}, \tag{3.3}$$

with terminal condition

$$\mathbf{p}(1) = \frac{1}{\varepsilon}\begin{pmatrix}\mathbf{y}_{\varepsilon,1}(1) - z_1\\ (-\Delta)^{-1}(\mathbf{y}_{\varepsilon,2}(1) - \tau_\varepsilon z_2)\end{pmatrix}. \tag{3.4}$$

Here $\mathbf{A}^*$ is given as adjoint of $\mathbf{A}$:

$$\mathbf{A}^* := \begin{pmatrix} 0 & \Delta \\ I & 0\end{pmatrix}.$$

Hence, equation (3.3) with terminal condition (3.4) is a wave equation in the second coordinate $\mathbf{p}_{\varepsilon,2}$ with

$$\partial_t\mathbf{p}_{\varepsilon,2} = -\tau_\varepsilon\mathbf{p}_{\varepsilon,1}.$$

It will be convenient to introduce the notation

$$\mathbf{P}^s := Y^{s-1} \times Y^s,$$

which will be used for $s = 0, 1, 2$. The index s with $\mathbf{Y}^s$ and $\mathbf{P}^s$ denotes the regularity of the wave function for the primal state $\mathbf{y}$ and the adjoint state $\mathbf{p}$. We may note that $(\mathbf{Y}^s)^* = \mathbf{P}^{(1-s)}$. By Corollary 1.2 we have the following regularity result for $\mathbf{p}_\varepsilon$.

Corollary 3.6. *Let $\mathbf{y}_\varepsilon \in C(I; \mathbf{Y}^1)$. Then* (3.3) - (3.4) *admits a unique solution $\mathbf{p}_\varepsilon \in C(I; \mathbf{P}^2) \cap C^1(I; \mathbf{P}^1)$.*

Theorem 3.7. *Let $(\tau_\varepsilon, y_\varepsilon, u_\varepsilon)$ be a local solution of (P_ε). Then there exists $\mathbf{p}_\varepsilon \in C(I; \mathbf{P}^2) \cap C^1(I; \mathbf{P}^1)$ such that the following optimality system holds:*

$$
\text{(3.5)} \qquad \begin{cases}
\partial_t \mathbf{y}_\varepsilon = \tau_\varepsilon \mathbf{A} \mathbf{y}_\varepsilon + \tau_\varepsilon \mathbf{B} u_\varepsilon, \quad \mathbf{y}_\varepsilon(0) = \Theta_{\tau_\varepsilon} \mathbf{y}_0 \\
-\partial_t \mathbf{p}_\varepsilon = \tau_\varepsilon \mathbf{A}^* \mathbf{p}_\varepsilon, \quad \mathbf{p}_\varepsilon(1) = \dfrac{1}{\varepsilon} \begin{pmatrix} \mathbf{y}_{\varepsilon,1}(1) - z_1 \\ (-\Delta)^{-1}(\mathbf{y}_{\varepsilon,2}(1) - \tau_\varepsilon z_2) \end{pmatrix} \in \mathbf{P}^2 \\
\tau_\varepsilon(\varepsilon u_\varepsilon + \mathbf{B}^* \mathbf{p}_\varepsilon, u - u_\varepsilon)_{L^2(I; L^2(\omega))} \geq 0, \ \textit{for all } u \in U_{ad} \\
1 + \dfrac{\varepsilon}{2} \|u_\varepsilon\|^2_{L^2(I; L^2(\omega))} + \langle \mathbf{A} \mathbf{y}_\varepsilon + \mathbf{B} u_\varepsilon, \ \mathbf{p}_\varepsilon \rangle_{L^2(I; \mathbf{Y}^0), L^2(I; \mathbf{P}^1)} + \\
\qquad (y_2, \mathbf{p}_{\varepsilon,2}(0))_{L^2(\Omega)} - (z_2, \mathbf{p}_{\varepsilon,2}(1))_{L^2(\Omega)} = 0.
\end{cases}
$$

The optimal control u_ε has the additional regularity

$$u_\varepsilon \in C(I; L^2(\omega)) \ \textit{and} \ \partial_t u_\varepsilon \in L^\infty(I; L^2(\omega)).$$

Moreover, if $\mathbf{y}_0 \in \mathbf{Y}^2$ then

$$\mathbf{y}_\varepsilon \in C(I; \mathbf{Y}^2) \cap C^1(I; \mathbf{Y}^1).$$

We refer to the four assertions in (3.5) as primal- and adjoint equations, optimality- and transversality condition. Note also, that $\mathbf{B}^* \mathbf{p}_\varepsilon$ is the restriction operator to ω given by $(\mathbf{B}^* \mathbf{p}_\varepsilon)(x) = \mathbf{p}_{\varepsilon,2}(x)$ for $x \in \omega$.

Proof. Let us take $u \in U_{ad}$ and set $h = u - u_\varepsilon$. Further let $\tilde{\mathbf{y}} \in C(I; \mathbf{Y}^1)$ denote the solution to the sensitivity equation of the primal equation, i.e. to

$$
\begin{aligned}
\partial_t \tilde{\mathbf{y}} &= \tau_\varepsilon(\mathbf{A} \tilde{\mathbf{y}} + \mathbf{B} h) \\
\tilde{\mathbf{y}}(0) &= 0.
\end{aligned}
$$

Then the directional derivative $\partial_u J_\varepsilon(\tau_\varepsilon, u_\varepsilon)h$ at $(\tau_\varepsilon, u_\varepsilon)$ in the admissible directions h satisfies $\partial_u J_\varepsilon(\tau_\varepsilon, u_\varepsilon)h \geq 0$ and it is given by

$$\begin{aligned}
\partial_u J_\varepsilon(\tau_\varepsilon, u_\varepsilon)h &= \varepsilon\tau_\varepsilon(u_\varepsilon, h)_{L^2(I;L^2(\omega))} + (\tilde{\mathbf{y}}(1), \mathbf{p}_\varepsilon(1))_{L^2(\Omega)} \\
&= \varepsilon\tau_\varepsilon(u_\varepsilon, h)_{L^2(I;L^2(\omega))} + \int_0^1 \frac{d}{dt}(\tilde{\mathbf{y}}(t), \mathbf{p}_\varepsilon(t))_{L^2(\Omega)} \\
&= \varepsilon\tau_\varepsilon(u_\varepsilon, h)_{L^2(I;L^2(\omega))} + \tau_\varepsilon \int_0^1 \big(\langle A\tilde{\mathbf{y}} + \mathbf{B}h, \mathbf{p}_\varepsilon\rangle - \langle \tilde{\mathbf{y}}, \mathbf{A}^*\mathbf{p}_\varepsilon\rangle\big)\, dt \\
&= \tau_\varepsilon(\varepsilon u_\varepsilon + \mathbf{B}^*\mathbf{p}_\varepsilon, h)_{L^2(I;L^2(\omega))} \geq 0,
\end{aligned}$$

proves the optimality condition. Similarly let $\hat{\mathbf{y}}$ denote the solution of the sensitivity equation of the primal equation with respect to $\tau > 0$, i.e. of

$$\begin{aligned}
\partial_t \hat{\mathbf{y}} &= \tau_\varepsilon \mathbf{A}\hat{\mathbf{y}} + A\mathbf{y}_\varepsilon + \mathbf{B}u_\varepsilon \\
\hat{\mathbf{y}}(0) &= \begin{pmatrix} 0 \\ y_2 \end{pmatrix}.
\end{aligned}$$

Then

$$\begin{aligned}
0 = \partial_\tau J_\varepsilon(\tau_\varepsilon, u_\varepsilon) &= 1 + \frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} + \frac{1}{\varepsilon}\langle \mathbf{y}_\varepsilon(1) - \Theta_\tau \mathbf{z},\, \hat{\mathbf{y}}(1)\rangle_{\mathbf{Y}^0} \\
&\quad - \frac{1}{\varepsilon}\langle \mathbf{y}_\varepsilon(1) - \Theta_\tau \mathbf{z},\, \Theta'_\tau \mathbf{z}\rangle_{\mathbf{Y}^0} \\
&= 1 + \frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} + (\hat{\mathbf{y}}(1), \mathbf{p}_\varepsilon(1))_{L^2(\Omega)} - (\hat{\mathbf{p}}_{\varepsilon,2}(1), z_2)_{L^2(\Omega)} \\
&= 1 + \frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} + \int_0^1 \frac{d}{dt}(\hat{\mathbf{y}}(t), \mathbf{p}_\varepsilon(t))_{L^2(\Omega)} + (y_2, \mathbf{p}_{\varepsilon,2}(0))_{L^2(\Omega)} \\
&\quad - (z_2, \mathbf{p}_{\varepsilon,2}(1))_{L^2(\Omega)} \\
&= 1 + \frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} + \int_0^1 \big(\langle \tau_\varepsilon \mathbf{A}\hat{\mathbf{y}}(t) + \mathbf{A}\mathbf{y}_\varepsilon(t) + \mathbf{B}u_\varepsilon(t), \mathbf{p}_\varepsilon(t)\rangle \\
&\quad - (\hat{\mathbf{y}}(t), \tau_\varepsilon \mathbf{A}^*\mathbf{p}_\varepsilon(t))\big)\, dt + (y_2, \mathbf{p}_{\varepsilon,2}(0))_{L^2(\Omega)} - (z_2, \mathbf{p}_{\varepsilon,2}(1))_{L^2(\Omega)} \\
&= 1 + \frac{\varepsilon}{2}\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} + \int_0^1 \langle \mathbf{A}\mathbf{y}_\varepsilon(t) + \mathbf{B}u_\varepsilon(t), \mathbf{p}_\varepsilon(t)\rangle\, dt \\
&\quad + (y_2, \mathbf{p}_{\varepsilon,2}(0))_{L^2(\Omega)} - (z_2, \mathbf{p}_{\varepsilon,2}(1))_{L^2(\Omega)}
\end{aligned}$$

which proves the transversality condition. The optimality condition is equivalent to

$$u_\varepsilon(t) = \mathrm{P}_U\left(-\frac{1}{\varepsilon}\chi_\omega \mathbf{p}_{\varepsilon,2}(t)\right) \qquad \text{f.a.a. } t \in I,$$

which implies that $u_\varepsilon \in C(I; L^2(\omega))$, hence the point-wise representation holds everywhere on the closed interval. Here P_U denotes the canonical projection in $L^2(\omega)$ onto U. With Lemma 3.8 below, we conclude $\partial_t u_\varepsilon \in L^\infty(I; L^2(\omega))$, which in turn gives the higher regularity $\mathbf{y}_\varepsilon \in C(I; \mathbf{Y}^2)$. □

Lemma 3.8. *Let $q \in C^1(I; L^2(\omega))$ be given. Then u defined by*

$$u(t) = \mathrm{P}_U(q(t))$$

is in $C(I; L^2(\omega))$ with $\partial_t u \in L^\infty(I; L^2(\omega))$ and $\|\partial_t u\|_{L^\infty(I;L^2(\omega))} \leq 2\|\partial_t q\|_{L^\infty(I;L^2(\omega))}$.

Proof. The definition of U implies that

$$u(t) = \frac{q(t)}{\max(1, \|q(t)\|_{L^2(\omega)}/\gamma)},$$

which proves $u \in C^{0,1}(I; L^2(\omega))$. Then the weak time derivative of u satisfies

$$\partial_t u(t) = \begin{cases} \partial_t q(t) & \|q(t)\|_{L^2(\omega)} \leq \gamma \\ \gamma \frac{\|q(t)\|^2_{L^2(\omega)} \partial_t q(t) - (q(t), \partial_t q(t))_{L^2(\omega)} q(t)}{\|q(t)\|^3_{L^2(\omega)}} & \|q(t)\|_{L^2(\omega)} > \gamma. \end{cases}$$

Due to the regularity of q we have $\partial_t u \in L^\infty(I; L^2(\omega))$. □

The following two corollaries concern themselves with the transversality condition of the regularized problems.

Corollary 3.9. *(Pointwise transversality). Let $\mathbf{y}_0 \in \mathbf{Y}^2$ and let $(\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon)$ be a local solution of (P_ε) with associated adjoint state $\mathbf{p}_\varepsilon$. Then we have*

$$\begin{aligned} 1 + \frac{\varepsilon}{2}\|u_\varepsilon(t)\|^2_{L^2(\omega)} + (\mathbf{A}\mathbf{y}_\varepsilon(t) + \mathbf{B}u_\varepsilon(t),\ \mathbf{p}_\varepsilon(t))_{L^2(\Omega)} \\ +(y_2, \mathbf{p}_{\varepsilon,2}(0))_{L^2(\Omega)} - (z_2, \mathbf{p}_{\varepsilon,2}(1))_{L^2(\Omega)} = 0 \end{aligned}$$

for all $t \in \bar{I}$.

Proof. Due to the regularity of $\mathbf{y}_\varepsilon$ and $\mathbf{p}_\varepsilon$, we have $\mathbf{y}_\varepsilon \in C^1(I; \mathbf{Y}^1)$, $\mathbf{A}\mathbf{y}_\varepsilon \in C^1(I; \mathbf{Y}^0) = C^1(I; (\mathbf{P}^1)^*)$, and $\mathbf{p}_\varepsilon \in C^1(I; \mathbf{P}^1)$. Hence the mapping $t \mapsto \langle \mathbf{A}\mathbf{y}_\varepsilon(t),\ \mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1}$ is continuously differentiable. Additionally, the optimality conditions imply

$$(\varepsilon u_\varepsilon(t) + \mathbf{B}^*\mathbf{p}_\varepsilon(t),\ \partial_t u_\varepsilon(t))_{L^2(\omega)} = 0 \tag{3.6}$$

for almost all $t \in I$. Let us define the function

$$g(t) := \frac{\varepsilon}{2}\|u_\varepsilon(t)\|^2_{L^2(\omega)} + \langle \mathbf{A}\mathbf{y}_\varepsilon(t) + \mathbf{B}u_\varepsilon(t),\ \mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1}.$$

Differentiating w.r.t. t, we obtain for almost all $t \in I$

$$\begin{aligned}\partial_t g(t) &= (\varepsilon u_\varepsilon(t) + \mathbf{B}^*\mathbf{p}_\varepsilon(t),\ \partial_t u_\varepsilon(t))_{L^2(\omega)}\\ &\quad + \langle \mathbf{A}\mathbf{y}_\varepsilon(t) + \mathbf{B}u_\varepsilon(t),\ \partial_t\mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1} + \langle \mathbf{A}\partial_t\mathbf{y}_\varepsilon(t),\ \mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1},\end{aligned}$$

which proves that $\partial_t g \in L^2(I)$. Taking (3.6) into account, we obtain

$$\begin{aligned}\partial_t g(t) &= \langle \mathbf{A}\mathbf{y}_\varepsilon(t) + \mathbf{B}u_\varepsilon(t),\ \partial_t\mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1} + \langle \mathbf{A}\partial_t\mathbf{y}_\varepsilon(t),\ \mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1}\\ &= \frac{1}{\tau_\varepsilon}\big(\langle \partial_t\mathbf{y}_\varepsilon(t),\ \partial_t\mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1} - \langle \partial_t\mathbf{y}_\varepsilon(t),\ \partial_t\mathbf{p}_\varepsilon(t)\rangle_{(\mathbf{P}^1)^*,\mathbf{P}^1}\big) = 0,\end{aligned}$$

hence $\partial_t g(t) = 0$ on I, and $g(t) = \text{constant}$ on I. This implies the claim. □

The transversality condition of Corollary 3.9 for $t = 1$ can be simplified using the explicit expression for $\mathbf{p}_\varepsilon(1)$.

Corollary 3.10. *Let $(\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon)$ be a local solution of (P_ε) with associated adjoint state $\mathbf{p}_\varepsilon$. Then we have*

$$(\mathbf{A}\mathbf{y}_\varepsilon(1),\, \mathbf{p}_\varepsilon(1))_{L^2(\Omega)^2} = \langle \Theta_{\tau_\varepsilon}\mathbf{z},\, \mathbf{A}^*\mathbf{p}_\varepsilon(1)\rangle_{L^2(\Omega)^2}.$$

If additionally $\mathbf{z} \in \mathbf{Y}^2$ then it holds

$$(\mathbf{A}\mathbf{y}_\varepsilon(1),\, \mathbf{p}_\varepsilon(1))_{L^2(\Omega)^2} = \langle \mathbf{A}\Theta_{\tau_\varepsilon}\mathbf{z},\, \mathbf{p}_\varepsilon(1)\rangle_{(\mathbf{P}^0)^*,\mathbf{P}^0}. \tag{3.7}$$

Proof. Using the definition of $\mathbf{p}_\varepsilon(1)$ and of the operator $\mathbf{A}$, we find

$$\begin{aligned}(\mathbf{A}\mathbf{y}_\varepsilon(1),\, \mathbf{p}_\varepsilon(1))_{L^2(\Omega)^2} &= (\mathbf{y}_{\varepsilon,2}(1),\, \mathbf{p}_{\varepsilon,1}(1))_{L^2(\Omega)} + (\Delta\mathbf{y}_{\varepsilon,1}(1),\, \mathbf{p}_{\varepsilon,2}(1))_{L^2(\Omega)}\\ &= \frac{1}{\varepsilon}(\mathbf{y}_{\varepsilon,2}(1),\, \mathbf{y}_{\varepsilon,1}(1) - z_1)_{L^2(\Omega)} - \frac{1}{\varepsilon}(\mathbf{y}_{\varepsilon,1}(1),\, \mathbf{y}_{\varepsilon,2}(1) - \tau_\varepsilon z_2)_{L^2(\Omega)}\\ &= -\frac{1}{\varepsilon}(\mathbf{y}_{\varepsilon,2}(1),\, z_1)_{L^2(\Omega)} + \frac{1}{\varepsilon}(\mathbf{y}_{\varepsilon,1}(1),\, \tau_\varepsilon z_2)_{L^2(\Omega)}\\ &= -\frac{1}{\varepsilon}(\mathbf{y}_{\varepsilon,2}(1) - \tau_\varepsilon z_2,\, z_1)_{L^2(\Omega)} + \frac{1}{\varepsilon}(\mathbf{y}_{\varepsilon,1}(1) - z_1,\, \tau_\varepsilon z_2)_{L^2(\Omega)}.\end{aligned}$$

Again by the definition of $\mathbf{p}_\varepsilon(1)$, we have

$$(\mathbf{A}\mathbf{y}_\varepsilon(1),\, \mathbf{p}_\varepsilon(1))_{L^2(\Omega)^2} = \langle \Delta\mathbf{p}_{\varepsilon,2}(1),\, z_1\rangle_{H^{-1}(\Omega),H^1_0(\Omega)} + (\mathbf{p}_{\varepsilon,1}(1),\, \tau_\varepsilon z_2)_{L^2(\Omega)}.$$

If $\mathbf{z} \in \mathbf{Y}^2$, then $\mathbf{A}\mathbf{z} \in \mathbf{Y}^1 = (\mathbf{P}^0)^*$, which finishes the proof. □

Remark 3.11. The relationship between primal and adjoint variables as well as $\mathbf{z}$, expressed in equation (3.7) is remarkable in its own right. In fact, if the regularized problem (P_ε) had been defined with the terminal conditions as constraint $\mathbf{y}(1) = \Theta_\tau \mathbf{z}$, rather than as penalty, then this would again result in (3.7).

Corollary 3.12. *Let the assumptions of the Corollary 3.9 be satisfied and assume that $\mathbf{z} \in \mathbf{Y}^2$. Then the transversality condition at $t = 1$ is given by*

$$1 + \frac{\varepsilon}{2}\|u_\varepsilon(1)\|^2_{L^2(\omega)} + \langle \mathbf{A}\Theta_{\tau_\varepsilon}\mathbf{z} + \mathbf{B}u_\varepsilon(1),\ \mathbf{p}_\varepsilon(1)\rangle_{(\mathbf{P}^0)^*,\mathbf{P}^0} + (y_2, \mathbf{p}_{\varepsilon,2}(0))_{L^2(\Omega)} - (z_2, \mathbf{p}_{\varepsilon,2}(1))_{L^2(\Omega)} = 0. \tag{3.8}$$

Proof. The claim follows directly from the result of Corollary 3.9 for $t = 1$ using (3.7). □

4 Passing to the limit $\varepsilon \to 0^+$ in the optimality system

Here we are concerned with the asymptotic behavior of (3.5) as $\varepsilon \to 0^+$. For the primal equation this was addressed in Theorem 3.2. Throughout this section we will impose the regularity assumption

$$\mathbf{y}_0, \mathbf{z} \in \mathbf{Y}^2. \tag{4.1}$$

Furthermore, we will utilize the controllability assumption. For convenience, let us reformulate the observability inequality (2.1) for the transformed system on the interval I. If the original wave equation is controllable for $T > 0$ then it holds

$$\|\mathbf{p}(1)\|^2_{\mathbf{P}^0} \le C\ T\ \|\mathbf{B}^*\mathbf{p}\|^2_{L^2(I;L^2(\omega))}$$

for all functions $\mathbf{p}$ satisfying the adjoint equation $-\mathbf{p}_t = T\mathbf{A}^*\mathbf{p}$, where the constant C is given by (2.1).

We start with two lemmata.

Lemma 4.1. *For each $t \in \bar{I}$ we have*

$$\begin{aligned}(u_\varepsilon(t), \mathbf{B}^*\mathbf{p}_\varepsilon(t))_{L^2(\omega)} &= \max\left(-\gamma\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)},\ -\frac{1}{\varepsilon}\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|^2_{L^2(\omega)}\right)\\ &= -\varepsilon\|u_\varepsilon(t)\|^2_{L^2(\omega)} - \gamma\|\varepsilon u_\varepsilon(t) + \mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)}\end{aligned} \tag{4.2}$$

and

$$(4.3)\quad (u_\varepsilon,\, \mathbf{B}^*\mathbf{p}_\varepsilon)_{L^2(I;L^2(\omega))} = -\varepsilon\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} - \gamma\|\varepsilon u_\varepsilon + \mathbf{B}^*\mathbf{p}_\varepsilon\|_{L^1(I;L^2(\omega))}.$$

Proof. Inactive case $\frac{1}{\varepsilon}\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)} \le \gamma$: Then $u_\varepsilon(t) = -\frac{1}{\varepsilon}\mathbf{B}^*\mathbf{p}_\varepsilon(t)$, and

$$(u_\varepsilon(t), \mathbf{B}^*\mathbf{p}_\varepsilon(t))_{L^2(\omega)} = -\frac{1}{\varepsilon}\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|^2_{L^2(\omega)} \ge -\gamma\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)}.$$

Since $(u_\varepsilon(t), \mathbf{B}^*\mathbf{p}_\varepsilon(t))_{L^2(\omega)} = -\varepsilon\|u_\varepsilon(t)\|^2_{L^2(\omega)}$ and $\varepsilon u_\varepsilon(t) + \mathbf{B}^*\mathbf{p}_\varepsilon(t) = 0$, we find (4.2) for the inactive case.

Active case $\frac{1}{\varepsilon}\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)} > \gamma$: Here $u_\varepsilon(t) = -\frac{\gamma}{\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)}}\mathbf{B}^*\mathbf{p}_\varepsilon(t)$, which gives

$$(u_\varepsilon(t), \mathbf{B}^*\mathbf{p}_\varepsilon(t))_{L^2(\omega)} = -\gamma\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)} \ge -\frac{1}{\varepsilon}\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|^2_{L^2(\omega)}.$$

Moreover

$$\begin{aligned}\varepsilon\|u_\varepsilon(t)\|^2_{L^2(\omega)} &+ \gamma\|\varepsilon u_\varepsilon(t) + \mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)} \\ &= \varepsilon\gamma^2 + \gamma\left(1 - \frac{\varepsilon\gamma}{\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)}}\right)\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)} = \gamma\|\mathbf{B}^*\mathbf{p}_\varepsilon(t)\|_{L^2(\omega)},\end{aligned}$$

which proves the second equality in (4.2). Equality (4.3) follows from (4.2). □

Lemma 4.2. *Let* $\mathbf{z} \in \mathbf{Y}^2$. *Then there exists* $\bar\varepsilon > 0$ *and* $\delta > 0$ *such that* $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0} \ge \delta$ *for all* $\varepsilon \in (0, \bar\varepsilon]$,

Proof. If the claim was false, then there exists a sequence $\{\varepsilon_n\}_{n=1}^\infty$ with $\lim_{n\to\infty} \varepsilon_n = 0$ and $\lim_{n\to\infty} \|\mathbf{p}_{\varepsilon_n}(1)\|_{\mathbf{P}^0} = 0$. By Theorem 1.1, moreover, $\lim_{n\to\infty} \mathbf{p}_{\varepsilon_n} = 0$ in $C(0,T;\mathbf{P}^0)$. From (3.8) we have

$$\begin{aligned}1 &+ \frac{\varepsilon_n}{2}\|u_{\varepsilon_n}(1)\|^2_{L^2(\omega)} + (\Delta z_1,\, \mathbf{p}_{\varepsilon_n,2}(1))_{L^2(\Omega)} + \langle \mathbf{p}_{\varepsilon_n,1}(1),\, \tau_{\varepsilon_n} z_2\rangle_{H^{-1},H^1} \\ &+ (\chi_\omega u_{\varepsilon_n}(1),\, \mathbf{p}_{\varepsilon_n,2}(1))_{L^2(\Omega)} + (y_2, \mathbf{p}_{\varepsilon_n,2}(0))_{L^2(\Omega)} - (z_2, \mathbf{p}_{\varepsilon_n,2}(1))_{L^2(\Omega)} = 0.\end{aligned}$$

Since $\mathbf{z} \in \mathbf{Y}^2$ by (4.1) all addends tend to zero except for the first one. This is impossible. □

In Corollary 3.12 the requirement that $z \in \mathbf{Y}^2$ could have been avoided by replacing $\langle \mathbf{A}\Theta_{\tau_\varepsilon}\mathbf{z},\, \mathbf{p}_\varepsilon(1)\rangle_{(\mathbf{P}^0)^*,\mathbf{P}^0}$ by $\langle \Theta_{\tau_\varepsilon}\mathbf{z},\, \mathbf{A}^*\mathbf{p}_\varepsilon(1)\rangle_{L^2(\Omega)^2}$. For the proof of Lemma 4.2, however, it was essential to pass the operator $\mathbf{A}$ on $\mathbf{z}$.

In the following theorem we consider, according to Theorem 3.2, a weakly-star convergent subsequence of $\{(\tau_\varepsilon, y_\varepsilon, u_\varepsilon)\}_{\varepsilon>0}$, denoted by the same symbol with

$$(4.4)\quad (\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon) \rightharpoonup^* (\tau^*, \tilde{\mathbf{y}}, \tilde{u}) \text{ in } \mathbb{R}\times(L^\infty(I;\mathbf{Y}^1)\cap H^1(I;\mathbf{Y}^0))\times L^2(I;L^2(\omega)),$$

as $\varepsilon \to 0^*$.

Theorem 4.3. *Suppose that* (1.1) *is controllable for some* $T < \tau^*$ *and let* $\{(\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon)\}$ *denote a sequence of solutions to* (P_ε) *converging weakly* in* $\mathbb{R}\times(L^\infty(I;\, \mathbf{Y}^1)\cap H^1(I;\, \mathbf{Y}^0))\times L^2(I;\, L^2(\omega))$ *to a solution* $(\tau^*, \tilde{\mathbf{y}}, \tilde{u})$ *of* $(\tilde{P})$ *as* $\varepsilon \to 0^+$.

Then, if $\tilde{u}$ *is bang-bang or if the sequence* $\{\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}\}_{\varepsilon\in(0,1]}$ *is bounded in* $\mathbf{P}^1$*, there exists a non-trivial* $\tilde{\mathbf{p}} \in C(I;\mathbf{P}^0)$ *such that*

$$(4.5)\qquad \begin{cases} \tilde{\mathbf{y}}_t = \tau^*\mathbf{A}\tilde{\mathbf{y}} + \tau^*\mathbf{B}\,\tilde{u},\ \ \tilde{\mathbf{y}}(0) = \Theta_{\tau^*}\mathbf{y}_0,\ \ \tilde{\mathbf{y}}(1) = \Theta_{\tau^*}\mathbf{z}, \\ -\,\tilde{\mathbf{p}}_t = \tau^*\mathbf{A}^*\tilde{\mathbf{p}}, \\ \tau^*(\mathbf{B}^*\,\tilde{\mathbf{p}},\, u-\tilde{u})_{L^2(I;L^2(\omega))} \geq 0 \text{ for all } u \in U_{ad}. \end{cases}$$

Proof. The first equation in (4.5) follows from (4.4) and (3.5). The terminal condition for $\tilde{y}$ was proven in Theorem 3.2.

For $\varepsilon \in [0,\bar{\varepsilon}]$, with $\bar{\varepsilon}$ defined in Lemma 4.2, consider

$$(4.6)\qquad \tilde{\mathbf{p}}_\varepsilon(1) = \frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}.$$

This sequence contains a weakly convergent subsequence in $\mathbf{P}^0$ with limit $\tilde{\mathbf{p}}(1) \in \mathbf{P}^0$, where we denote the subsequence by the same symbol. Let $\tilde{\mathbf{p}}_\varepsilon$ and $\tilde{\mathbf{p}} \in C(I;\mathbf{P}^0)$ denote the solutions of the adjoint equations with terminal conditions $\tilde{\mathbf{p}}_\varepsilon(1)$ and $\tilde{\mathbf{p}}(1)$, respectively. It follows that $\tilde{\mathbf{p}}_\varepsilon \rightharpoonup^* \tilde{\mathbf{p}}$ weakly star in $L^\infty(I;\mathbf{P}^0)$. Let us show that the weak limit $\tilde{\mathbf{p}}$ satisfies (4.5).

We divide the second and third equations of (3.5) by $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}$ and pass to the limit $\varepsilon \to 0^+$ to obtain the second equation of (4.5).

To prove the third equation of (4.5), we have to show the convergence $(\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon, u_\varepsilon)_{L^2(I;L^2(\omega))} \to (\mathbf{B}^*\tilde{\mathbf{p}}, \tilde{u})_{L^2(I;L^2(\omega))}$. If on one hand $\tilde{u}$ is bang-bang, we have by Theorem 3.2 the strong convergence $u_\varepsilon \to \tilde{u}$ in $L^2(I;L^2(\omega))$, which proves $(\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon, u_\varepsilon)_{L^2(I;L^2(\omega))} \to (\mathbf{B}^*\tilde{\mathbf{p}}, \tilde{u})_{L^2(I;L^2(\omega))}$. If on the other hand the sequence $\{\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}\}$ is bounded in $\mathbf{P}^1$ then we can argue as follows. We have weak convergence of another subsequence $\tilde{\mathbf{p}}_\varepsilon(1) \rightharpoonup \tilde{\mathbf{p}}(1)$ in $\mathbf{P}^1$, which implies weak* convergence of $\tilde{\mathbf{p}}_\varepsilon \rightharpoonup^* \tilde{\mathbf{p}}$ in $L^\infty(I;\mathbf{P}^1) \cap H^1(I;\mathbf{P}^0)$ to $\tilde{\mathbf{p}} \in C(I;\mathbf{P}^1) \cap H^1(I;\mathbf{P}^0)$. Arzela-Ascoli's theorem and the compact embedding $\mathbf{P}^1$ in $\mathbf{P}^0$ imply that $\mathbf{p}_\varepsilon \to \tilde{\mathbf{p}}$ in $C(I;\mathbf{P}^0)$ as $\varepsilon \to 0$. Hence, in both cases we have for a subsequence

$$(\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon, u_\varepsilon)_{L^2(I;L^2(\omega))} \to (\mathbf{B}^*\tilde{\mathbf{p}}, \tilde{u})_{L^2(I;L^2(\omega))}. \tag{4.7}$$

This allows us to pass to the limit in the fourth equation of (3.5) to obtain the third equation of (4.5).

It remains to argue that $\tilde{\mathbf{p}}(1)$, or equivalently that $\tilde{\mathbf{p}}$ is nontrivial. By Lemma (4.1) we have

$$(u_\varepsilon, \mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon)_{L^2(I;L^2(\omega))} = -\varepsilon\sigma_\varepsilon \|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} - \gamma\|\varepsilon\sigma_\varepsilon u_\varepsilon + \mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|_{L^1(I;L^2(\omega))},$$

where $\sigma_\varepsilon = \frac{1}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}$, which is is bounded due to Lemma 4.2. This implies that

$$\begin{aligned}
\gamma\|\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|_{L^1(I;L^2(\omega))} &= \gamma\|-\varepsilon\sigma_\varepsilon u_\varepsilon + \varepsilon\sigma_\varepsilon u_\varepsilon + \mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|_{L^1(I;L^2(\omega))} \\
&\le \gamma\varepsilon\sigma_\varepsilon\|u_\varepsilon\|_{L^1(I;L^2(\omega))} + \gamma\|\varepsilon\sigma_\varepsilon u_\varepsilon + \mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|_{L^1(I;L^2(\omega))} \\
&= \gamma\varepsilon\sigma_\varepsilon\|u_\varepsilon\|_{L^1(I;L^2(\omega))} - \varepsilon\sigma_\varepsilon\|u_\varepsilon\|^2_{L^2(I;L^2(\omega))} - (u_\varepsilon, \mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon)_{L^2(I;L^2(\omega))}.
\end{aligned}$$

Due to controllability with $T < \tau^*$ there exists $\hat{\varepsilon} \in (0, \bar{\varepsilon}]$ such that for $\varepsilon \in (0, \hat{\varepsilon}]$ the system is controllable for all τ_ε, and there exists a constant $\bar{c}$ such that

$$1 = \|\tilde{\mathbf{p}}_\varepsilon(1)\|^2_{\mathbf{P}^0} \le \bar{c}\|\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|^2_{L^2(I;L^2(\omega))} \le \bar{c}\|\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|_{L^1(I;L^2(\omega))}\|\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|_{L^\infty(I;L^2(\omega))}.$$

Combining the last two estimates we find for $\varepsilon \in (0, \hat{\varepsilon}]$

$$1 \le \frac{\bar{c}}{\gamma}\|\mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon\|_{L^\infty(I;L^2(\omega))}\left[\gamma\varepsilon\sigma_\varepsilon\|u_\varepsilon\|_{L^1(I;L^2(\omega))} - \varepsilon\sigma_\varepsilon\|u_\varepsilon\|_{L^2(I;L^2(\omega))} - (u_\varepsilon, \mathbf{B}^*\tilde{\mathbf{p}}_\varepsilon)_{L^2(I;L^2(\omega))}\right].$$

The first and the second addend in the brackets on the right hand side tend to zero for $\varepsilon \to 0^+$, the third addend converges according to (4.7). If the weak subsequential limit of $\tilde{\mathbf{p}}_\varepsilon$ in $L^2(I;L^2(\omega))$ was zero this contradicts the above inequality and concludes the proof. □

Theorem 4.3 asserts that the optimality system (4.5) is qualified, i.e. $\tilde{\mathbf{p}} \neq 0$. Moreover $u^*(t) = -\gamma \frac{\mathbf{B}^* \tilde{\mathbf{p}}(t)}{\|\mathbf{B}^* \tilde{\mathbf{p}}(t)\|_{L^2(\omega)}}$ for $\mathbf{B}^* \tilde{\mathbf{p}}(t) \neq 0$.

In case $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon \in (0,\bar{\varepsilon}]}$ is bounded in $\mathbf{P}^0$, there is no need for the normalizing step in (4.6).

Corollary 4.4. *Suppose that* (1.1) *is controllable for some* $T < \tau^*$ *and let* $\{(\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon)\}$ *denote a sequence of solutions to* (P_ε) *converging weakly* in* $\mathbb{R} \times (L^\infty(I;\, \mathbf{Y}^1) \cap H^1(I;\, \mathbf{Y}^0)) \times L^2(I;\, L^2(\omega))$ *to a solution* $(\tau^*, \tilde{\mathbf{y}}, \tilde{u})$ *of* $(\tilde{P})$ *as* $\varepsilon \to 0^+$.

Let us assume that the sequence $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon \in (0,1]}$ *is bounded in* $\mathbf{P}^1$. *Then for a subsequence we have that* $\mathbf{p}_\varepsilon \to \tilde{\mathbf{p}}$ *strongly in* $C(I; \mathbf{P}^0)$, $\mathbf{p}_\varepsilon(1) \rightharpoonup \tilde{\mathbf{p}}(1)$ *weakly in* $\mathbf{P}^1$, *with* $\tilde{\mathbf{p}} \neq 0$ *satisfying* (4.5).

Proof. Using a compactness argument as in the proof of the previous theorem, we have $\mathbf{p}_\varepsilon \to \tilde{\mathbf{p}}$ in $C(I; \mathbf{P}^0)$ for a subsequence. Passing to the limit in (3.5) gives (4.5). To argue that $\tilde{\mathbf{p}}$ is qualified we proceed as in the proof of Theorem 4.3 and, using Lemma 4.2, we obtain

$$0 < \delta^2 \leq \|\mathbf{p}_\varepsilon(1)\|^2_{\mathbf{P}^0} \leq \frac{\bar{c}}{\gamma} \|\mathbf{B}^* \, \mathbf{p}_\varepsilon\|_{L^\infty(I;\, L^2(\omega))} \Big(\gamma \varepsilon \|u_\varepsilon\|_{L^\infty(I;\, L^2(\omega))} \\ - \varepsilon \|u_\varepsilon\|_{L^\infty(I;\, L^2(\omega))} - (u_\varepsilon, \mathbf{B}^* \mathbf{p}_\varepsilon)_{L^\infty(I;\, L^2(\omega))} \Big).$$

Hence it is impossible that $\mathbf{p}_\varepsilon$ converges weakly to zero, which proves that $\tilde{\mathbf{p}}$ is qualified. ☐

Remark 4.5. A similar result can be proven if $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon \in (0,1]}$ is bounded in $\mathbf{P}^0$ and $\tilde{u}$ is bang-bang. Then for a subsequence $\mathbf{p}_\varepsilon \rightharpoonup^* \tilde{\mathbf{p}}$ weakly-star in $L^\infty(I; \mathbf{P}^0)$, $\mathbf{p}_\varepsilon(1) \rightharpoonup \tilde{\mathbf{p}}(1)$ weakly in $\mathbf{P}^0$, with $\tilde{\mathbf{p}} \neq 0$ satisfying (4.5).

Remark 4.6. Since $\mathbf{y}_0 \in \mathbf{Y}^2$ in case $y_2 = 0$ and $\mathbf{z} = 0$ the pointwise transversality condition can be used to show that $\{\mathbf{B}^* \mathbf{p}_\varepsilon(1)\}$ is bounded in $L^2(\omega)$. In fact the transversality condition (3.8) implies that

$$1 + \frac{\varepsilon}{2} \|u_\varepsilon(1)\|^2_{L^2(\omega)} + (u_\varepsilon(1),\, \mathbf{B}^* \mathbf{p}_\varepsilon(1))_{L^2(\omega)} = 0.$$

By Lemma 4.2 we obtain

$$\begin{aligned} 1 + \frac{\varepsilon}{2} \|u_\varepsilon(1)\|^2_{L^2(\omega)} &= -(u_\varepsilon(1),\, \mathbf{B}^* \mathbf{p}_\varepsilon(1))_{L^2(\omega)} \\ &= \varepsilon \|u_\varepsilon(1)\|^2_{L^2(\omega)} + \gamma \|\varepsilon u_\varepsilon(1) + \mathbf{B}^* \mathbf{p}_\varepsilon(1)\|_{L^2(\omega)} \\ &\geq \varepsilon \|u_\varepsilon(1)\|^2_{L^2(\omega)} + \gamma \|\mathbf{B}^* \mathbf{p}_\varepsilon(1)\|_{L^2(\omega)} - \gamma \|\varepsilon u_\varepsilon(1)\|_{L^2(\omega)}, \end{aligned}$$

and the boundedness of $\{\mathbf{B}^*\mathbf{p}_\varepsilon(1)\}_{\varepsilon>0}$ in $L^2(\omega)$ follows. In case $\omega = \Omega$ it follows that $\{\mathbf{p}_{\varepsilon,2}(1)\}_{\varepsilon>0}$ is bounded in $L^2(\Omega)$. See also [1], where this is used in the context of time-optimal control of parabolic equations. For the hyperbolic equation considered here, we do not know whether $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon>0}$ is bounded in $\mathbf{P}^0$. Possibly this could be achieved by also applying an independent control to the y_t component of the state which is not valid in practice, however.

In the final result of this section we address transversality of the original problem $(\tilde{P})$. Here we are confronted with the difficulty of an inherent regularity gap. In view of the observability inequality (2.1) the (normalized) family of adjoint states has to be considered in $\mathbf{P}^0$, as we did, to guarantee that $\tilde{\mathbf{p}}$ is nontrivial. This allows only $C(I;\,\mathbf{P}^0)$ regularity of $\tilde{\mathbf{p}}$. Since $\tilde{\mathbf{y}} \in C(I;\,\mathbf{Y}^1)$ one cannot pass to the limit in the integral form of the transversality condition as given in the last equation of (3.5). One can resort to the pointwise transversality condition at $t = 1$, and exploit the regularity assumption $\mathbf{z} \in \mathbf{Y}^2$.

But to pass to the limit in the product $(u_\varepsilon(1), \mathbf{B}^*\mathbf{p}_\varepsilon(1))$ requires that $u_\varepsilon(1) \to \tilde{u}(1)$ in $L^2(\Omega)$, which appears to necessitate additional assumptions.

Theorem 4.7. *Suppose that* (1.1) *is controllable for some* $T < \tau^*$ *and let* $\{(\tau_\varepsilon, \mathbf{y}_\varepsilon, u_\varepsilon)\}$ *denote a sequence of solutions to* (P_ε) *converging weakly* in* $\mathbb{R} \times (L^\infty(I;\,\mathbf{Y}^1) \cap H^1(I;\,\mathbf{Y}^0)) \times L^2(I;\,L^2(\omega))$ *to a solution* $(\tau^*, \tilde{\mathbf{y}}, \tilde{u})$ *of* $(\tilde{P})$ *as* $\varepsilon \to 0^+$. *Let us assume that* $\tilde{u}$ *is continuous from the left at* $t = 1$.

If $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon\in(0,1]}$ *is bounded in* $\mathbf{P}^1$, *then there exists a nontrivial* $\tilde{\mathbf{p}} \in C(I;\,\mathbf{P}^1)$ *such that* (4.5) *holds and*

$$(4.8)\quad 1 + \langle \mathbf{A}\tilde{\mathbf{y}}(1) + \mathbf{B}\tilde{u}(1),\, \tilde{\mathbf{p}}(1)\rangle_{(\mathbf{P}^0)^*,\mathbf{P}^0} + (y_2, \tilde{\mathbf{p}}_2(0))_{L^2(\Omega)} - (z_2, \tilde{\mathbf{p}}_2(1))_{L^2(\Omega)} = 0.$$

If $\{\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}\}_{\varepsilon\in(0,1]}$ *is bounded in* $\mathbf{P}^1$ *and* $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0} \to \infty$ *for* $\varepsilon \to 0^+$, *then there exists a nontrivial* $\tilde{\mathbf{p}} \in C(I;\,\mathbf{P}^1) \cap H^1(I;\,\mathbf{P}^0)$ *such that* (4.5) *holds and*

$$(4.9)\quad \langle \mathbf{A}\tilde{\mathbf{y}}(1) + \mathbf{B}\tilde{u}(1),\, \tilde{\mathbf{p}}(1)\rangle_{(\mathbf{P}^0)^*,\mathbf{P}^0} + (y_2, \tilde{\mathbf{p}}_2(0))_{L^2(\Omega)} - (z_2, \tilde{\mathbf{p}}_2(1))_{L^2(\Omega)} = 0.$$

Proof. Let $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon\in(0,1]}$ be bounded in $\mathbf{P}^1$. Then by Corollary 4.4 there exists a subsequence converging strongly in $C(I;\mathbf{P}^0)$ to $\tilde{\mathbf{p}}$, where $\tilde{\mathbf{p}} \neq 0$ satisfies (4.5). Let us denote this subsequence by $\mathbf{p}_\varepsilon$ again.

Next we aim at passing to the limit in (3.8). First note that we have by strong convergence $\mathbf{B}^*\mathbf{p}_\varepsilon(1) \to \mathbf{B}^*\tilde{\mathbf{p}}(1)$ in $L^2(\omega)$.

If on one hand $\mathbf{B}^*\tilde{\mathbf{p}}(1) = 0$ then we have $(u_\varepsilon(1), \mathbf{B}^*\mathbf{p}_\varepsilon(1)) \to 0 = (\tilde{u}(1), \mathbf{B}^*\tilde{\mathbf{p}}(1))$. If on the other hand $\mathbf{B}^*\tilde{\mathbf{p}}(1) \neq 0$ then there is ε' such that $\|\mathbf{B}^*\mathbf{p}_\varepsilon(1)\|_{L^2(\omega)} > \gamma\varepsilon$ for all $\varepsilon < \varepsilon'$, which implies that $u_\varepsilon(1) = \gamma \frac{\mathbf{B}^*\mathbf{p}_\varepsilon(1)}{\|\mathbf{B}^*\mathbf{p}_\varepsilon(1)\|_{L^2(\omega)}}$ for all $\varepsilon < \varepsilon'$. Hence $u_\varepsilon(1) = \gamma \frac{\mathbf{B}^*\mathbf{p}_\varepsilon(1)}{\|\mathbf{B}^*\mathbf{p}_\varepsilon(1)\|_{L^2(\omega)}} \to \gamma \frac{\mathbf{B}^*\tilde{\mathbf{p}}(1)}{\|\mathbf{B}^*\tilde{\mathbf{p}}(1)\|_{L^2(\omega)}}$ strongly in $L^2(\omega)$. By continuity of $\tilde{u}$ at $t = 1$ we obtain $\tilde{u}(1) = \gamma \frac{\mathbf{B}^*\tilde{\mathbf{p}}(1)}{\|\mathbf{B}^*\tilde{\mathbf{p}}(1)\|_{L^2(\omega)}}$. Hence, it holds $(u_\varepsilon(1), \mathbf{B}^*\mathbf{p}_\varepsilon(1)) \to (\tilde{u}(1), \mathbf{B}^*\tilde{\mathbf{p}}(1))$.

Since $\mathbf{p}_\varepsilon(1) \rightharpoonup \tilde{\mathbf{p}}(1)$ in $\mathbf{P}^1$ and $\mathbf{z} \in \mathbf{Y}^1$, we can pass to the limit in (3.8) and obtain

$$1 + \langle \mathbf{A}\Theta_{\tau^*}\mathbf{z} + \mathbf{B}\tilde{u}(1),\, \tilde{\mathbf{p}}(1)\rangle_{(\mathbf{P}^0)^*,\mathbf{P}^0} + (y_2, \tilde{\mathbf{p}}_2(0))_{L^2(\Omega)} - (z_2, \tilde{\mathbf{p}}_2(1))_{L^2(\Omega)} = 0.$$

Since $\tilde{\mathbf{y}}(1) = \Theta_{\tau^*}\mathbf{z}$ by Theorem 3.2, we obtain (4.8).

The case that $\{\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}\}_{\varepsilon\in(0,1]}$ is bounded in $\mathbf{P}^1$ with $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0} \to \infty$ can be treated analogously, only that this time we apply the normalizing step (4.6) as in the proof of Theorem 4.3. In particular, one needs to divide (3.8) by $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}$ before passing to the limit $\varepsilon \to 0^+$. Moreover, by Theorem 4.3 we obtain that (4.5) is fulfilled. □

In order to be able to pass to the limit in the integrated form of the transversality condition we additionally assume that $\tilde{u}$ is bang-bang.

Corollary 4.8. *Let the assumptions of the previous Theorem 4.7 be fulfilled. In addition, let us assume that $\tilde{u}$ is bang-bang.*

If $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon\in(0,1]}$ is bounded in $\mathbf{P}^1$, then $\tilde{\mathbf{p}}$ given by Theorem 4.7 fulfills

$$1 + \langle \mathbf{A}\tilde{\mathbf{y}} + \mathbf{B}\tilde{u},\, \tilde{\mathbf{p}}\rangle_{L^2(I;(\mathbf{P}^1)^*),L^2(I;\mathbf{P}^1)} + (y_2, \tilde{\mathbf{p}}_2(0))_{L^2(\Omega)} - (z_2, \tilde{\mathbf{p}}_2(1))_{L^2(\Omega)} = 0.$$

If $\{\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}\}_{\varepsilon\in(0,1]}$ is bounded in $\mathbf{P}^1$ and $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0} \to \infty$ for $\varepsilon \to 0^+$, then $\tilde{\mathbf{p}}$ given by Theorem 4.7 fulfills

$$\langle \mathbf{A}\tilde{\mathbf{y}} + \mathbf{B}\tilde{u},\, \tilde{\mathbf{p}}\rangle_{L^2(I;(\mathbf{P}^1)^*),L^2(I;\mathbf{P}^1)} + (y_2, \tilde{\mathbf{p}}_2(0))_{L^2(\Omega)} - (z_2, \tilde{\mathbf{p}}_2(1))_{L^2(\Omega)} = 0.$$

Proof. If $\tilde{u}$ is bang-bang, then $u_\varepsilon \to \tilde{u}$ strongly in $L^2(I; L^2(\omega))$ by Theorem 3.2 in addition to (4.4). Moreover, we have $\mathbf{y}_\varepsilon \to \tilde{\mathbf{y}}$ strongly in $C(I; \mathbf{Y}^1)$. Hence, it holds $\mathbf{A}\mathbf{y}_\varepsilon + \mathbf{B}u_\varepsilon \to \mathbf{A}\tilde{\mathbf{y}} + \mathbf{B}\tilde{u}$ strongly in $L^2(I; \mathbf{Y}^0) = L^2(I; (\mathbf{P}^1)^*)$.

If $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon\in(0,1]}$ is bounded in $\mathbf{P}^1$, then we have $\mathbf{p}_\varepsilon \rightharpoonup \tilde{\mathbf{p}}$ in $L^2(I;\mathbf{P}^1)$, where $\tilde{\mathbf{p}}$ satisfies (4.5). Hence, we can pass to the limit in the last equation of (3.5) to obtain the claim in the first case.

If $\{\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}\}_{\varepsilon\in(0,1]}$ is bounded in $\mathbf{P}^1$, then the scaled sequence $\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}$ is bounded in $C(I;\mathbf{P}^1)$. Dividing the last equation of (3.5) by $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}$, gives the claim in the second case. □

Remark 4.9. If $|\tau_\varepsilon - \tau^*| = O(\varepsilon)$, then due to Corollary 3.3 and the relation $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^1} = \frac{1}{\varepsilon}\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon}\mathbf{z}\|_{\mathbf{Y}^0}$ the sequence $\{\mathbf{p}_\varepsilon(1)\}$ is bounded in $\mathbf{P}^1$.

Remark 4.10. In case $\omega = \Omega$ and $\{\mathbf{p}_\varepsilon(1)\}_{\varepsilon\in(0,1]}$ bounded in $\mathbf{P}^1$ we have

$$\tilde{u}(t) = \gamma\frac{\tilde{\mathbf{p}}_2(t)}{\|\tilde{\mathbf{p}}_2(t)\|_{L^2(\Omega)}} \text{ for } t \text{ sufficiently close to } 1^-$$

by Lemma 4.2 and the strong convergence $\mathbf{p}_\varepsilon \to \tilde{\mathbf{p}}$ in $C(I;\mathbf{P}^0)$. Since $\tilde{\mathbf{p}} \in C(I;\mathbf{P}^1)$ the assumption that $\tilde{u}(1)$ is continuous at $t = 1$ is automatically satisfied in this case.

Remark 4.11. Let us remark that we cannot use the auxiliary sequence $\hat{\mathbf{p}}_\varepsilon := \{\frac{\mathbf{p}_\varepsilon(1)}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^1}}\}$ to prove the results of Theorems 4.3 and 4.7 in the case that $\{\mathbf{p}_\varepsilon(1)\}$ is not bounded in $\mathbf{P}^1$. While the limiting process in the first-order system can be done as in the proof of the Theorem 4.7, it seems to be impossible to prove that a subsequence of $\hat{\mathbf{p}}_\varepsilon$ converges to a non-zero limit point.

Remark 4.12. Our formulation of the regularized problems (P_ε) realizes the terminal constraint by means of penalty. We could alternatively consider

$$(4.10) \qquad \begin{cases} \min \tau(1 + \frac{\varepsilon}{2}\|u\|^2_{L^2(I;L^2(\omega))}) \\ \text{subject to } \tau \geq 0, \text{ and} \\ \mathbf{y}_t = \tau\mathbf{A}\,\mathbf{y} + \mathbf{B}\,u, \text{ on } (0,1] \\ \mathbf{y}(0) = \Theta_\tau\mathbf{y}_0,\ \mathbf{y}(1) = \Theta_\tau\mathbf{z},\ u \in U_{ad}, \end{cases}$$

where the terminal condition is kept as explicit constraint. For every $\varepsilon > 0$ problem (4.10) has a solution $(\tau_\varepsilon, y_\varepsilon, u_\varepsilon)$ which satisfies the monotonicity properties, for $0 < \varepsilon_0 < \varepsilon_1$,

$$\tau^* \leq \tau_{\varepsilon_0} \leq \tau_{\varepsilon_1} \leq \tau^*(1 + \frac{\gamma\varepsilon_1}{2})$$

and

$$\|u_{\varepsilon_1}\|_{L^2(I;L^2(\omega))} \le \|u_{\varepsilon_0}\|_{L^2(I;L^2(\omega))} \le \|u^*\|_{L^2(I;L^2(\omega))},$$

which can be verified with arguments analogous to those in [9]. Developing a Lagrangian theory for (4.10), however, is impeded by the fact that the constraints in (4.10) are not differentiable in the natural norms. Specifically we define

$$e(\tau,u) : \mathbb{R}\times\ L^2(I;L^2(\omega)) \to \mathbf{Y}^1$$

given by $e(\tau,u) = \mathbf{y}(\tau,u)(1) = (y(\tau,u)(1),\, y_t(\tau,u)(1))$. Note that e is not differentiable with respect to τ. In fact, $e_\tau(\tau,u)$ would be the solution to

$$\tilde{\mathbf{y}}_t = \tau\mathbf{A}\tilde{\mathbf{y}} + \mathbf{A}\mathbf{y} + \mathbf{B}\,u$$
$$\tilde{\mathbf{y}}(0) = \begin{pmatrix} 0 \\ y_2 \end{pmatrix}$$

evaluated at 1. Since $\mathbf{A}\mathbf{y} \in C(I;\mathbf{Y}^0)$ only, this does not guarantee that $\tilde{\mathbf{y}}(1) \in \mathbf{Y}^1$, in general.

5 A parametric optimization problem

Throughout this section we fix $\varepsilon > 0$, and we consider, for any $\tau > 0$ the minimization problem with respect to the variable u:

$$(P_\varepsilon^\tau) \qquad \begin{cases} \displaystyle\min_{u\in U_{ad}} \tau\left(1 + \frac{\varepsilon}{2}\|u\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon}\|\mathbf{y}(1) - \Theta_\tau \mathbf{z}\|^2_{Y^0} \\ \text{subject to} \\ \mathbf{y}_t = \tau\mathbf{A}\mathbf{y} + \tau\mathbf{B}u \text{ on } (0,1], \\ \mathbf{y}(0) = \Theta_\tau \mathbf{y}_0. \end{cases}$$

Clearly this problem admits a unique solution $(\mathbf{y}_\tau, u_\tau) \in (C(I;\mathbf{Y}^1)\cap C^1(I;\mathbf{Y}^0))\times L^2(I;L^2(\omega))$ and the necessary and sufficient optimality condition is given by

$$(5.1) \qquad \begin{cases} \partial_t \mathbf{y}_\tau = \tau\mathbf{A}\mathbf{y}_\tau + \tau\mathbf{B}u_\tau, \quad \mathbf{y}_\tau(0) = \Theta_\tau \mathbf{y}_0 \\ -\partial_t \mathbf{p}_\tau = \tau\mathbf{A}^*\mathbf{p}_\tau,\ \mathbf{p}_\tau(1) = \dfrac{1}{\varepsilon}\begin{pmatrix} \mathbf{y}_{\varepsilon,1}(1) - z_1 \\ (-\Delta)^{-1}(\mathbf{y}_{\varepsilon,2}(1) - \tau z_2) \end{pmatrix} \\ (\varepsilon u_\tau + \mathbf{B}^*\mathbf{p}_\tau, u - u_\tau)_{L^2(I;L^2(w))} \ge 0, \text{ for all } u \in U_{ad}. \end{cases}$$

Here we use the index τ for the solution to (P_ε^τ), since our attention focuses on the dependence of the solution on this parameter. The main objective of this section is to verify that the derivative of the value function of the cost associated to (P_ε^τ) is given by the transversality condition, which is the last equation in (3.5). Some preliminaries will be required. Let us note that these results apparently cannot be directly derived from well-known abstract sensitivity results. This is due to the fact that these results require us to express the constraint in an abstract setting, in which the partial differential equation in (P_ε^τ) is transformed to an equality constraint of the form $e(y,u) = 0$. The natural domain for e is $(L^2(I;\mathbf{Y}^1) \cap H^1(I;\mathbf{Y}^0) \cap \{\mathbf{y} : (y_2)_t - \tau\Delta y_1 \in L^2(\Omega)\}) \times L^2(I;L^2(\omega))$. This space endowed with the natural norm is a Banach space. However, the domain for e depends on the parameter τ, and the abstract results appear not to cover this case.

Theorem 5.1. *Let $\mathbf{y}_0 \in \mathbf{Y}^2$. Then for every compact subset $J \subset (0,\infty)$ the mapping $\tau \to (\mathbf{y}_\tau, u_\tau, \mathbf{p}_\tau)$ is globally Lipschitz continuous from J to $(C(I;\mathbf{Y}^2) \cap C^1(I;\mathbf{Y}^1)) \times L^2(I;L^2(\omega)) \times (C(I;\mathbf{P}^2) \cap C^1(I;\mathbf{P}^1)$.*

Proof. Let $J \subset (0,\infty)$ be compact and choose arbitrary $\overline{\tau} \in J, \tau \in J$. Further set

$$(\delta\tau, \boldsymbol{\delta}\mathbf{y}, \delta u, \boldsymbol{\delta}\mathbf{p}) = (\overline{\tau} - \tau, \mathbf{y}_{\overline{\tau}} - \mathbf{y}_\tau, u_{\overline{\tau}} - u_\tau, \mathbf{p}_{\overline{\tau}} - \mathbf{p}_\tau),$$

where $(\mathbf{y}_\tau, u_\tau, \mathbf{p}_\tau)$ is the solution to (P_ε^τ), and analogously for $(\mathbf{y}_{\overline{\tau}}, u_{\overline{\tau}}, \mathbf{p}_{\overline{\tau}})$. Then we have

(5.2)
$$\begin{cases} \partial_t\boldsymbol{\delta}\mathbf{y} = \tau(\mathbf{A}\boldsymbol{\delta}\mathbf{y} + \mathbf{B}\delta u) + \delta\tau(\mathbf{A}\mathbf{y}_{\overline{\tau}} + \mathbf{B}u_{\overline{\tau}}),\ \boldsymbol{\delta}\mathbf{y}(0) = \begin{pmatrix} 0 \\ \delta\tau y_2 \end{pmatrix}, \\[2ex] -\partial_t\boldsymbol{\delta}\mathbf{p} = \tau\mathbf{A}^*\boldsymbol{\delta}\mathbf{p} + \delta\tau\mathbf{A}^*\mathbf{p}_{\overline{\tau}},\ \boldsymbol{\delta}\mathbf{p}(1) = \dfrac{1}{\varepsilon}\begin{pmatrix} \boldsymbol{\delta}\mathbf{y}_1(1) \\ (-\Delta)^{-1}(\boldsymbol{\delta}\mathbf{y}_2(1) - \delta\tau z_2) \end{pmatrix}, \\[2ex] (\varepsilon u_\tau + \mathbf{B}^*\mathbf{p}_\tau, u - u_\tau)_{L^2(I;L^2(w))} \geq 0, \\ (\varepsilon u_{\overline{\tau}} + \mathbf{B}^*\mathbf{p}_{\overline{\tau}}, u - u_{\overline{\tau}})_{L^2(I;L^2(w))} \geq 0, \quad \text{for all } \ u \in U_{ad}. \end{cases}$$

Since $\mathbf{y}_0 \in \mathbf{Y}^2$ we have by Corollary 1.2

$$\{\mathbf{y}_\tau, \mathbf{p}_\tau\}_{\tau\in J} \text{ is bounded in } W,$$

where $W = (C(I;\mathbf{Y}^2) \cap C^1(I;\mathbf{Y}^1)) \times (C(I;\mathbf{P}^2) \cap C^1(I;\mathbf{P}^1))$. Moreover there exists a constant K_1 such that

$$(5.3) \qquad \begin{cases} \|\boldsymbol{\delta}\mathbf{y}\|_{C(I;\mathbf{Y}^2)\cap C^1(I;\mathbf{Y}^1)} \le K_1(|\delta\tau| + |\delta u|_{L^2(I;L^2(\omega))}) \\ \|\boldsymbol{\delta}\mathbf{p}\|_{C(I;\mathbf{P}^1)\cap C^1(I;\mathbf{P}^0)} \le K_1(|\delta\tau| + \|\boldsymbol{\delta}\mathbf{y}(1)\|_{\mathbf{Y}^0}). \end{cases}$$

Let us note here that the scalar form of the first equation in (5.2) is given by

$$\begin{cases} \partial_{tt}\boldsymbol{\delta}\mathbf{y}_1 = \tau(\Delta\boldsymbol{\delta}\mathbf{y}_1 + \delta u) + \tau(\Delta\mathbf{y}_{\overline{\tau},1} + u_{\overline{\tau}})\delta\tau + \delta\tau\partial_t\mathbf{y}_{\overline{\tau},2}. \\ \boldsymbol{\delta}\mathbf{y}_1(0) = 0,\ \partial_t\boldsymbol{\delta}\mathbf{y}_1(0) = \delta\tau\, y_2. \end{cases}$$

Setting $u = u_{\overline{\tau}}$ and $u = u_\tau$ in the two inequalities of (5.2) we obtain, after rearranging terms,

$$(5.4) \qquad \varepsilon\|\delta u\|^2_{L^2(I;L^2(\omega))} \le -(\mathbf{B}^*\boldsymbol{\delta}\mathbf{p}, \delta u)_{L^2(I;L^2(\omega))}.$$

Next we take the inner product in $L^2(I;L^2(\Omega))$ of the first equation in (5.2) with $\boldsymbol{\delta}\mathbf{p}$ and of the second equation with $\boldsymbol{\delta}\mathbf{y}$. After integration by parts and subtraction of the two resulting equations we obtain

$$(\boldsymbol{\delta}\mathbf{y}(1), \boldsymbol{\delta}\mathbf{p}(1)) - (\boldsymbol{\delta}\mathbf{y}(0), \boldsymbol{\delta}\mathbf{p}(0)) = \tau(\mathbf{B}\,\delta u, \boldsymbol{\delta}\mathbf{p}) + \delta\tau\big((\mathbf{A}\mathbf{y}_{\overline{\tau}} + \mathbf{B}u_{\overline{\tau}}, \boldsymbol{\delta}\mathbf{p}) - (\mathbf{A}^*\mathbf{p}_{\overline{\tau}}, \boldsymbol{\delta}\mathbf{y})\big),$$

where all inner products are taken in $L^2(I;L^2(\Omega))$. The initial condition for $\boldsymbol{\delta}\mathbf{p}(1)$ and (5.4) imply that

$$\begin{aligned} &\frac{1}{\varepsilon}\|\boldsymbol{\delta}\mathbf{y}(1)\|^2_{\mathbf{Y}^0} + \tau\varepsilon\|\delta u\|_{L^2(I;L^2(\omega))} \\ &\quad \le \frac{|\delta\tau|}{\varepsilon}\|z_2\|_{H^{-1}}\|\boldsymbol{\delta}\mathbf{y}_2(1)\|_{H^{-1}} + K_1|\delta\tau|\,\|y_2\|_{H^1}\,(|\delta\tau| + \|\boldsymbol{\delta}\mathbf{y}(1)\|_{\mathbf{Y}^0}) \\ &\qquad + K_1\tau\|\delta u\|_{L^2(I;L^2(\omega))}(|\delta\tau| + \|\boldsymbol{\delta}\mathbf{y}(1)\|_{Y^0}) \\ &\qquad + K_2|\delta\tau|(|\delta\tau| + \|\boldsymbol{\delta}\mathbf{y}(1)\|_{\mathbf{Y}^0}) + K_2|\delta\tau|(|\delta\tau| + \|\delta u\|_{L^2(I;L^2(\omega))}), \end{aligned}$$

where K_2 is a constant independent of $\tau_1, \tau_2 \in J$. By Young's inequality we find the existence of a constant K_3, independent of $\tau, \overline{\tau} \in J$, but depending on ε, such that

$$\|\boldsymbol{\delta}\mathbf{y}(1)\|_{\mathbf{Y}^0} + \|\delta u\|_{L^2(I;L^2(\omega))} \le K_3|\delta\tau|.$$

From the first equation of (5.3) we deduce that

$$\|\boldsymbol{\delta}\mathbf{y}\|_{C(I;\mathbf{Y}^2)\cap C^1(I;\mathbf{Y}^1)} \le K_4|\delta\tau|,$$

and by the second equation in (5.2)

$$\|\boldsymbol{\delta}\mathbf{p}\|_{C(I;\mathbf{P}^2)\cap C^1(I;\mathbf{P}^1)} \le K_4|\delta\tau|,$$

for a constant K_4 independent of $(\tau,\overline{\tau}) \in J \times J$ but depending on ε, and $\mathbf{y}_0 \in \mathbf{Y}^2$, $\mathbf{z} \in \mathbf{Y}^1$. ☐

Theorem 5.2. *Let* $\mathbf{y}_0 \in \mathbf{Y}^2$ *and let* $(\dot{\mathbf{y}}, \dot{u}, \dot{\mathbf{p}})$ *denote a weak* cluster-point of* $\frac{1}{\sigma}((\mathbf{y}_{\tau+\sigma}, u_{\tau+\sigma}, \mathbf{p}_{\tau+\sigma}) - (\mathbf{y}_\tau, u_\tau, \mathbf{p}_\tau))$ *in* $(L^\infty(I;\mathbf{Y}^2)\cap W^{1,\infty}(I;\mathbf{Y}^1))\times L^2(I;L^2(\omega))\times (L^\infty(I;\mathbf{P}^2)\cap W^{1,\infty}(I;\mathbf{P}^1))$ *as* $\sigma \to 0^+$. *Then* $(\dot{\mathbf{y}}, \dot{u}, \dot{\mathbf{p}})$ *in* $(C(I;\mathbf{Y}^2)\cap C^1(I;\mathbf{Y}^1))\times L^2(I;L^2(\omega)) \times (C(I;\mathbf{P}^2)\cap C^1(I;\mathbf{P}^1)$ *and it satisfies the sensitivity system*

$$(5.5)\qquad \begin{cases} \partial_t \dot{\mathbf{y}} = \tau(\mathbf{A}\dot{\mathbf{y}} + \mathbf{B}\dot{u}) + \mathbf{A}\mathbf{y}_\tau + \mathbf{B}u_\tau,\ \dot{\mathbf{y}}(0) = \begin{pmatrix} 0 \\ y_2 \end{pmatrix} \\ \\ -\partial_t \dot{\mathbf{p}} = \tau\mathbf{A}^*\dot{\mathbf{p}} + \mathbf{A}\mathbf{p}_\tau,\ \dot{\mathbf{p}}(1) = \dfrac{1}{\varepsilon}\begin{pmatrix} \dot{\mathbf{y}}_1(1) \\ (-\Delta)^{-1}(\dot{\mathbf{y}}_2(1) - z_2) \end{pmatrix} \\ \\ (\varepsilon u_\tau + \mathbf{B}^*\mathbf{p}_\tau, \dot{u})_{L^2(I;L^2(\omega))} = 0. \end{cases}$$

Proof. Existence of the weak cluster point follows from Theorem 5.1. Passing to the limit in the first two equations of (5.2) we obtain the first two equations of (5.5), which hold as equalities in $L^2(I;\mathbf{Y}^1)$ and $L^2(I;\mathbf{P}^1)$ respectively. Since $\dot{\mathbf{y}}(0) \in \mathbf{Y}^2$ and $\mathbf{A}y_\tau + \mathbf{B}u_\tau \in H^1(I;\mathbf{Y}^1)$ we find moreover that $\dot{\mathbf{y}} \in C(I;\mathbf{Y}^2)\cap C^1(I;\mathbf{Y}^1)$ and analogously $\dot{\mathbf{p}} \in C(I;\mathbf{P}^2)\cap C^1(I;\mathbf{P}^1)$. Setting $u = u_{\tau+\sigma}$ in the first inequality of (5.2) we find $(\varepsilon u_\tau + \mathbf{B}^*\mathbf{p}_\tau, \dot{u})_{L^2(I;L^2(\omega))} \ge 0$. From the second inequality with u_τ replaced by $u_{\overline{\tau}}$ and $u = u_{\tau+\sigma}$ it follows that $(\varepsilon u_\tau + \mathbf{B}^*\mathbf{p}_\tau, \dot{u})_{L^2(I;L^2(\omega))} \le 0$. Combined this implies the third equation in (5.5). ☐

Let $\mathcal{V}$ denote the value functional associated to (P_ε^τ), i.e.

$$(5.6)\qquad \mathcal{V}(\tau) = \tau\left(1 + \frac{\varepsilon}{2}\|u_\tau\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon}\|\mathbf{y}_\tau(1) - \Theta_\tau \mathbf{z}\|_{\mathbf{Y}^0},$$

and set $\mathbf{L}^2(\Omega) = L^2(\Omega)\times L^2(\Omega)$. The following result establishes smoothness of $\mathcal{V}$ and the relationship of $\frac{d}{d\tau}\mathcal{V}(\tau)$ to the transversality condition.

Theorem 5.3. *Let* $\mathbf{y}_0 \in \mathbf{Y}^2$*. Then* $\tau \to \mathcal{V}(\tau)$ *is continuously differentiable on* $(0,\infty)$ *and*

$$\frac{d}{d\tau}\mathcal{V}(\tau) = 1 + \frac{\varepsilon}{2}\|u_\tau\|^2_{L^2(I;L^2(\omega))} + (\mathbf{A}\mathbf{y}_\tau + \mathbf{B}u_\tau, \mathbf{p}_\tau)_{L^2(I;\mathbf{L}^2(\Omega))}$$
$$+ (y_2, \mathbf{p}_{\tau,2}(0))_{L^2(\Omega)} - (z_2, \mathbf{p}_{\tau,2}(1))_{L^2(\Omega)}.$$

Proof. Let $\tau \in (0,\infty)$ and let $(\dot{\mathbf{y}}, \dot{u}, \dot{\mathbf{p}})$ denote a weak cluster point according to Theorem 5.2 as $\sigma \to 0^+$. To save notation it will be convenient to set

$$b = (1 + \frac{\varepsilon}{2}\|u_\tau\|^2_{L^2(I;L^2(\omega))}) + \tau\varepsilon(u_\tau, \dot{u})_{L^2(I;L^2(\omega))}.$$

We compute

$$\lim_{n\to\infty}\frac{1}{\sigma_n}(\mathcal{V}(\tau+\sigma_n) - \mathcal{V}(\tau)) = b + \frac{1}{\varepsilon}(\mathbf{y}_\tau(1) - \Theta_\tau\mathbf{z}, \dot{\mathbf{y}}(1))_{\mathbf{Y}^0} - \frac{1}{\varepsilon}(\mathbf{p}_{\tau,2}(1), z_2)_{L^2(\Omega)}$$
$$= b + (\mathbf{p}_\tau(1), \dot{\mathbf{y}}(1))_{\mathbf{L}^2(\Omega)} - (\mathbf{p}_{\tau,2}(1), z_2)_{L^2(\Omega)}$$
$$= b + \int_0^1 \frac{d}{dt}(\mathbf{p}_\tau(t), \dot{\mathbf{y}}(t))_{L^2(\Omega)} + (\mathbf{p}_\tau(0), \dot{\mathbf{y}}(0))_{L^2(\Omega)} - (\mathbf{p}_{\tau,2}(1), z_2)_{L^2(\Omega)}$$
$$= b - \tau\int_0^1 (\mathbf{A}^*\mathbf{p}_\tau(t), \dot{\mathbf{y}}(t))_{\mathbf{L}^2(\Omega)} + \int_0^1 (\mathbf{p}_\tau(t), \tau\mathbf{A}\dot{\mathbf{y}}(t) + \tau\mathbf{B}\dot{u}(t) + \mathbf{A}\mathbf{y}_\tau(t)$$
$$+ \mathbf{B}u_\tau(t))_{L^2(\Omega)} + (\mathbf{p}_{\tau,2}(0), y_2)_{L^2(\Omega)} - (\mathbf{p}_{\tau,2}(1), z_2)_{L^2(\Omega)}$$
$$= (1 + \frac{\varepsilon}{2}\|u_\tau\|^2_{(I;L^2(\Omega))}) + \tau(\varepsilon u_\tau + \mathbf{B}^*\mathbf{p}_\tau, \dot{u})_{(I;L^2(\Omega))} + \int_0^1 (\mathbf{p}_\tau(t), \mathbf{A}\mathbf{y}_\tau(t)$$
$$+ \mathbf{B}u_\tau(t))_{\mathbf{L}^2(\Omega)} + (\mathbf{p}_{\tau,2}(0), y_2)_{L^2(\Omega)} - (\mathbf{p}_{\tau,2}(1), z_2)_{L^2(\Omega)}.$$

By (5.5) we have $(\varepsilon u_\tau + \mathbf{B}^*\mathbf{p}_\tau, \dot{u})_{(I;L^2(\Omega))} = 0$. Hence $\lim_{\sigma\to 0^+}\frac{1}{\sigma}(\mathcal{V}(\tau+\sigma) - \mathcal{V}(\tau))$ exists and it is given by

$$\frac{d}{dt}\mathcal{V}(\tau) = 1 + \frac{\varepsilon}{2}\|u_\tau\|^2_{L^2(I;L^2(\omega))} + (\mathbf{p}_\tau, \mathbf{A}\mathbf{y}_\tau + \mathbf{B}u_\tau)_{L^2(I;L^2(\omega))}$$
$$+ (\mathbf{p}_{\tau,2}(0), y_2)_{L^2(\Omega)} - (\mathbf{p}_{\tau,2}(1), z_2)_{L^2(\Omega)},$$

as desired. The expression on the right hand side coincides with the expression that arises in the transversality condition. Note that due to extra regularity requirement that $\mathbf{y}_0 \in \mathbf{Y}^2$ we have that $\mathbf{A}\mathbf{y} \in L^2(I; \mathbf{L}^2(\Omega))$, so that we can avoid the duality pairing that we used in the transversality condition in (3.5). Finally, by Theorem 5.1 $\tau \to \frac{d}{dt}\mathcal{V}(\tau)$ is Lipschitz continuous. □

We conclude this section with an asymptotic estimate of $\mathcal{V}(\tau)$ as $\tau \to \infty$.

Theorem 5.4. *Let us assume that there are positive constants τ_0, C_0 such that for all $\tau > \tau_0$ there exists a control $u_{0,\tau}$ that is admissible for the original time-optimal control problem* (P) *and which satisfies*

$$\|u_{0,\tau}\|_{L^\infty(I;L^2(\omega))} \leq C_0 \tau^{-1}.$$

Then there exist constants $\varepsilon_0 > 0$ and $c > 0$ independent of τ such that for all $\tau > \tau_0$ and $\varepsilon \in (0, \varepsilon_0)$ we have

$$|\mathcal{V}(\tau) - \tau| \leq c\, \epsilon\, \tau^{-1},$$

$$\left| \frac{d}{d\tau}\mathcal{V}(\tau) - 1 \right| \leq c\, \tau^{-1/2}.$$

Proof. Let us take $\tau > \tau_0$. Then $u_{0,\tau}$ is admissible for (P_ε^τ) and hence

$$\tau \frac{\varepsilon}{2}\|u_\tau\|^2_{L^2(I;L^2(\omega))} + \frac{1}{2\varepsilon}\|\mathbf{y}_\tau(1) - \Theta_\tau \mathbf{z}\|^2_{Y^0} \leq \tau \frac{\varepsilon}{2}\|u_{0,\tau}\|^2_{L^2(I;L^2(\omega))} \leq c\, \varepsilon\, \tau^{-1}.$$

This proves the first claim.

This estimate further implies that

$$\|\mathbf{p}_\tau(1)\|_{\mathbf{P}^1} = \frac{1}{\epsilon}\|\mathbf{y}_\tau(1) - \Theta_\tau \mathbf{z}\|_{Y^0} \leq c\, \tau^{-1/2}, \tag{5.7}$$

where c is used as a generic constant, independent of τ.

Testing the adjoint equations, cf. (3.3), $-\mathbf{p}'_{\tau,1} = \tau\Delta\mathbf{p}_2$ and $-\mathbf{p}_{\tau,2} = \tau\mathbf{p}_{\tau,1}$ by $\mathbf{p}'_{\tau,2}$ and $-\mathbf{p}'_{\tau,1}$, respectively, subtracting the resulting equations, and integrating on $(t, 1)$, $t \in (0, 1)$, gives the energy equation

$$\|\mathbf{p}_{\tau,1}(t)\|^2_{L^2(\Omega)} + \|\nabla\mathbf{p}_{\tau,2}(t)\|^2_{L^2(\Omega)} = \|\mathbf{p}_{\tau,1}(1)\|^2_{L^2(\Omega)} + \|\nabla\mathbf{p}_{\tau,2}(1)\|^2_{L^2(\Omega)}$$

for all $t \in (0, 1)$ and $\tau > 0$. Together with (5.7) this proves

$$\|\mathbf{p}_\tau\|_{L^\infty(I;\mathbf{P}^1)} \leq c\, \tau^{-1/2}.$$

Analogously one obtains the estimate

$$\|\mathbf{y}_\tau\|_{L^\infty(\mathbf{Y}^1)} \leq c\left(\|y_0\|_{\mathbf{Y}^1} + \|u_\tau\|_{L^2(I;L^2(\omega))}\right),$$

with $c > 0$ independent of τ, which shows that $\mathbf{A}\mathbf{y}_\tau$ is bounded in $L^\infty(I;\mathbf{Y}^0)$ uniformly with respect to τ. This implies that

$$\sup_{t\in I} \langle \mathbf{A}\mathbf{y}_\tau(t) + \mathbf{B}u_\tau(t),\ \mathbf{p}_\tau(t)\rangle_{\mathbf{Y}^0,\mathbf{P}^1} \leq c\,\tau^{-1/2},$$

which, together with Theorem 5.3 and (5.7), proves the second claim. □

In the case of $\omega = \Omega$, the assumption of the previous theorem on the existence of controls $u_{0,\tau}$ is satisfied due to a result by Fattorini [6].

6 Semi-smooth Newton algorithm for the regularized optimality system associated to (P_ε^τ)

The algorithm for solving (P), which will be described in the following section, relies on an efficient numerical method to solve (P_ε^τ). For this purpose we use a semi-smooth Newton method. This section is is devoted to justification of the method in the sense of verifying its local superlinear convergence.

In view of (5.1) a control u, which depends on τ and ε, is a solution to (P_ε^τ), if and only if

$$\mathcal{F}u = u + P_{U_{ad}}\left(\frac{1}{\varepsilon}\chi_\omega \mathbf{p}_{\tau,2}\right) = 0, \tag{6.1}$$

where $\mathbf{p}_\tau = \mathbf{p}_\tau(u)$ is defined through the primal and adjoint equations in (5.1). Here $\mathcal{F}$ is considered as an operator from $L^2(I;L^2(\omega))$ to itself. We shall verify below that $P_{U_{ad}} : C(I;L^2(\omega)) \to L^2(I;L^2(\omega))$ is Newton differentiable [9] with Newton derivative denoted by $DP_{U_{ad}}$. Since $u \to \mathbf{y}_\tau \to \mathbf{p}_{\tau,2}$ is a continuous linear mapping from $L^2(I;L^2(\omega)) \to C(I;L^2(\omega))$, a semi-smooth Newton iteration can be applied to (6.1). Given u_0, we compute $u_{k+1} = u_k + \delta u$ by solving

$$D\mathcal{F}(u_k)\,\delta u = -\mathcal{F}(u_k), \tag{6.2}$$

where $D\mathcal{F}(u) = I + DP_{U_{ad}}$.

Next we address Newton differentiability of radial projections in $L^2(\omega)$ and $L^2(I;L^2(\omega))$.

Proposition 6.1. *The projection $P_U : L^2(\omega) \to L^2(\omega)$ given by $P_U(p) = p\min(1, \frac{\gamma}{\|p\|_{L^2(\omega)}})$ is Newton differentiable with Newton derivative given by*

$$DP_U(p)h = \begin{cases} h & \text{if } \|p\|_{L^2(\omega)} \le \gamma \\ \dfrac{\gamma h}{\|p\|_{L^2(\omega)}} - \dfrac{\gamma p(p,\, h)_{L^2(\omega)}}{\|p\|^3_{L^2(\omega)}} & \text{if } \|p\|_{L^2(\omega)} > \gamma. \end{cases} \tag{6.3}$$

Proof. We need to argue that

$$\lim_{\|h\|\to 0} \frac{1}{\|h\|}\|P(p+h) - P_U(p) - DP_U(p+h)h\| = 0$$

for any $p \in L^2(\omega)$, see e.g. [9, page 261]. Throughout the proof $\|\cdot\|$ stands for $\|\cdot\|_{L^2(\omega)}$.

Let us first consider the mapping $G : L^2(\omega) \to \mathbb{R}$ given by

$$G(p) = \min(1, \frac{\gamma}{\|p\|}).$$

Then

$$DG(p)(h) = \begin{cases} 0 & \text{if } \|p\| \le \gamma \\ -\frac{\gamma}{\|p\|^3}(p,\, h)_{L^2} & \text{if } \|p\| > \gamma \end{cases}$$

is a Newton derivative for G at any $p \in L^2(\omega)$. Since G is Fréchet differentiable for every u with $\|p\| \neq \gamma$ it suffices to consider the case $\|p\| = \gamma$. Let $\{h_n\}$ be an arbitrary sequence in $L^2(\omega)$ converging to 0. It has subsequences h_n^1 and h_n^2 with $\|p + h_n^1\| \le \gamma$ and $\|p + h_n^2\| > \gamma$. For the first we have

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{\|h_n^1\|}|G(p+h_n^1) - G(p) - DG(p+h_n^1)\, h_n^1| \\ &= \lim_{n\to\infty} \frac{1}{\|h_n^1\|}(1 - 1 - 0) = 0. \end{aligned} \tag{6.4}$$

For the second one we find

(6.5)

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{\|h_n^2\|}|G(p+h_n^2) - G(p) - DG(p+h_n^2)h_n^2| \\ &= \lim_{n\to\infty} \frac{\gamma}{\|h_n^2\|}\left(\frac{1}{\|p+h_n^2\|} - \frac{1}{\|p\|} + \frac{1}{\|p\|^3}(p,\, h_n^2) - \frac{1}{\|p\|^3}(p,\, h_n^2) + \frac{1}{\|p+h_n^2\|}(p+h_n^2,\, h_n^2)\right) \\ &= \lim_{n\to\infty} \frac{\gamma}{\|h_n^2\|}\left(o\,(\|h_n^2\|) - (\frac{1}{\|p\|^3} - \frac{1}{\|p+h_n^2\|^3})(p,\, h_n^2) + \frac{1}{\|p+h_n^2\|^3}(h_n^2,\, h_n^2)\right) = 0. \end{aligned}$$

Here we used that

$$\frac{1}{\|p+h\|} - \frac{1}{\|p\|} + \frac{1}{\|p\|^3}(p,\, h) = o\,(\|h\|^2). \tag{6.6}$$

In the proof of the following proposition we shall use the fact that o in (6.6) is uniform with respect to p in sets of the form $\{p : \rho \le \|p\| \le R\}$, where $0 < \rho < R$.

With (6.5) Newton differentiability of G holds. Newton differentiability of P_U is now an easy consequence. In fact, we have

(6.7)

$$\begin{aligned}
&P_U\,(p+h) - P_U\,(p) - DP_U\,(p+h)\,h \\
&= (p+h)\,\min\left(1,\, \frac{\gamma}{\|p+h\|}\right) - p\,\min\left(1,\, \frac{\gamma}{\|p+h|}\right) + p\,\big(\min(1,\, \frac{\gamma}{\|p+h\|}) \\
&\qquad - \min\,(1,\, \frac{\gamma}{\|u\|})\big) - h\,\min\left(1,\, \frac{\gamma}{\|p+h\|}\right) - p\,DG\,(p+h)\,h \\
&= p\,(G(p+h) - G(p) - DG\,(p+h)\,h) = o\,(\|h\|).
\end{aligned}$$

□

Remark 6.2. Note that in (6.4) we assigned the value 0 to $DG(p)$ for p at the critical value $\|p\| = \gamma$. This corresponds to the equality sign in (6.3) which appears in the first, rather than the second line.

Similarly to considerations above, one can prove that the mapping

$$\tilde{D}G(p)(h) = \begin{cases} 0 & \text{if } \|p\| < \gamma \\ -\frac{\gamma}{\|p\|^3}(p,\, h)_{L^2} & \text{if } \|p\| \ge \gamma \end{cases}$$

is a Newton derivative of G too. Here, the proof has to be modified in the case $\|p\| = \gamma$ only for a subsequence h_n^3 with $\|p + h_n^3\| = \gamma$. Then one has

$$0 = \|p + h_n^3\|^2 - \|p\|^2 = \|h_n^3\|^2 + 2(p, h_n^3),$$

which implies $(p, h_n^3) = -\frac{1}{2}\|h_n^3\|^2$. Then (6.4) has to be replaced by

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{\|h_n^3\|}|G(p+h_n^3) - G(p) - DG(p+h_n^3)\,h_n^3| \\
&= \lim_{n\to\infty} \frac{1}{\|h_n^3\|}(1 - 1 + \frac{1}{\gamma^2}(p + h_n^3, h_n^3)) \\
&= \frac{1}{2\gamma^2} \lim_{n\to\infty} \frac{1}{\|h_n^3\|}(h_n^3, h_n^3) = 0.
\end{aligned}$$

For our purpose the projection operates pointwise in time on elements $p \in C(I;\, L^2(\omega))$. We denote it by the same symbol.

Proposition 6.3. *The projection* $P_{U_{ad}} : C(I;\, L^2(\omega)) \to L^2(I;\, L^2(\omega))$ *given by* $P_{U_{ad}}(p)(t) = p(t)\min(1, \frac{\gamma}{\|p(t)\|_{L^2(\omega)}})$ *is Newton differentiable with Newton derivative given by* $(DP_U)p(t))$ *as in* (6.3).

Proof. Let $p \in C(I;\, L^2(\omega))$ Further let $\{h_n\}$ be a sequence in $C(I;\, L^2(\omega))$ with $h_n \to 0$ in $C(I;\, L^2(\omega))$. Let

$$K_1 = \{t : \|p(t)\| < \gamma\},\ K_2 = \{t : \|p(t)\| = \gamma\},\ \text{and } K_3 = \{t : \|p(t)\| \geq \gamma\}.$$

We need to estimate

$$\lim_{n\to\infty}(\textstyle\int_0^1 \|P_U(p(t)+h_n(t)) - P_U(p(t)) - DP_U(p(t)+h_n(t))h_n(t)\|^2 dt)^{\frac{1}{2}}$$
$$\leq \lim_{n\to\infty}\big((\textstyle\int_{K_1} q(t)dt)^{\frac{1}{2}} + (\int_{K_2} q(t)dt)^{\frac{1}{2}} + (\int_{K_3} q(t)dt)^{\frac{1}{2}}\big),$$

where $q(t)$ denotes the integrand of the integral on the left hand side. By Lebesgue's bounded convergence theorem

$$\lim_{n\to\infty}\int_{K_1} q(t)dt = \int_{K_1}\lim_{n\to\infty}\|p(t)+h_n(t)-p(t)-h_n(t)\|^2 dt = 0$$

and hence $\lim_{n\to\infty}\frac{1}{\|h_n\|_{C(I;L^2(\omega))}}\left(\int_{K_1} q(t)dt\right)^{\frac{1}{2}} = 0$. Similarly by (6.6)

$$\lim_{n\to\infty}\frac{1}{\|h_n\|_{C(I;L^2(\omega))}}\left(\int_{K_3} q(t)dt\right)^{\frac{1}{2}} = \lim_{n\to\infty}\frac{1}{\|h_n\|_{C(I;L^2(\omega))}}\left(\int_{K_3} o\,(\|h_n(t)\|^4)dt\right)^{\frac{1}{2}} = 0.$$

To obtain the estimate on K_2 we express $h_n(t) = h_n^1(t) + h_n^2(t)$, where

$$h_n^1(t) = \begin{cases} h_n(t) \text{ if } \|p(t)+h_n(t)\| \leq \gamma \\ 0 \text{ if } \|p(t)+h_n(t)\| > \gamma, \end{cases} \quad h_n^2(t) = \begin{cases} 0 \text{ if } \|p(t)+h_n(t)\| \leq \gamma \\ h_n(t) \text{ if } \|p(t)+h_n(t)\| > \gamma. \end{cases}$$

By (6.7) we find using a splitting similar to (6.5)

$$\left(\int_{K_2} q(t)dt\right)^{\frac{1}{2}} = \left(\int_{K_2} \Big(G(p(t)+h_n(t)) - G(p(t)) - DG(p(t)+h_n(t))h_n(t)\Big)^2 \|p(t)\|^2 dt\right)^{\frac{1}{2}}$$

$$= \gamma \left(\int_{K_2} \Big(G(p(t)+h_n^2(t)) - G(p(t)) - DG(p_n(t)+h_n^2(t))h_n^2(t)\Big)^2 dt\right)^{\frac{1}{2}}$$

$$\leq \gamma^2 \left(\int_{K_2} \Big(\frac{1}{\|p(t)+h_n^2(t)\|} - \frac{1}{\|p(t)\|} + \frac{1}{\|p(t)\|^3}(p(t),\, h_n^2(t))\Big)^2 dt\right)^{\frac{1}{2}}$$

$$+ \gamma^2 \left(\int_{K_2} \left(\left(\frac{1}{\|p(t)\|^3} - \frac{1}{\|p(t)+h_n^2(t)\|^3}\right)(p(t),\, h_n^2(t))\right)^2 dt\right)^{\frac{1}{2}}$$

$$+ \gamma^2 \left(\int_{K_2} \frac{1}{\|p(t)+h_n^2(t)\|^6} \|h_n^2(t)\|^4 dt\right)^{\frac{1}{2}}$$

$$\leq \gamma^2 \left(\int_{K_2} o\,(h_n^2(t))^2 dt\right)^{\frac{1}{2}} + \frac{1}{\gamma^3}\left(\int_{K_2} \|h_n^2(t)\|^4 \big(\|p(t)\|^2 + \|p(t)\|\cdot\|h_n^2(t)\| + \|h_n^2(t)\|^2\big)^2 dt\right)^{\frac{1}{2}}$$

$$+ \frac{1}{\gamma}\left(\int_{K_2} \|h_n^2(t)\|^4 dt\right)^{\frac{1}{2}} = o\,(\|h_n\|_{C(I;L^2(\omega))}),$$

and consequently $\lim_{n\to\infty} \frac{1}{\|h_n\|_{C(I;L^2(\omega))}} \left(\int_{K_2} q(t)dt\right)^{\frac{1}{2}} = 0$. This concludes the proof. □

Let us reconsider (6.2) and set $\mathbf{p}_k = \mathbf{p}\,(u_k)$, the solution to the first two equations in (5.1) with u_τ replaced by u_k. Let us further note that the Fréchet derivative of $u \to \mathbf{p}\,(u)$ at u_k in direction δ_u, denoted by $\mathbf{p}' = \mathbf{p}'(u_k)\,\delta_u$ satisfies

$$\begin{cases} \partial_t \mathbf{y}' = \tau \mathbf{A}\mathbf{y}' + \tau \mathbf{B}\,\delta u, & \mathbf{y}'(0) = 0 \\ -\,\partial_t \mathbf{p}' = \tau \mathbf{A}^* \mathbf{p}', & \mathbf{p}'(1) = \dfrac{1}{\varepsilon}\begin{pmatrix} \mathbf{y}_1'(1) \\ (-\Delta)^{-1}\mathbf{y}_2'(1)\end{pmatrix}. \end{cases} \tag{6.8}$$

We further set

$$\mathcal{I} = \{t \in I : \|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)} \leq \varepsilon\gamma\} \text{ and } \mathcal{A} = \{t \in I : \|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)} > \varepsilon\gamma\}.$$

Then (6.2) can equivalently be expressed as

$$\delta u + \frac{\gamma \chi_{\mathcal{A}} \chi_\omega}{\|\mathbf{p}_{k,2}\|}\left(\mathbf{p}_2' - \frac{\mathbf{p}_{k,2}}{\|\mathbf{p}_{k,2}\|}(\frac{\mathbf{p}_{k,2}}{\|\mathbf{p}_{k,2}\|}, \mathbf{p}_2')_{L^2(\mathcal{A};L^2(\omega))}\right) + \frac{1}{\varepsilon}\chi_{\mathcal{I}}\chi_\omega \mathbf{p}_2' = -u_k - P_{U_{ad}}\left(\frac{1}{\varepsilon}\chi_\omega \mathbf{p}_{k,2}\right),$$

where the norms are taken in $L^2(I; L^2(\omega))$. Setting $u_{k+1} = u_k + \delta u$ the Newton update can be expressed as

(6.9)

$$u_{k+1} + \frac{1}{\varepsilon}\chi_{\mathcal{I}}\chi_\omega \mathbf{p}_{k+1,2} + \frac{\gamma \chi_{\mathcal{A}} \chi_\omega}{\|\mathbf{p}_{k,2}\|}\left(\mathbf{p}_{k+1,2} - \frac{\mathbf{p}_{k,2}}{\|\mathbf{p}_{k,2}\|}(\frac{\mathbf{p}_{k_2}}{\|\mathbf{p}_{k,2}\|}, \mathbf{p}_{k+1,2} - \mathbf{p}_{k,2})_{L^2(I;L^2(\omega))}\right) = 0.$$

Proposition 6.4. *The Newton update u_{k+1} is the unique solution to*

(6.10)

$$\begin{cases} \min_{u \in U_{ad}} \tilde{J}(\mathbf{y}, u) = \frac{\tau}{2\varepsilon}\int_0^1 \max\left(\frac{\|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)}}{\gamma}, \varepsilon\right)\|u(t)\|^2\, dt + \frac{1}{2\varepsilon}\|\mathbf{y}(1) - \Theta_\tau \mathbf{z}\|^2_{\mathbf{Y}^0} \\ \text{subject to} \\ \mathbf{y}_t = \tau \mathbf{A}\mathbf{y} + \tau \mathbf{B} u \text{ on } (0,1], \quad \mathbf{y}(0) = \Theta_\tau \mathbf{y}_0, \\ (u(t), \mathbf{p}_{k,2}(t))_{L^2(\omega)} = -\gamma \|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)} \text{ for } t \in \mathcal{A}. \end{cases}$$

Proof. Existence of an unique solution to (6.10) follows from the linear quadratic structure of the problem. To derive the necessary and sufficient optimality system for (6.10) we consider the Lagrangian associated to (6.10), which is given by

$$\mathcal{L}(\mathbf{y}, u, \mathbf{p}, \mu) = \tilde{J}(\mathbf{y}, u) + \langle \mathbf{p}, -\mathbf{y}_t + \tau \mathbf{A}\mathbf{y} + \tau \mathbf{B} u\rangle$$

$$(\mu(\cdot), (u(\cdot), \mathbf{p}_{k,2}(\cdot))_{L^2(\omega)} + \gamma \|\mathbf{p}_{k,2}(\cdot)\|_{L^2(\omega)})_{L^2(\mathcal{A},\mathbb{R})}.$$

It follows that the adjoint equation for (6.10) is given by

$$-\partial_t \mathbf{p} = \tau A^* \mathbf{p}, \quad \mathbf{p}(1) = \begin{pmatrix} \mathbf{y}_1(1) - z_1 \\ (-\Delta)^{-1}(\mathbf{y}_2(1) - \tau z_2) \end{pmatrix}$$

and that

$$\tau \max\left(\frac{\|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)}}{\gamma}, \varepsilon\right) u + \tau \mathbf{p}_2 + \chi_{\mathcal{A}}\, \mu\, \mathbf{p}_{k,2} = 0 \text{ on } (0,1) \times \omega.$$

Consequently

(6.11) $$u + \frac{1}{\varepsilon}\mathbf{p}_2 = 0 \text{ on } \mathcal{I} \times \omega,$$

$$\frac{\|\mathbf{p}_{k,2}\|}{\gamma}u + \mathbf{p}_2 + \frac{1}{\tau}\mu\mathbf{p}_{k,2} = 0 \text{ on } \mathcal{A}\times\omega. \tag{6.12}$$

The latter equation implies that

$$\frac{\|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)}}{\gamma}(u(t),\mathbf{p}_{k,2}(t))_{L^2(\omega)} + (\mathbf{p}_2(t),\mathbf{p}_{k,2}(t))_{L^2(\omega)} + \frac{1}{\tau}\mu(t)\|\mathbf{p}_{k,2}(t)\|^2_{L^2(\omega)} = 0.$$

By the last equality in (6.10)

$$-\|\mathbf{p}_{k,2}\|^2_{L^2(\omega)} + (\mathbf{p}_2(t),\mathbf{p}_{k,2}(t))_{L^2(\omega)} + \frac{1}{\tau}\mu(t)\|\mathbf{p}_{k,2}(t)\|^2_{L^2(\omega)} = 0$$

and hence

$$\mu(t) = \tau - \frac{\tau}{\|\mathbf{p}_{k,2}(t)\|^2_{L^2(\omega)}}(\mathbf{p}_2(t),\mathbf{p}_{k,2}(t))_{L^2(\omega)} \text{ for } t\in\mathcal{A}.$$

By (6.12) we find for $t\in\mathcal{A}$

$$u(t) + \frac{\gamma}{\|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)}}\left(\mathbf{p}_2(t) + \mathbf{p}_{k,2}(t)\left(1 - \frac{(\mathbf{p}_2(t),\mathbf{p}_{k,2}(t))_{L^2(\omega)}}{\|\mathbf{p}_{k,2}(t)\|^2_{L^2(\omega)}}\right)\right) = 0,$$

or equivalently for $t\in\mathcal{A}$

$$u(t) + \frac{\gamma\chi_\omega}{\|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)}}\left(\mathbf{p}_2(t) - \frac{\mathbf{p}_{k,2}(t)}{\|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)}}\left(\mathbf{p}_2(t)-\mathbf{p}_{k,2}(t), \frac{\mathbf{p}_{k,2}(t)}{\|\mathbf{p}_{k,2}(t)\|_{L^2(\omega)}}\right)_{L^2(\omega)}\right). \tag{6.13}$$

Combined, (6.11) and (6.13) imply that (6.9) is satisfied with $(u_{k+1},\mathbf{p}_{k+1}) = (u,\mathbf{p})$. ☐

Uniform boundedness of the generalized derivatives of F is addressed next.

Proposition 6.5. *For every $u\in L^2(I;L^2(\omega))$*

$$\|D\mathcal{F}(u)^{-1}\|_{\mathcal{L}(I;L^2(\omega))} \le 1.$$

Proof. Let $\mathbf{p} = \mathbf{p}(u)$ denote the solution to the adjoint equation as in (P^τ_ε) with u_τ replaced by u and let $\mathbf{p}' = \mathbf{p}'(h)$ denote the linearization of $u \to \mathbf{p}(u)$

at u in direction $h \in L^2(I; L^2(\omega))$, i.e. $\mathbf{p}'$ satisfies (6.8) with $\delta u = h$.
We set

$$\mathcal{A} = \{t \in I : \|\mathbf{p}_2(t)\|_{L^2(\omega)} > \varepsilon\gamma\}, \quad \mathcal{I} = \{t \in I : \|\mathbf{p}_2(t)\|_{L^2(\omega)} \leq \varepsilon\gamma\},$$

and we define

$$q(t) = \frac{\mathbf{p}_2(t)}{\|\mathbf{p}_2(t)\|_{L^2(\omega)}} \chi_\omega \chi_{\mathcal{A}}.$$

We have the orthogonal decomposition

$$h = h_1 + h_2 = (h, q)_{L^2(I;L^2(\omega))} q + (h - (h, q)_{L^2(I;L^2(\omega))} q).$$

According to Proposition 6.3 we find

$$D\mathcal{F}(u)h = h + \frac{1}{\varepsilon}\chi_\omega \chi_{\mathcal{I}} \mathbf{p}_2' + \frac{\gamma \chi_\omega \chi_{\mathcal{A}}}{\|\mathbf{p}_2'\|_{L^2(\omega)}} (\mathbf{p}_2' - \frac{\mathbf{p}_2(\mathbf{p}_2, \mathbf{p}_2')_{L^2(\omega)}}{\|\mathbf{p}_2\|^2_{L^2(\omega)}}),$$

where dependence on t is suppressed. We note that $(\mathbf{p}_2' - \frac{\mathbf{p}_2(\mathbf{p}_2, \mathbf{p}_2')_{L^2(\omega)}}{\|\mathbf{p}_2\|_{L^2(\omega)}}, h_1)_{L^2(\omega)} = 0$. Consequently we have

$$(D\mathcal{F}(u)h, h)_{L^2(I,L^2(\omega))} = |h|^2_{L^2(I;L^2(\omega))} + \frac{1}{\varepsilon}(\mathbf{p}_2'(h), h)_{L^2(I;L^2(\omega))}. \tag{6.14}$$

We find

$$\begin{aligned}
\tau(\mathbf{p}_2'(h), h)_{L^2(I;L^2(\omega))} &= \langle \mathbf{p}(h), \partial_t \mathbf{y}' - \tau \mathbf{A}\mathbf{y}' \rangle \\
&= (\mathbf{p}'(h)(1), \mathbf{y}'(1)) + \langle \partial_t \mathbf{p}'(h), \mathbf{y}' \rangle = 2(\mathbf{p}'(h)(1), \mathbf{y}'(1)) \\
&= \frac{2}{\varepsilon}(\|\mathbf{y}_1'(1)\|^2_{L^2(\Omega)} + \|\mathbf{y}_2'(1)\|^2_{H^{-1}(\Omega)}) \geq 0
\end{aligned}$$

Together with (6.14) this implies the claim. □

We are now prepared to verify local convergence of the semi-smooth Newton iteration (6.2).

Theorem 6.6. *Assume that* $\|u_0 - u_\tau\|_{L^2(I;L^2(\omega))}$ *is sufficiently small. Then the semi-smooth Newton iteration* (6.2) *converges superlinearly to* u_τ *in* $L^2(I; L^2(\omega))$.

Proof. The mapping $\mathcal{F}$ is continuous and by Proposition 6.3 is also Newton differentiable at every $u \in L^2(I; L^2(\omega))$. By Proposition 6.5 the inverses of $D\mathcal{F}(u)$ are uniformly bounded. The conclusion therefore follows from well-known results on semi-smooth Newton methods, cf. [9, page 238].

□

7 A semi-smooth Newton algorithm

In this section a Newton-type algorithm for solving (P) proposed. Minimization with respect to $(\tau, \mathbf{y}, u, \mathbf{p})$ is carried out in two nested loops, an outer τ- and an inner $(\mathbf{y}, u, \mathbf{p})$-loop. Key ingredients of the method are provided by Theorem 5.3 stating that the gradient of $\mathcal{V}$ is given by the transversality condition and, secondly, that for each given τ, the optimal control problem (P_ε^τ) can be solved efficiently by a semi-smooth Newton method. Selected examples illustrate the feasibility of the approach.

While we present this approach here for optimal control of the wave equation, this is a procedure that could be of much wider relevance even in the context of time optimal control of ordinary differential equations. As for the latter, while we do not intend to give a survey of the literature here, let us very briefly point out some of the common techniques that are employed. These include the indirect methods based on Pontryagin's minimal principle, and subsequent realization of the two point boundary problems by multiple shooting techniques, see e.g. [4]. The literature indicates that if good initial conditions can be supplied then this approach is fast and accurate. A second popular approach are direct methods that transform the control problem to nonlinear programming problems, see e.g. [18] and the references given there. In this case the unknowns are, transferred to be switching times [16, 17], or arc-time optimization [11]. Also in these cases, a-priori information, for example on the switching structure, or availability of a feasible solution are required.

7.1 Description of the algorithm

Let us explain some details of the algorithm. The basic building block is a Newton algorithm for the minimization of the real function $\mathcal{V}$, defined in (5.6) by

$$\mathcal{V}(\tau) = \tau \left(1 + \frac{\varepsilon}{2} \|u_\tau\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon} \|\mathbf{y}_\tau(1) - \Theta_\tau \mathbf{z}\|^2_{Y^0},$$

where $(\mathbf{y}_\tau, u_\tau)$ is the unique solution of the parametric optimization problem (P_ε^τ). The first and second derivatives of $\mathcal{V}$ are computed on the basis of Theorems 5.2 and 5.3. The resulting Newton algorithm is embedded in an outer loop that drives the parameter ϵ to zero. The overall solution procedure is depicted in Algorithm 1.

Algorithm 1 Outline of the solution algorithm

Initial guess: ϵ_0, τ_0.
Parameters: $\theta \in (0,1)$, $r_{\text{tol}} > 0$.
Initialization: $k := 0$, $i = 0$, $r_0 := +\infty$.
repeat {Solve (P) - ε-loop}
  Solve (P_ε^τ), solution $(\mathbf{y}_{\tau_k}, u_{\tau_k}, \mathbf{p}_{\tau_k})$.
  repeat {Solve (P_ε) - τ-loop}
    Solve sensitivity system (7.1)
    Compute $s_k := \frac{d}{d\tau} V(\tau_k)$ and $h_k := \frac{d^2}{d\tau^2} V(\tau_k)$
    Set $\sigma = 2$
    repeat {Line-search in τ}
      Set $\sigma := \sigma/2$.
      Set $\tilde{\tau}_k := \tau_k - \sigma h_k^{-1} s_k$.
      Solve (P_ε^τ), solution $(\mathbf{y}_{\tilde{\tau}_k}, u_{\tilde{\tau}_k}, \mathbf{p}_{\tilde{\tau}_k})$.
    until $\mathcal{V}(\tau_k) - \mathcal{V}(\tilde{\tau}_k) \geq \theta\sigma |s_k|^2$
    Set $(\tau_{k+1}, \mathbf{y}_{\tau_{k+1}}, u_{\tau_{k+1}}, \mathbf{p}_{\tau_{k+1}}) := (\tilde{\tau}_k, \mathbf{y}_{\tilde{\tau}_k}, u_{\tilde{\tau}_k}, \mathbf{p}_{\tilde{\tau}_k})$.
    Set $k := k + 1$.
  until $|\frac{d}{d\tau} V(\tau_k)| < \varepsilon_i$.
  Compute $\tilde{r}_{i+1} := \|\mathbf{y}_{\tau_k}(1) - \Theta_{\tau_k} \mathbf{z}\|_{\mathbf{Y}^0}$.
  if $\frac{\tilde{r}_{i+1} - r_i}{r_i} \leq 0.1$ **then** {Global search necessary}
    for $j = 1, 2, \ldots$ **do**
      if $\mathcal{V}(\tilde{\tau}^j) < \mathcal{V}(\tau_k)$ or $\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^j) < 0$ **then** {Accept step}
        Set $\tau_k := \tilde{\tau}^j$
        continue
      end if
      if $|\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^j) - 1| \leq \sqrt{0.6}\, |\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^{j-1}) - 1|$ **then**
        {Asymptotic regime reached}
        break
      end if
    end for
    {Mesh refinement necessary}
    Do uniform spatial and temporal mesh refinement.
    continue
  end if
  Set $r_{i+1} := \tilde{r}_{i+1}$.
  Set $\varepsilon_{i+1} := \varepsilon_i / 2$.
  Set $i = i + 1$.
until $r_t < r_{\text{tol}}$

Some remarks are in order. As already mentioned, the inner parametric optimization problem (P_ε^τ) is solved by a standard semi-smooth Newton method combined with an Armijo-type linesearch. The semi-smooth Newton method is stopped as soon as the residual at the current iterate $u_{\tau_k}^i$ satisfies

$$\|\mathcal{F}\left(u_{\tau_k}^i\right)\|_{L^2(I;L^2(\omega))} \le \max\left(10^{-2}\left|\frac{d}{d\tau}\mathcal{V}(\tau_{k-1})\right|,\, 10^{-6}\right),$$

see (6.1) for the definition of $\mathcal{F}$.

To apply Newton's method to the minimization of $\mathcal{V}$, we need to describe how the second derivative of $\mathcal{V}$ is computed. For given τ let $(\dot{\mathbf{y}}, \dot{u}, \dot{\mathbf{p}})$ denote the solution of the sensitivity system

$$\begin{cases} \partial_t\dot{\mathbf{y}} = \tau(\mathbf{A}\dot{\mathbf{y}} + \mathbf{B}\dot{u}) + \mathbf{A}\mathbf{y}_\tau + \mathbf{B}u_\tau,\ \dot{\mathbf{y}}(0) = \begin{pmatrix} 0 \\ y_2 \end{pmatrix}, \\ -\partial_t\dot{\mathbf{p}} = \tau\mathbf{A}^*\dot{\mathbf{p}} + \mathbf{A}\mathbf{p}_\tau,\ \dot{\mathbf{p}}(1) = \frac{1}{\varepsilon}\begin{pmatrix} \dot{\mathbf{y}}_1(1) \\ (-\Delta)^{-1}(\dot{\mathbf{y}}_2(1) - z_2) \end{pmatrix}, \\ \varepsilon\dot{u} + DP_{U_{ad}}\left(-\frac{1}{\varepsilon}\mathbf{B}^*\mathbf{p}_\tau\right)\mathbf{B}^*\dot{\mathbf{p}} = 0. \end{cases} \tag{7.1}$$

It can be shown with similar arguments as in Section 5 that (7.1) is uniquely solvable. In fact, this system can be written as an equation $D\mathcal{F}(u_\tau)\dot{u} = v(\tau, \delta\tau)$ for some right-hand side $v(\tau, \delta\tau)$. We observed that the computational effort to solve (7.1) was comparable to the effort of computing one Newton step for (P_ε^τ). Once the solution of (7.1) is available, the second derivative of $\mathcal{V}$ is obtained from

$$\begin{aligned} \frac{d^2}{d\tau^2}\mathcal{V}(\tau) &= \varepsilon(u_\tau, \dot{u})_{L^2(I;L^2(\omega))} + (\mathbf{A}\dot{\mathbf{y}} + \mathbf{B}\dot{u}, \mathbf{p}_\tau)_{L^2(I;\mathbf{L}^2(\Omega))} \\ &\quad + (\mathbf{A}\mathbf{y}_\tau + \mathbf{B}u_\tau, \dot{\mathbf{p}})_{L^2(I;\mathbf{L}^2(\Omega))} + (y_2, \dot{\mathbf{p}}_2(0))_{L^2(\Omega)} - (z_2, \dot{\mathbf{p}}_2(1))_{L^2(\Omega)}. \end{aligned}$$

As stopping criteria for the solution of the regularized problem (P_ε) we chose $|\frac{d}{d\tau}\mathcal{V}(\tau_k)| < \varepsilon_i$. That is, the precision in which the inner problem is solved is directly coupled to the regularization parameter. This prevents the algorithm to spend too much time to solve problems with large regularization parameter.

We observed that the resulting algorithm sometimes gets trapped in local minima of $\mathcal{V}$. These local minima did not vanish for $\varepsilon \to 0$. Hence a globalization step is needed. It is based on monitoring the development of the residual at the terminal time. As can be seen from Corollary 3.3, we expect the residual $\tilde{r}_{i+1} := \|\mathbf{y}_{\tau_k}(1) - \Theta_{\tau_k}\mathbf{z}\|_{\mathbf{Y}^0}$ to be of order $o(\sqrt{\varepsilon_k})$. Since in each outer iteration the regularization parameter ε is reduced by a factor of $1/2$, the accepted residual r_i should decrease by a factor of $1/\sqrt{2}$. If only a small or even no reduction in the residual is achieved, we assume that the iterates τ_k are trapped in a local minimum of $\mathcal{V}$. That is, we start a global search if

$$\rho(\tilde{r}_{i+1}, r_i) := \frac{r_i - \tilde{r}_{i+1}}{r_i} < 0.1.$$

In the global search, we compute new trial values $\tilde{\tau}^j := \tau_k + 0.1 \cdot 2^j$. We accept $\tilde{\tau}^j$ if $\mathcal{V}(\tilde{\tau}^j) < \mathcal{V}(\tau_k)$ or $\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^j) < 0$. Then the Newton method is restarted at $\tilde{\tau}^j$. The global search algorithm is stopped if

$$\left|\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^j) - 1\right| \leq \sqrt{0.6}\ \left|\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^{j-1}) - 1\right|,$$

as we expect due to Theorem 5.4

$$\frac{|\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^j) - 1|}{|\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^{j-1}) - 1|} \to \frac{1}{\sqrt{2}},$$

for $j \to \infty$. That is, we stop the global search if $\frac{d}{d\tau}\mathcal{V}(\tilde{\tau}^j)$ asymptotically reaches $+1$, since in this case there will be no local minima τ of $\mathcal{V}$ with $\tau > \tilde{\tau}^j$. If the global search is stopped then we assume that the discretization error dominates the residual and that with the current discretization no further reduction of the residual will be possible. Hence the discretization is refined, and the Newton method for $\mathcal{V}$ is restarted on the finer grid at the old iterate τ_k.

Turning to the discretization scheme, the state-, adjoint-, and control variables were discretized by finite elements. The amplitude components $\mathbf{y}_1$ and $\mathbf{p}_2$ are discretized by P1-elements, while P0- elements are used for the velocity unknowns $\mathbf{y}_2$ and $\mathbf{p}_1$, and for the controls.

For time discretization, we used a cg(1)dG(0)-scheme as described in [13], which corresponds to a Crank-Nicolson time-stepping procedure.

7.2 Numerical experiments

Let us report on the outcome of our computational experiments. We modified the cost functional of the penalized problem (P_ε) to

$$J_\varepsilon(\tau, u) = \tau\left(1 + \frac{\varepsilon}{2}\|u\|^2_{L^2(I;L^2(\omega))}\right) + \frac{1}{2\varepsilon_y}\|\mathbf{y}(1) - \Theta_\tau\mathbf{z}\|^2_{\mathbf{Y}^0}$$

with

$$\varepsilon_y = 0.1 \cdot \varepsilon.$$

That is, the violation of the terminal constraint is penalized with larger weight.

We chose $\Omega = (0,1)^2$ for our computations. The control bound was set to $\gamma = 3$. The target state was $z_1 = z_2 = 0$. The initial state was given as $y_1(x_1, x_2) = x_1 x_2 (1 - x_1)(1 - x_2)$.

The spatial domain was discretized using a uniform triangulation, and the time interval was split into equidistant subintervals. We will report on the results for the following hierarchy of discretizations: $(N, M) = (50, 10)$, $(200, 20)$, $(800, 40)$, and $(3200, 80)$, where N is the numbers of triangles and M the numbers of time intervals. The resulting mesh size h is $h = 2/\sqrt{N}$, the resulting length of the temporal subintervals $\Delta t = 1/M$.

The parameters for the algorithm as described in the previous section were chosen as

$$\varepsilon_0 = 0.1, \quad \tau_0 = 1.9, \quad \theta = 10^{-3}.$$

The algorithm was stopped as soon as the terminal residual satisfies $\|\mathbf{y}_\varepsilon(1) - \Theta_\tau \mathbf{z}\|_{\mathbf{Y}^0} \leq 10^{-3}$, which corresponds to the choice $r_{\text{tol}} = 10^{-3}$.

Example 1

Here the initial velocity was $y_2 = 0$. That is, the control objective was to steer the system from a given deflection into zero. The control domain was chosen to be $\omega := \Omega$.

We repost on the convergence for $\varepsilon \to 0$. In Figure 1 we show the convergence of τ_ε for different discretizations. Moreover, we depict the evolution of $|\tau_\varepsilon - \tau^*|$ for the finest discretization, where we use as value for τ^* the optimal time for the smallest ε_i, i.e. $\tau^* = \tau_{\varepsilon_i}$. In Table 1 we report on the convergence of τ_ε and $\|\mathbf{y}_\varepsilon(1) - \Theta_\tau \mathbf{z}\|_{\mathbf{Y}^0}$ for the finest discretization. We observe the convergence rate

$$|\tau_\varepsilon - \tau^*| = O(\varepsilon).$$

As argued in Remark 4.9 this implies that

$$\|\mathbf{y}_\varepsilon(1) - \Theta_\tau \mathbf{z}\|_{\mathbf{Y}^0} = O(\varepsilon),$$

which can be seen in the table as well. This convergence rate implies that $\{\mathbf{p}_\varepsilon(1)\}$ is bounded in $\mathbf{P}^1$, where $\mathbf{p}_\varepsilon$ are the solutions of the undiscretized

ε	τ_ε	$\lvert\tau_\varepsilon - \tau^*\rvert$	$\lVert\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon}\mathbf{z}\rVert_{\mathbf{Y}^0}$
$1.0000 \cdot 10^{-1}$	$3.68 \cdot 10^{-1}$	$1.52 \cdot 10^{-2}$	$2.49 \cdot 10^{-2}$
$5.0000 \cdot 10^{-2}$	$3.77 \cdot 10^{-1}$	$7.03 \cdot 10^{-3}$	$1.03 \cdot 10^{-2}$
$2.5000 \cdot 10^{-2}$	$3.80 \cdot 10^{-1}$	$3.56 \cdot 10^{-3}$	$4.48 \cdot 10^{-3}$
$1.2500 \cdot 10^{-2}$	$3.82 \cdot 10^{-1}$	$1.82 \cdot 10^{-3}$	$2.12 \cdot 10^{-3}$
$6.2500 \cdot 10^{-3}$	$3.83 \cdot 10^{-1}$	$6.67 \cdot 10^{-4}$	$1.02 \cdot 10^{-3}$
$3.1250 \cdot 10^{-3}$	$3.84 \cdot 10^{-1}$		$5.10 \cdot 10^{-4}$

Table 1: Convergence history for Example 1

problem. In Figure 2, we plotted the evolution of $\lVert\mathbf{p}_{\varepsilon,h}(1)\rVert_{\mathbf{P}^1}$ for the solutions of the discrete problems for the 4 different discretizations, and we observe that the $\mathbf{P}^1$-norms of $\mathbf{p}_{\varepsilon,h}(1)$ are bounded uniformly with respect to ε and with respect to the discretization. This suggests that the strong requirements of Theorem 4.7 are satisfied for this example, and that (4.8) holds in the limit $\varepsilon \to 0$.

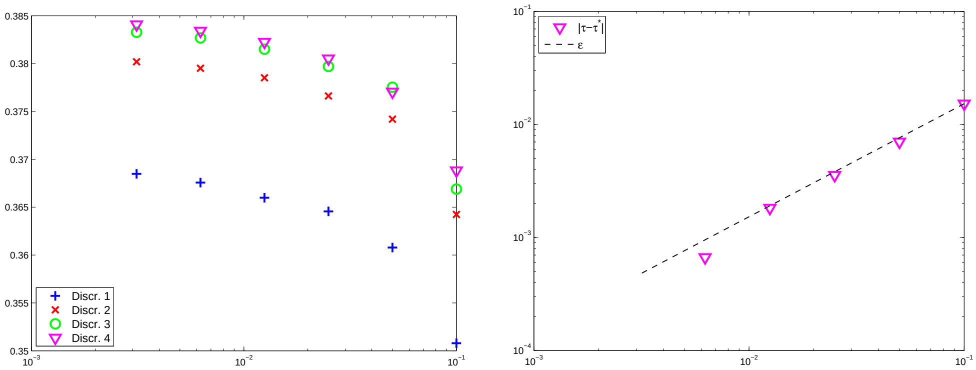

Figure 1: (Example 1) τ_ε vs. ε for different discretizations; $\lvert\tau_\varepsilon - \tau^*\rvert$ vs. ε for finest discretization

Example 2

In our second example we chose the initial velocity to be $y_2(x_1, x_2) = x_1^2 + y_2^2$. All other data are the same as in the previous example. The convergence history can be found in Table 2. We observed the same convergence rates

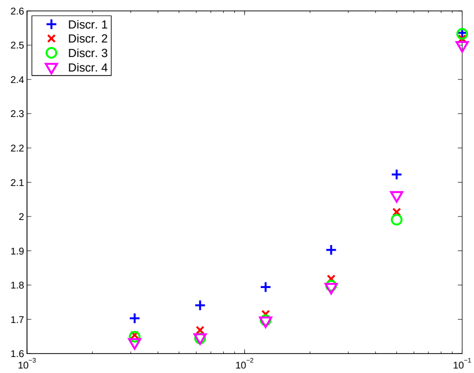

Figure 2: (Example 1) $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^1}$ vs. ε for different discretizations

as in the previous example, i.e. $\|\mathbf{y}_\varepsilon(1) - \Theta_\tau \mathbf{z}\|_{\mathbf{Y}^0} = O(\varepsilon)$. Moreover, $\{\mathbf{p}_\varepsilon(1)\}$ is bounded in $\mathbf{P}^1$, which indicates that strict transversality holds for the original problem.

ε	τ_ε	$\|\tau_\varepsilon - \tau^*\|$	$\|\mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon}\mathbf{z}\|_{\mathbf{Y}^0}$
$1.0000 \cdot 10^{-1}$	$4.00 \cdot 10^{-1}$	$3.15 \cdot 10^{-2}$	$2.87 \cdot 10^{-2}$
$5.0000 \cdot 10^{-2}$	$4.18 \cdot 10^{-1}$	$1.45 \cdot 10^{-2}$	$1.21 \cdot 10^{-2}$
$2.5000 \cdot 10^{-2}$	$4.24 \cdot 10^{-1}$	$7.81 \cdot 10^{-3}$	$5.68 \cdot 10^{-3}$
$1.2500 \cdot 10^{-2}$	$4.28 \cdot 10^{-1}$	$3.72 \cdot 10^{-3}$	$2.71 \cdot 10^{-3}$
$6.2500 \cdot 10^{-3}$	$4.31 \cdot 10^{-1}$	$1.34 \cdot 10^{-3}$	$1.33 \cdot 10^{-3}$
$3.1250 \cdot 10^{-3}$	$4.32 \cdot 10^{-1}$		$7.00 \cdot 10^{-4}$

Table 2: Convergence history for Example 2

Example 3

Here again the initial velocity was $y_2 = 0$, however, the control domain ω was chosen to be a proper subset of Ω: $\omega = \Omega \setminus [0, 0.5]^2$. For this control domain, we have controllability for times $T \geq 1$.

Here the convergence rates are worse than in example 1. In fact, we observe $|\tau_\varepsilon - \tau^*| = O(\sqrt{\varepsilon})$.

Another interesting observation is that on the finest discretization a smaller value of ε and thus more iterations to reach the prescribed accuracy for the

ε	τ_ε	$\lvert\tau_\varepsilon - \tau^*\rvert$	$\lVert \mathbf{y}_\varepsilon(1) - \Theta_{\tau_\varepsilon} \mathbf{z} \rVert_{\mathbf{Y}^0}$
$1.0000 \cdot 10^{-1}$	$3.98 \cdot 10^{-1}$	$2.25 \cdot 10^{-1}$	$3.97 \cdot 10^{-2}$
$5.0000 \cdot 10^{-2}$	$4.45 \cdot 10^{-1}$	$1.77 \cdot 10^{-1}$	$2.41 \cdot 10^{-2}$
$2.5000 \cdot 10^{-2}$	$4.86 \cdot 10^{-1}$	$1.37 \cdot 10^{-1}$	$1.50 \cdot 10^{-2}$
$1.2500 \cdot 10^{-2}$	$5.21 \cdot 10^{-1}$	$1.02 \cdot 10^{-1}$	$9.25 \cdot 10^{-3}$
$6.2500 \cdot 10^{-3}$	$5.47 \cdot 10^{-1}$	$7.64 \cdot 10^{-2}$	$5.81 \cdot 10^{-3}$
$3.1250 \cdot 10^{-3}$	$5.67 \cdot 10^{-1}$	$5.58 \cdot 10^{-2}$	$3.74 \cdot 10^{-3}$
$1.5625 \cdot 10^{-3}$	$5.84 \cdot 10^{-1}$	$3.95 \cdot 10^{-2}$	$2.52 \cdot 10^{-3}$
$7.8125 \cdot 10^{-4}$	$5.98 \cdot 10^{-1}$	$2.50 \cdot 10^{-2}$	$1.73 \cdot 10^{-3}$
$3.9063 \cdot 10^{-4}$	$6.12 \cdot 10^{-1}$	$1.15 \cdot 10^{-2}$	$1.21 \cdot 10^{-3}$
$1.9531 \cdot 10^{-4}$	$6.23 \cdot 10^{-1}$		$9.02 \cdot 10^{-4}$

Table 3: Convergence history for Example 3

terminal residual are needed, see Figure 3.

Moreover, the norms $\|\mathbf{p}_{\varepsilon,h}(1)\|_{\mathbf{P}^1}$ are not bounded uniformly with respect to the discretization, see Figure 4. However, $\frac{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^1}}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}$ is bounded uniformly with respect to ε and discretization, and we therefore expect that the conclusion (4.9) of Theorem 4.7 is true for the limit $\varepsilon \to 0$.

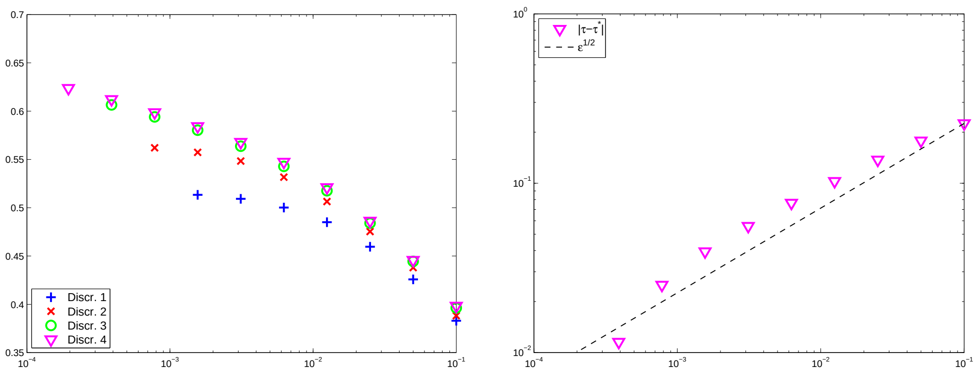

Figure 3: (Example 3) τ_ε vs. ε for different discretizations; $|\tau_\varepsilon - \tau^*|$ vs. ε for finest discretization

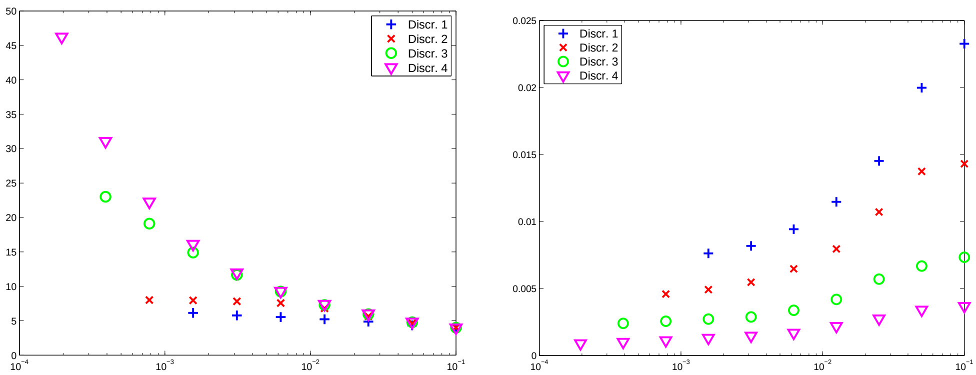

Figure 4: (Example 3) $\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^1}$ and $\frac{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^1}}{\|\mathbf{p}_\varepsilon(1)\|_{\mathbf{P}^0}}$ vs. ε for different discretizations

8 Concluding remarks

Clearly many interesting questions can be addressed as extensions of the present research. Here we used an $L^2(\omega)$-norm constraint on the controls spatially. Pointwise controls are equally important and require, in part, different treatment. Boundary control problems, as well, are a natural subsequent problem to be addressed. Here we used a Newton type methods to determine τ and $(\mathbf{y}, u, \mathbf{p})$ separately in a nested approach. The question arises of treating the combined problem in the variables $(\tau, \mathbf{y}, u, \mathbf{p})$ by a Newton-type algorithm. Discretization issues should also be further investigated.

References

[1] V. Barbu. *Optimal control of variational inequalities*, volume 100 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[2] D. Barcenas, H. Leiva, and T. Maya. The transversality condition for infinite dimensional control systems. *Revista Notas de Matematica*, 4:25–36, 2008.

[3] C. Bardos, G. Lebeau, and J. Rauch. Un exemple d'utilisation des notions de propagation pour le contrôle et la stabilisation de problèmes

hyperboliques. *Rend. Sem. Mat. Univ. Politec. Torino*, pages 11–31 (1989), 1988. Nonlinear hyperbolic equations in applied sciences.

[4] R. Bulirsch. Die Mehrzielmethode zur numerischen Lösung von nichtlinearen Randwertproblemen und Aufgaben der optimalen Steuerung. *DLR, Oberpfaffenhofen, Germany, Report of the Carl-Cranz-Gesellschaft (1971)*. Reprint: Department of Mathematics, Munich University of Technology (1985).

[5] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[6] H. O. Fattorini. The time optimal problem for distributed control of systems described by the wave equation. In *Control theory of systems governed by partial differential equations (Conf. Naval Surface Weapons Center, Silver Spring, Md., 1976)*, pages 151–175. Academic Press, New York, 1977.

[7] H. O. Fattorini. *Infinite-dimensional optimization and control theory*, volume 62 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999.

[8] H. Hermes and J. P. LaSalle. *Functional analysis and time optimal control*. Academic Press, New York, 1969. Mathematics in Science and Engineering, Vol. 56.

[9] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[10] K. Ito and K. Kunisch. Semismooth Newton methods for time-optimal control for a class of ODEs. *SIAM J. Control Optim.*, 48(6):3997–4013, 2010.

[11] C. Y. Kaya and J. L. Noakes. Computational method for time-optimal switching control. *J. Optim. Theory Appl.*, 117(1):69–92, 2003.

[12] H. B. Keller. *Numerical methods for two-point boundary-value problems.* Blaisdell Publishing Co. Ginn and Co., Waltham, Mass.-Toronto, Ont.-London, 1968.

[13] A. Kröner, K. Kunisch, and B. Vexler. Semismooth Newton methods for optimal control of the wave equation with control constraints. *SIAM Journal Contr. Opt.*, 2011. To appear.

[14] E. B. Lee and L. Markus. *Foundations of optimal control theory.* John Wiley & Sons Inc., New York, 1967.

[15] J.-L. Lions. *Control of distributed singular systems.* Gauthier-Villars, Kent, 1985.

[16] H. Maurer and N. P. Osmolovskii. Second order sufficient conditions for time-optimal bang-bang control. *SIAM J. Control Optim.*, 42(6):2239–2263, 2004.

[17] E. Meier and A. E. Bryson, Jr. Efficient algorithm for time-optimal control of a two-link manipulator. *J. Guidance Control Dynam.*, 13(5):859–866, 1990.

[18] O. von Stryk and R. Bulirsch. Direct and indirect methods for trajectory optimization. *Ann. Oper. Res.*, 37(1-4):357–373, 1992. Nonlinear methods in economic dynamics and optimal control (Vienna, 1990).

[19] E. Zuazua. Propagation, observation, and control of waves approximated by finite difference methods. *SIAM Rev.*, 47(2):197–243, 2005.