sciendo
Control and Cybernetics
vol. 50 (2021) No. 4
pages: 427- 455
DOI: 10.2478/candc-2021-0026

Time-domain decomposition for optimal control problems governed by semilinear hyperbolic systems with mixed two-point boundary conditions*

by

Richard Krug[1], Günter Leugering[1], Alexander Martin[1], Martin Schmidt[2] and Dieter Weninger[1]

[1]Friedrich-Alexander-Universität Erlangen-Nürnberg, Department of Data Science, Cauerstr. 11, 91058 Erlangen, Germany,
[2]Trier University, Department of Mathematics, Universitätsring 15, 54296 Trier, Germany
richard.krug@fau.de, alexander.martin@fau.de, dieter.weninger@fau.de
guenter.leugering@fau.de, martin.schmidt@uni-trier.de

Abstract: In this article, we study the time-domain decomposition of optimal control problems for systems of semilinear hyperbolic equations and provide an in-depth well-posedness analysis. This is a continuation of our work, Krug et al. (2021) in that we now consider *mixed* two-point boundary value problems. The more general boundary conditions significantly enlarge the scope of applications, e.g., to hyperbolic problems on metric graphs with cycles. We design an iterative method based on the optimality systems that can be interpreted as a decomposition method for the original optimal control problem into *virtual control* problems on smaller time domains.

Keywords: time-domain decomposition, optimal control, semilinear hyperbolic systems, convergence

1. Introduction

In this article, we are concerned with an iterative and completely parallel time-domain decomposition technique for optimality systems associated to optimal control problems for semilinear, one-dimensional, hyperbolic systems with general mixed two-point boundary conditions. The main interest is to design an iterative method such that the time-wise decomposed optimality system at each iteration is by itself an iteration of parallel optimality systems on the smaller time-domains. In the following, we explain the context more specifically and provide (non-exhaustive) references to the literature.

*Submitted: July 2021; Accepted: November 2021

Time-domain decomposition for partial differential equations (PDEs) has been a subject of intense research in the past. Let a partial differential equation with respect to space and time be given with time domain $[0, T]$. It often occurs in the applications that the time horizon is very large and, hence, the computation of the entire time evolution of the system is very costly. The desire is to apply a time-domain decomposition that allows to reduce the original problem on the long time horizon to similar problems on a short time interval such that the sub-problems can be treated in parallel. To be more precise, one introduces a coarse time discretization of $[0, T]$ into a disjoint union of subintervals $I_k := [T_k, T_{k+1}]$ with $[0, T] = \mathrm{cl}(\cup_{k=1}^{K}(T_k, T_{k+1}))$ such that on each subinterval I_k, the same PDE is solved together with time-like transmission conditions at the breakpoints T_k that couple the states at the current iteration $n + 1$ with those at iteration n.

This approach is not new, in principle, and can be traced back to the contributions to the seminal paper by Lions, Maday and Turinici (2001), in which the so-called "parareal"-scheme has been introduced, which, in turn, has later been identified as a variant of the common multiple-shooting method; see, e.g., Gander and Vandewalle (2007). These methods, see also Maday, Salomon and Turinici (2006), Maday and Turinici (2002), which consist of a coupling of coarse grain discrete-in-time solutions at the break points with a parallel computation of full (respectively, small grain) solutions on the subintervals, were first developed for the mere simulation of nonlinear PDEs. In Lagnese and Leugering (2002), the authors, for the first time, considered the time-domain decomposition of optimal control problems for the time-dependent Maxwell system. Later, in Lagnese and Leugering (2004), a broad number of such problems–even combined with a spatial domain decomposition for PDEs on networked domains–have been investigated. We also refer to Barker and Stoll (2015), Gander and Kwok (2016), Gander, Kwok and Salomon (2020), Wu and Huang (2018), Wu and Liu (2020), as well as Heinkenschloss (2005) and Ulbrich (2007), where methods related to multiple-shooting have been provided along with applications for the heat equation also in the context of optimal control problems. A distinguishing feature of the method in Lagnese and Leugering (2002, 2003, 2004) is the fact mentioned above, namely, that the iterative time-domain decomposition is applied to the optimality system of the original optimal control problem on the time domain $[0, T]$ in such a way that the decomposed problems are by themselves the optimality systems, corresponding to the so-called virtual control problems on the subintervals I_k. Thus, the fully parallel iteration can be seen as one of optimal control problems on the subintervals.

The analysis here is provided on the continuous level and relies on the fact that the state variables evolve forwardly in time, whereas the adjoint variable progresses backwardly. A time-domain decomposition for semilinear optimal control problems that enjoys this distinguishing feature has been considered in Krug et al. (2021) for separated two-point boundary conditions. It turns out that in a number of important applications, including problems on metric graphs with cycles, such separated two-point initial boundary value problems are not applicable. The goal of this article is to extend the time-domain decomposition method of Lagnese and Leugering (2003, 2004), that has been the conceptual basis for Krug et al. (2021), to optimal control problems involving semilinear hyperbolic systems of conservation laws with more general boundary conditions. Moreover, in this article, we restrict ourselves to coefficients constant with respect to time. This allows us to add a complete well-posedness analysis based on semilinear perturbations of linear semi-groups.

We emphasize that the resulting algorithm is derived in analogy to the classic spatial domain decomposition for elliptic problems by P. L. Lions (1990) which was, in turn, interpreted by Glowinski and LeTallec (1989) as a variant of a Uzawa-type saddle-point iteration. The relaxation that we introduce in Section 5, represented by the parameter $\varepsilon \in [0, 1)$, is related to a damped Richardson ansatz. This indicates that, in general, very rapid convergence may not take place. However, in very special cases, in particular for separated boundary conditions, distributed controls and simple homogeneous state equations, we observe almost "two-step-convergence", as in the optimized non-overlapping Schwarz iterations. See Krug et al. (2021) for details and numerical experiments and Lagnese and Leugering (2004) for a general discussion.

We are particularly interested in applications that focus on processes on metric graphs or networks containing cycles. As an example, we will focus on networks of semilinear strings or rods. Systems that are related to gas flow in pipe networks are easily seen to fit into the framework of this article—for the model, see, e.g., Hante et al. (2017) or Leugering et al. (2017). Such problems on metric graphs with cycles, where the edges, which are representative of the spatial domains of the corresponding PDEs, are coupled at the vertices of the graph, can be transformed into mixed two-point initial boundary value problems with a possibly large number of state variables; see Example 1 for further explanation.

The remainder of the article is organized as follows. We begin with the problem statement in Section 2, where we include a detailed discussion of an example of a network of controlled strings or rods, possibly including cycles. In this example, a local nonlinear damping term is present along some or all strings involved, together with nonlinear boundary conditions. In Section 3, we introduce the time-domain decomposition method for the overall optimality system into systems on the subintervals, and show in which way these decomposed systems are themselves optimality systems for "virtual" optimal control problems on the subintervals. In Section 4, we discuss the well-posedness of the underlying systems. In Section 5, we discuss the convergence of the iteration for unconstrained controls, while the constrained case is considered in Section 6, however, for the linear case only. As the nonlinearities are not assumed to be explicitly given by specific functions and the corresponding Nemytskij operators, we will rely on bounds, regularity, and smallness assumptions and use the control structure to compensate for nonlinear effects. Finally, the paper closes with some concluding remarks in Section 7.

2. Problem statement

Let $y(t, x) \in \mathbb{R}^d$, $t \in [0, T]$, $x \in [0, L]$, denote the state and let

$$\Lambda(x) = \operatorname{diag}\left(\lambda_1(x), \dots, \lambda_m(x), \lambda_{m+1}(x), \dots, \lambda_d(x)\right) \in \mathbb{R}^{d\times d}$$

with

$$\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_m(x) > 0 > \lambda_{m+1}(x) \geq \cdots \geq \lambda_d(x)$$

for all $(x) \in [0, L]$ represent the physics of the system, taken in characteristic coordinates to make the mathematical description simpler. We use the block-matrix abbreviation

$$\lambda(x) = \operatorname{diag}\left(\Lambda^+(x), \Lambda^-(x)\right),$$

with

$$\Lambda^+(x) := \operatorname{diag}(\lambda_1(x), \ldots, \lambda_m(x)) \quad \text{and} \quad \Lambda^-(x) := \operatorname{diag}(\lambda_{m+1}(x), \ldots, \lambda_d(x)).$$

Accordingly, we denote the first m components of the state by y^+ and the remaining $d - m$ components by y^-, so that $y = (y^+, y^-)$. In order to describe our boundary conditions, we introduce a block matrix K as follows

$$K := \begin{bmatrix} K^{00} & K^{01} \\ K^{10} & K^{11} \end{bmatrix}, K^{00} \in \mathbb{R}^{m \times m}, K^{01} \in \mathbb{R}^{m \times d-m}, K^{10} \in \mathbb{R}^{d-m \times m}, K^{11} \in \mathbb{R}^{d-m \times d-m}. \tag{1}$$

Referring to this block structure and for the sake of simplicity, we introduce boundary controls u for negative components only. The corresponding input operator is given by

$$B := \begin{pmatrix} 0 \\ B_b \end{pmatrix},$$

while B_d signifies the input operator for distributed controls v. We consider mixed two-point boundary value problems for systems of hyperbolic semilinear equations of the form

$$\partial_t y + \Lambda(x)\partial_x y + My = f(y) + B_d u, \qquad (t, x) \in (0, T) \times (0, L), \tag{2a}$$

$$\begin{pmatrix} y^+(t, 0) \\ y^-(t, L) \end{pmatrix} = K \begin{pmatrix} y^+(t, L) \\ y^-(t, 0) \end{pmatrix} + Bv, \qquad t \in (0, T), \tag{2b}$$

$$y(0, x) = y_0(x), \qquad x \in (0, L), \tag{2c}$$

$$u(t) \in U_{\text{ad}}^{\text{d}}, \qquad \text{a.e. in } (0, T), \tag{2d}$$

$$v(t) \in U_{\text{ad}}^{\text{b}}, \qquad \text{a.e. in } (0, T). \tag{2e}$$

We have introduced a matrix $M \in \mathbb{R}^{d \times d}$ to allow for a linear coupling on the distributed PDE-level, which is useful if one resorts to linear systems of equations or to those with small nonlinearities. The functions f_j, $j = 1, \ldots, d$, are differentiable and satisfy Lipschitz conditions, to be specified below. Moreover, u and v are taken to represent the boundary and distributed controls, respectively, where $u(t, x) \in \mathbb{R}^d$, $v(t) \in \mathbb{R}^m$, are constrained a. e. by closed and convex sets U_{ad}^{d} and U_{ad}^{b}. If $U_{\text{ad}}^{\text{d}} = \mathbb{R}^d$, $U_{\text{ad}}^{\text{b}} = \mathbb{R}^m$, then controls are unrestricted. Finally, $y_0(x) \in \mathbb{R}^d$ for $x \in [0, L]$ denotes the initial data. We should note that the boundary conditions in (2) are in accordance with the standard formulation as in Chapter 6 of Bastin and Coron (2016). Under these conventions, System (2) is a controlled hyperbolic and semilinear system. In addition to (2), we consider the natural tracking-type cost function

$$
\begin{aligned}
J(w,y) := \frac{\kappa}{2}\int_0^L \|y(T)-y_t^d\|\,\mathrm{d}x + \frac{\mu}{2}\int_0^T\int_0^L \|y-y_d\|^2\,\mathrm{d}x\,\mathrm{d}t \\
+\frac{\nu}{2}\int_0^T\int_0^L \|u\|^2\,\mathrm{d}x\,\mathrm{d}t + \frac{\rho}{2}\int_0^T \|v(t)\|^2\,\mathrm{d}t
\end{aligned}
\tag{3}
$$

with $w=(u,v)$. The considered control problem is thus given by

$$
\min_{w,y} \quad J(w,y) \quad \text{s.t.} \quad (w,y) \text{ satisfies (2).} \tag{4}
$$

It is a matter of standard variations to derive the adjoint system from the Lagrangian function

$$
\mathcal{L}(w,y,p) = J(w,y) + \sum_{i=1}^d \int_0^T\int_0^L (\partial_t y + \Lambda(x)\partial_x y + My - f(y) - B_d u)_i\, p_i\,\mathrm{d}x\,\mathrm{d}t.
$$

The details are left to the reader. We obtain the following optimality conditions, governing the adjoint variable p:

$$
\begin{aligned}
\partial_t p + \Lambda(x)\partial_x p(\partial_x\Lambda(x) - M^\top p + f'(y)^\top)p = \mu(y-y^d), \qquad & (t,x)\in(0,T)\times(0,L),\\
\begin{pmatrix} p^+(t,L)\\ p^-(t,0)\end{pmatrix} = \tilde{K}\begin{pmatrix} p^+(t,0)\\ p^-(t,L)\end{pmatrix}, \qquad & t\in(0,T),\\
p(T,x) = -\kappa(y(T)-y_t^d), \qquad & x\in(0,L).
\end{aligned}
\tag{5}
$$

Here, the boundary matrix $\tilde{K}$ is given by

$$
\tilde{K} := \operatorname{diag}(\Lambda^+(L)^{-1}, |\Lambda^-(0)^{-1}|)K^\top \operatorname{diag}(\Lambda^+(0), |\Lambda^-(L)|), \tag{6}
$$

where we used the absolute value signs to indicate the absolute values of the diagonal entries. By taking the directional derivative of $\mathcal{L}(w,y,p)$ w.r.t. u in the direction $\hat{u}$ we obtain

$$
\int_0^T\int_0^L \left(\nu v - B_d^\top p\right)(\hat{v}-v)\,\mathrm{d}x\,\mathrm{d}t \ge 0 \quad \text{for all } \hat{v}\in U_{\mathrm{ad}}^{\mathrm{d}}, \tag{7}
$$

and, similarly, for v we get

$$
\int_0^T \left(\rho(t)u - B_b^\top|\Lambda^-(L)|p^-(t,L)\right)(\hat{u}(t)-u(t))\,\mathrm{d}t \ge 0 \quad \text{for all } \hat{u}\in U_{\mathrm{ad}}^{\mathrm{b}}. \tag{8}
$$

Obviously, in the case of no constraints on the control, the optimality conditions (7) and (8) reduce to

$$v(t) = \frac{1}{\rho} B_b^\top |\Lambda^-(L)| p(t, L), \qquad t \in (0, T), \tag{9}$$

$$u(t) = \frac{1}{\nu} B_d^\top p, \qquad (t, x) \in (0, T) \times (0, L). \tag{10}$$

EXAMPLE 1 (A NETWORK OF STRINGS) *We consider a star-graph consisting of m strings or rods connected at a multiple node located at $x = 0$. The individual strings are stretched along an interval $[0, L]$. Each string is represented by a displacement $w_i(t, x)$ for $x \in [0, L]$ and $t \in [0, \infty)$. Indeed, we assume that there is a spatio-temporal axial loading $c_i(t, x)$. These strings or rods form a network located in the plane and $w_i(t, x)$ is either the out-of-the-place displacement of the i^{th} string or the longitudinal displacement of the i^{th} rod. We assume that the strings (or rods) satisfy a semilinear damped wave equation such that at $x = 0$, the displacements are equal for all times and the sum of forces is 0. At the simple nodes, i.e., at $x = L$, the strings $i = 2, \dots, m$ are subject to dissipative controlled boundary conditions, while string $i = 1$ is clamped. The corresponding system can be written down as*

$$\partial_{tt} w_i - \partial_x (c_i \partial_x w_i) + b_i(\partial_t w_i)) = B_d u in\ (0, T) \times (0, L),\ i = 1, \dots, m, \tag{11a}$$

$$w_i(t, 0) = w_j(t, 0) in\ (0, T),\ i, j = 1, \dots, m, \tag{11b}$$

$$\sum_{i=1}^{m} c_i(t, 0) \partial_x w_i(t, 0) = 0 in\ (0, T), \tag{11c}$$

$$w_m(t, L) = 0, in\ (0, T), \tag{11d}$$

$$\partial_x w_i(t, L) = v_i(t) in\ (0, T),\ i = 2, \dots, m, \tag{11e}$$

$$w_i(0, x) = w_{i0}(x) in\ (0, L),\ i = 1, \dots, m, \tag{11f}$$

$$\partial_t w_i(0, x) = w_{i1}(x) in\ (0, L),\ i = 1, \dots, m. \tag{11g}$$

We now transform (11) *into the format of* (2). *In a first step, we transform* (11a) *into a* 2×2*-system, assuming, for simplicity, $c_i(x) = c_i$. To this end, we set*

$$z_{i1} := \frac{1}{2} \left(\partial_t w_i - \sqrt{c_i} \partial_x w_i \right), \quad z_{i2} := \frac{1}{2} \left(\partial_t w_i + \sqrt{c_i} \partial_x w_i \right).$$

Hence

$$\partial_t w_i = (z_{i1} + z_{i2}), \quad \partial_x w_i = \frac{1}{\sqrt{c_i}} (z_{i2} - z_{i1})$$

and, therefore,

$$\partial_t \begin{pmatrix} z_{i1} \\ z_{i2} \end{pmatrix} + \begin{bmatrix} \sqrt{c}_i & 0 \\ 0 & -\sqrt{c}_i \end{bmatrix} \partial_x \begin{pmatrix} z_{i1} \\ z_{i2} \end{pmatrix} = -b_i \left(\frac{1}{2} (z_{i1} + z_{i2}) \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

We define

$$y_i = z_{i1}\ for\ i = 1, \dots, m, \quad y_i = z_{(i-m)2}\ for\ i = m + 1, \dots, d.$$

For the sake of simplicity, we assume that the tensions are equal at $x = 0$ for all times; i.e., $c_i(t,0) = c_j(t,0)$ holds for all $t \in [0,T]$. Then, the transmission conditions (11b) *and* (11c) *can be equivalently formulated as*

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}(t,0) = -\frac{1}{m}\begin{bmatrix} m-2 & -2 & \cdots & -2 \\ -2 & m-2 & \cdots & -2 \\ \vdots & \vdots & \ddots & \vdots \\ -2 & -2 & \cdots & m-2 \end{bmatrix}\begin{pmatrix} y_{m+1} \\ \vdots \\ y_d \end{pmatrix}(t,0). \tag{12}$$

Notice that without the assumption on the axial loads at $x = 0$, the matrix on the right-hand side becomes non-symmetric, which, in turn, is not a problem in principle. We introduce the matrix S such that

$$(S\varphi)_i := \left(\frac{2}{m}\sum_{j=1}^{m}\varphi_j - \varphi_i\right).$$

Thus, (12) *reads as $Y^+(0) = SY^-(0)$. The matrix S has nice properties. It can be interpreted as a scattering matrix. In particular,*

$$\sum_{i=1}^{m}(S\varphi)_i = \sum_{i=1}^{m}\varphi_i \quad \textit{and} \quad SS\varphi = \varphi$$

holds. At $x = L$, we have, at least formally, for sufficiently regular states,

$$\partial_t w_m(t,L) = 0 \implies z_{11}(t,L) + z_{12}(t,L) = 0 \implies y_m(t,L) = -y_d(t,L)$$

for the clamped string and

$$\partial_x w_i(t,L) = v_i(t) \implies z_{i2}(t,L) - z_{i1}(t,L) = \sqrt{c_i}v_i(t)$$

for the other strings. This provides the boundary conditions at the end $x = L$:

$$\begin{aligned} y_m(t,L) &= -y_d(t,L), \\ y_i(t,L) &= y_{m+i}(t,L) + h_i(t, 2y_{m+i}(t,L) + v_i(t)), \ \ i = 1,\dots,m-1. \end{aligned}$$

Thus, with $K^{00} = 0, K^{11} = 0$, as well as $K^{01} = S$ and $K^{10} = \operatorname{diag}(-1,\dots,-1)$, we obtain the boundary condition

$$\begin{pmatrix} y^+(t,0) \\ y^-(t,L) \end{pmatrix} = \begin{bmatrix} 0 & K^{01} \\ K^{10} & 0 \end{bmatrix}\begin{pmatrix} y^+(t,L) \\ y^-(t,0) \end{pmatrix}.$$

Thus, our example is of the format of (2)*, however, with separated boundary conditions.*

Remark 1 *We remark that there are many more examples—in particular for systems of semilinear hyperbolic balance laws on metric graphs—that exactly fit into this framework. These are, e.g., networks of open channels with the dynamics governed by the shallow water equations with wall friction (see, e.g., Leugering and Schmidt, 2002), or networks of gas pipelines (see Leugering et al. 2017). Moreover, networks of semilinear Timoshenko beams (se Leugering and Rodriguez, 2020) can be written in the framework of* (2) *as well.*

EXAMPLE 2 (A NETWORK OF STRINGS WITH A CYCLE) *Here, we consider the simplest network that contains a cycle, namely a triangle. We take the same model equations as in the example above, now for three strings:*

$$
\begin{aligned}
&\partial_{tt} w_i - \partial_{xx} w_i + b_i(\partial_t w_i) = B_d u && \text{in } (0,T)\times(0,L),\ i = 1,2,3,\\
&w_1(t,0) = w_3(t,L) && \text{in } (0,T),\\
&w_1(t,L) = w_2(t,0) && \text{in } (0,T),\\
&w_2(t,L) = w_3(t,0) && \text{in } (0,T),\\
&\partial_x w_1(t,0) = \partial_x w_3(t,L) + u_1(t)\ \partial_x w_1(t,L) = \partial_x w_2(t,0) + u_2(t), &&\\
&\partial_x w_2(t,L) = \partial_x w_3(t,0) && \text{in } (0,T),\\
&w_i(0,x) = w_{i0}(x) && \text{in } (0,L),\ i = 1,2,3,\\
&\partial_t w_i(0,x) = w_{i1}(x) && \text{in } (0,L),\ i = 1,2,3.
\end{aligned}
$$

In this case it is easily deduced by the same analysis as in the first example that the boundary conditions now are given by

$$
\begin{pmatrix} y^+(t,0) \\ y^-(t,L) \end{pmatrix} = \begin{bmatrix} K^{00} & 0 \\ 0 & K^{11} \end{bmatrix} \begin{pmatrix} y^+(t,L) \\ y^-(t,0) \end{pmatrix} + B_b \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},
$$

where now

$$
K^{00} := \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad K^{11} := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad K := \operatorname{diag}(K^{00}, K^{11})
$$

$$
B_b := \frac{1}{2}\begin{bmatrix} B_b^+ \\ B_b^- \end{bmatrix}, \qquad B_b^+ := \begin{bmatrix} -1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B_b^- := \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ -1 & 0 \end{bmatrix}.
$$

It is remarkable that $K^3 = id$, which reflects the fact that the longest path has the length of 3.

3. Time-domain decomposition

We now embark on time-domain decomposition of the optimal control problem by decomposing the optimality system (2), (5)–(8). The procedure pursued in this article is the same as in Krug et al. (2021) and is, in turn, very much inspired by Lagnese and Leugering (2003), where the linear wave equation and the Maxwell equations are considered. As discussed at the beginning, the novelty here is the more general boundary condition and the well-posedness analysis. We introduce a coarse time discretization so that we obtain $[T_k, T_{k+1}] \times (0,L)$ with

$$
0 = T_0 < T_1 < \cdots < T_k < T_{k+1} < \cdots < T_K < T_{K+1} = T.
$$

We now formulate the time-domain decomposition procedure for the general optimality system (2), (5)–(8) and then focus on the case of unconstrained controls u, v

and quadratic costs as in (3) for a proof of convergence in the subsequent sections. We denote the restrictions of y_k, p_k, u_k, v_k to $Q_k := I_k \times (0, L)$ with $I_k := (T_k, T_{k+1})$ by

$$y_k := y|_{Q_k}, \quad p_k := p|_{Q_k}, \quad u_k := u|_{Q_k}, \quad v_k := v|_{I_k}.$$

The idea is to satisfy the continuity conditions

$$\begin{aligned} y_k(T_{k+1}) &= y_{k+1}(T_{k+1}), \quad k = 0, \ldots, K-1, \\ p_k(T_{k+1}) &= p_{k+1}(T_{k+1}), \quad k = 0, \ldots, K-1, \end{aligned} \tag{13}$$

in the limit of an iterative procedure. We therefore use the decoupling

$$\begin{aligned} y_k^{n+1}(T_{k+1}) + \beta p_k^{n+1}(T_{k+1}) &= \phi_{k,k+1}^n, \quad k = 0, \ldots, K-1, \\ y_k^{n+1}(T_k) - \beta p_k^{n+1}(T_k) &= \phi_{k,k-1}^n, \qquad k = 1, \ldots, K, \end{aligned} \tag{14}$$

together with the update rule

$$\begin{aligned} \phi_{k,k+1}^n &= (1-\varepsilon)\left(y_{k+1}^n(T_{k+1}) + \beta p_{k+1}^n(T_{k+1})\right) + \varepsilon\left(y_k^n(T_{k+1}) + \beta p_k^n(T_{k+1})\right), k = 0, ..., K-1, \\ \phi_{k,k-1}^n &= (1-\varepsilon)\left(y_{k-1}^n(T_k) - \beta p_{k-1}^n(T_k)\right) + \varepsilon\left(y_k^n(T_k) - \beta p_k^n(T_k)\right), k = 1, \ldots, K, \end{aligned} \tag{15}$$

and $\beta > 0$, $0 \le \varepsilon < 1$, $n = 0, 1, 2, \ldots$

Remark 2 *Suppose that the iteration* (14), (15) *converges for $n \to \infty$, where y_k, p_k, u_k, v_k solve* (2), (5)–(8) *on Q_k. Then,* (14) *holds without iteration indices n and $n+1$. As a result, the iteration updates* (15) *and the decoupling* (14) *reduce to*

$$\begin{aligned} (1-\varepsilon)\left(y_k(T_{k+1}) + \beta p_k(T_{k+1})\right) &= (1-\varepsilon)\left(y_{k+1}(T_{k+1}) + \beta p_{k+1}(T_{k+1})\right), \\ (1-\varepsilon)\left(y_k(T_k) - \beta p_k(T_k)\right) &= (1-\varepsilon)\left(y_{k-1}(T_k) - \beta p_{k-1}(T_k)\right), \end{aligned}$$

where we may divide by $(1-\varepsilon)$ and shift the second equation by $k \to k+1$ to obtain

$$\begin{aligned} y_k(T_{k+1}) + \beta p_k(T_{k+1}) &= y_{k+1}(T_{k+1}) + \beta p_{k+1}(T_{k+1}), \\ y_{k+1}(T_{k+1}) - \beta p_{k+1}(T_{k+1}) &= y_k(T_{k+1}) - \beta p_k(T_{k+1}). \end{aligned}$$

Addition of the last two equations leads to

$$y_k(T_{k+1}) = y_{k+1}(T_{k+1}), \quad p_k(T_{k+1}) = p_{k+1}(T_{k+1}).$$

Thus, (13) *is satisfied and in the limit and the continuity conditions hold. Therefore, the non-overlapping domain decomposition* (14), (15) *appears reasonable.*

In view of this remark, we propose the time-domain decomposition

$$\begin{aligned} \partial_t y_k^n + \Lambda(x)\partial_x y_k^n + M y_k^n &= f_k(y_k^n) + B_d u_k^n, & (t,x) &\in Q_k, \\ \partial_t p_k^n + \Lambda(x)\partial_x p_k^n + (\partial_x\Lambda(x) - M^T)p_k^n &= (y_k^n - y_k), & (t,x) &\in Q_k, \\ \begin{pmatrix} y_k^{+,n}(t,0) \\ y^{-,n}(t,L) \end{pmatrix} &= K \begin{pmatrix} y^{+,n}(t,L) \\ y^{-,n}(t,0) \end{pmatrix} + B_b v_k^n(t), & t &\in I_k, \\ \begin{pmatrix} p^{+,n}(t,L) \\ p^{-,n}(t,0) \end{pmatrix} &= \tilde{K} \begin{pmatrix} p^{+,n}(t,0) \\ p^{-,n}(t,L) \end{pmatrix}, & t &\in I_k, \end{aligned} \tag{16}$$

of system (5)–(8) together with (7), (8) and (14), (15), which have to be extended by

$$y_0^n(0,\cdot) = y_0 \quad \text{and} \quad p_K^n(T_{K+1},\cdot) = 0.$$

For $k = 1,\ldots,K-1$, we now introduce the so-called virtual controls $g_{k,k-1}(x)$ for $x \in (0,L)$ in the sense of Lagnese and Leugering (2003, 2004) and consider the following virtual control problem on Q_k:

$$\begin{aligned}
\min_{g_{k,k-1},w_k,y_k} \quad & J_k(w_k,y_k) + \frac{1}{2\beta}\left(\|y_k(T_{k+1}) - \phi_{k,k+1}\|^2 + \|g_{k,k-1}\|^2\right) && \\
\text{s.t.} \quad & \partial_t y_k + \Lambda(x)\partial_x y_k + My_k = f_k(y_k) + B_d u, && (t,x) \in Q_k, \\
& \begin{pmatrix} y_k^+(t,0) \\ y^-(t,L) \end{pmatrix} = K \begin{pmatrix} y^+(t,L) \\ y^-(t,0) \end{pmatrix} + B_b v_k(t), && t \in I_k, \\
& y(T_k,x) = \phi_{k,k-1} + g_{k,k-1}, && x \in (0,L), \\
& u(t) \in U_{\text{ad}}^{\text{d}} && \text{a.e. in } I_k, \\
& v(t) \in U_{\text{ad}}^{\text{b}} && \text{a.e. in } I_k,
\end{aligned} \tag{17}$$

where $J_k(w_k,y_k)$ is given by (3), restricted to Q_k. Suppose that the controls $g_{k,k-1}$, u_k, v_k, and the state y_k are optimal for the virtual control problem (17). Then, formally, the corresponding optimality system for y_k, p_k is such that y_k, p_k satisfy (7), (8) and (14), (15), (16). In particular, the optimal virtual control $g_{k,k-1}$ is given by

$$g_{k,k-1} = \beta p_k(T_k).$$

The proof of this statement is straightforward and is, therefore, left to the reader.

REMARK 3 *The virtual control problems for $k = 1,\ldots,K-1$ have to be complemented by a corresponding problem for $k = 0$ and $k = K$, respectively. Clearly, for $k = 0$ no additional virtual control is needed as*

$$y_0(T_0) = y_0(0) = y(0) = y_0$$

is given data, while at $k = K$,

$$p_K(T_{K+1}) = p_K(T) = p(T)$$

is prescribed and, therefore, no penalty term for the upper transmission condition is needed.

As for the existence of optimal controls for problem (4), we refer to the next section.

4. Well-posedness

4.1. Strong and mild solutions

The semilinear format of (2) provides the possibility of handling questions concerning well-posedness of both the forward and the backward problem in a unifying manner.

To this end, we define the first-order differential expression $\mathcal{A}_0$ by

$$\mathcal{A}_0 y := -\Lambda(x) d_x y - My.$$

Let $\mu \geq 0$ be a weight and introduce the weight function

$$\gamma_\mu(x) = \begin{cases} e^{-\mu x}, & i = 1, \ldots, m, \\ e^{\mu(x-1)}, & i = m+1, \ldots, d. \end{cases}$$

We can then introduce a weighted Hilbert space

$$L^2_\mu((0,L))^d =: L^2(0,L;\mathbb{R}^d;\mu) =: \mathcal{H}_\mu,$$

such that with $\Phi = (\phi_i)_{i=1}^d$ and $\Psi = (\psi_i)_{i=1}^d$,

$$(\Phi,\Psi)_\mu := \sum_{i=1}^m \int_0^L \phi_i \psi_i e^{\mu(x-L)} + \sum_{i=m+1}^d \int_0^L \phi_i \psi_i e^{-\mu(x)}, \ \mu > 0$$

is finite, which is isomorphic to the standard space $L^2((0,L))^d$. Now, we can write

$$D(\mathcal{A}) := \left\{ y \in \mathcal{H}_\mu \colon \mathcal{A}_0 y \in \mathcal{H}_\mu, \begin{pmatrix} y^+(0) \\ y^-(L) \end{pmatrix} = K \begin{pmatrix} y^+(L) \\ y^-(0) \end{pmatrix} \right\}$$

and integration by parts in the weighted space as in Bastin and Coron (2016). The adjoint operator $\mathcal{A}^*$ is given by

$$\mathcal{A}^* \varphi = \Lambda(x) d_x \varphi + \tilde{M} \varphi$$

with

$$\begin{aligned} \tilde{M} &:= -D(x)^{-1} M^T D(x) + \mu \operatorname{diag}(\Lambda^+(x), |\Lambda^-(x)|), \\ D(x) &:= \operatorname{diag}(D^+(x), D^-(x)), \\ D^+(x) &:= \operatorname{diag}(\exp(\mu(L-x)), i = 1, \ldots m), \\ D^-(x) &:= \operatorname{diag}(\exp(-\mu x), i = m+1, \ldots d), \end{aligned}$$

and

$$D(\mathcal{A}^*) := \left\{ \varphi \in \mathcal{H}_\mu \colon \mathcal{A}^* \in \mathcal{H}_\mu, \begin{pmatrix} \varphi^+(L) \\ \varphi^-(0) \end{pmatrix} = \exp(-\mu L) \tilde{K} \begin{pmatrix} \varphi^+(0) \\ \varphi^-(L) \end{pmatrix} \right\},$$

where $\tilde{K}$ is given by (6). According to Appendix A in Bastin and Coron (2016), $\mathcal{A}$ and $\mathcal{A}^*$ are densely defined and quasi-dissipative in $\mathcal{H}_\mu$ for some $\mu > 0$. Hence, according to the Lumer–Phillips Theorem, $\mathcal{A}$ (and, accordingly, $\mathcal{A}^*$) generates a C_0-semi-group $T(t)$ in $\mathcal{H}_\mu$ and, thus, the homogeneous problem (2) with $f \equiv 0$, $u \equiv 0$, and $v \equiv 0$ admits a unique solution also in the case of non-vanishing u.

Remark 4 *It is important to note that one can easily shift the generator of the semi-group in order to also get a semi-group in the space without the weight function. Consequently, we will omit the dependence on μ in the sequel.*

In order to handle the boundary control, we may apply the standard shifting method to move the non-homogeneous boundary condition involving the control to the state equation at the cost, however, of extra regularity to be assumed for the control. To this end, we introduce a steady-state solution

$$c(x) := \begin{pmatrix} c^+(x) \\ c^-(x) \end{pmatrix}$$

of

$$\Lambda \partial_x c + Mc = 0, \begin{pmatrix} c^+(0) \\ c^-(L) \end{pmatrix} = K \begin{pmatrix} c^+(L) \\ c^-(0) \end{pmatrix} + B_b, \quad y = z + cv,$$

where u satisfies

$$u \in H^1(0,T) \quad \text{and} \quad u(0) = u(T) = 0. \tag{18}$$

Then, z has to solve

$$\partial_t z + \Lambda(x)\partial_x z + Mz = f(z) + B_d vc \frac{\mathrm{d}}{\mathrm{d}t} v, \qquad (t,x) \in (0,T) \times (0,L), \tag{19a}$$

$$\begin{pmatrix} z^+(t,0) \\ z^-(t,L) \end{pmatrix} = K \begin{pmatrix} z^+(t,L) \\ z^-(t,0) \end{pmatrix} + Bv, \qquad t \in (0,T), \tag{19b}$$

$$z(0,x) = y_0(x), \qquad x \in (0,L). \tag{19c}$$

We are then left with a semilinear PDE in the context of C_0 semi-groups—a case well studied in the literature. We refer, in particular, to Vrabie (2003); see Theorem 1.11.1 and Corollary 11.3.1. Indeed, we assume that there is a neighborhood $D \subset \mathcal{H}$ such that

$$f \colon D \to \mathcal{H} \text{ is continuous and locally Lipschitz.}$$

Under this condition, there is an analogue to the concept of semi-global classical solutions as introduced by Li Tatsien even for quasi-linear equations. The result says that starting in D, there is a time $T > 0$ such that (19) admits a unique solution, which, in turn, is given by the generalized variation of constants formula for (19), i.e.,

$$y(t,\cdot) = z(t,\cdot) + cv(t), \quad z(t,\cdot) = T(t)y_0 + \int_0^t T(t-s)\left(f(z+cv) + B_d u - \frac{\mathrm{d}}{\mathrm{d}t} v\right) \mathrm{d}s. \tag{20}$$

In particular, in order to handle our applications, we assume that f is in fact a Nemytskij operator in the sense of, e.g., Section 1.3 of Roubicek (2005):

$$\text{(i)} \quad f = (f_i)_i \text{ is a Caratheodory mapping,} \tag{21}$$

$$\text{(ii)} \quad |f_i(r_1, \dots, r_d)| \le \gamma_i c_i \sum_{j=1}^{d} |r_j|^\kappa \quad \text{for all } i = 1, \dots, d. \tag{22}$$

Then, according to Theorem 1.27 in Roubicek (2005), f is bounded and continuous from the space $(L^{2\kappa}(0,L))^d$ into $L^2(0,L)^d$. In addition, we will assume homogeneity of degree κ and strict dissipativity of f, i.e.,

$$\text{(iii)} \quad f(\lambda y) = \lambda^\kappa f(y), \tag{23}$$

$$\text{(iv)} \quad (f(v), v)_\mu \leq -C\|v\|^{\kappa+1}. \tag{24}$$

Example 3 *In our applications, we typically have the structure*

$$f_i(y) = -b(y_i^+ + y_i^-)|y_i^+ + y_i^-| \quad \textit{for} \quad i = 1\ldots, d \quad \textit{with} \quad \beta(s) := |s|^\rho s.$$

In this case, $\kappa = 2$ *holds in* (21) *and, hence, the mapping* $f : L^4(0,L)^d \to \mathcal{H}$ *is bounded and continuous. Therefore, with Assumption* (18) *and initial data in* $L^4(0,L)^d$*, we obtain a unique solution* y*.*

4.2. Weak solutions

With the assumptions on v, the strong formulation (19) can be handled with the classic semi-group properties of $T(t)$ and integration by parts w.r.t. time in the term involving $\partial_t v_i$ to achieve well-posedness of (19) and, hence, of (2). The price to pay is that the solutions have to leave the domain of the generator of the semi-group. This necessitates a concept of weak solutions. Of course, one can work with the dual of the domain for very weak solutions. The other natural way is to work with variational arguments and a priori estimates.

To this end, we multiply the state equation (2) by $\phi \in D(\mathcal{A})^*$, such that $\mathcal{A}^*\phi = -w \in L^{\kappa+1}(0,L)^d$ holds. Then

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int_0^L y(t)\cdot\phi\,\mathrm{d}x + \int_0^L y(t)\cdot w\,\mathrm{d}x &= \int_0^L f(y(t))\cdot\phi\,\mathrm{d}x \\ &+ \int_0^L u\cdot B_d^T\phi\,\mathrm{d}x - B_b^T|\Lambda^-(L)|\phi^-(t,L)\cdot v(t),\ t(0,T). \end{aligned} \tag{25}$$

We now integrate w.r.t. time and obtain the format of weak solutions:

$$\begin{aligned} \int_0^L y(t)\cdot\phi\,\mathrm{d}x + \int_0^t\int_0^L y(t)\cdot w\,\mathrm{d}x\,\mathrm{d}s &= \int_0^L y_0\cdot\phi\,\mathrm{d}x \int_0^t\int_0^L f(y(t))\cdot\phi\,\mathrm{d}x\,\mathrm{d}s \\ &+ \int_0^t\int_0^L u\cdot B_d^T\phi\,\mathrm{d}x\,\mathrm{d}s - \int_0^t B_b^T|\Lambda^-(L)|\phi^-(s,L)\cdot v(s)\,\mathrm{d}s,\ t\in(0,T). \end{aligned} \tag{26}$$

The idea is to approximate the data needed in the weak formulation in (26) by those corresponding to strong solutions discussed in the previous subsection and pass to the limit in (26). To this end, we need some a priori estimate.

Lemma 1 *For solutions y of* (2)*, we find a suitable constant $C > 0$ such that*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|y(t)\|^2 + (f(y(t), y(t)) \leq C\left(\|B_d u(t)\|\|y(t)\| + \|v(t)\|\right).$$

Proof We use arguments similar to the case of homogeneous boundary conditions of the previous subsection in the context of the generation result. To this end, we go back to the weighted spaces. Indeed,

$$\begin{aligned}
&(-\Lambda\partial_x\phi - M\phi, \phi)_\mu \\
&= -\int_0^L \sum_{i=1}^m \lambda_i\partial_x\phi_i\phi_i e^{\mu(L-x)} - \sum_{i=m+1}^d \lambda_i\partial_x\phi_i\phi_i e^{-\mu x}\,\mathrm{d}x - (M\phi,\phi)_\mu \\
&= -\frac{1}{2}\sum_{i=1}^m \left(\lambda_i(L)|\phi_i(L)|^2 - \lambda_i(0)|\phi_i(0)|^2 e^{-\mu L}\right) \\
&\quad + \frac{1}{2}\sum_{i=m+1}^d \left(\lambda_i(L)|\phi_i(L)|^2 e^{-\mu L} - \lambda_i(0)|\phi_i(0)|^2\right) \\
&\quad + \frac{1}{2}\sum_{i=1}^m \int_0^L |\phi_i|^2(|\lambda_i|e^{\mu(L-x)})_x\,\mathrm{d}x \\
&\quad - \frac{1}{2}\sum_{i=m+1}^d \int_0^L |\phi_i|^2(|\lambda_i|e^{-\mu(x)})_x\,\mathrm{d}x - (M\phi,\phi)_\mu \\
&=: \frac{1}{2}(B+I).
\end{aligned}$$

Due to the Lipschitz-property of the λ_i and the fact that they are bounded away from zero (say, by $4\delta > 0$) we can bound B as

$$B \leq -2\delta\left(\sum_{i=1}^m |\phi_i(L)|^2 + \sum_{i=m+1}^d\right)\sum_{i=1}^m \lambda_i(0)|\phi_i(L)|^2 e^{-\mu L} + \sum_{i=m+1}^d |\lambda_i(L)||\phi_i(L)|^2 e^{-\mu L}, \tag{27}$$

whereas I satisfies

$$I \leq \left(\max_{i\in\{1,\ldots,d\}}\sup_{x\in(0,L)}|\lambda_i|\mu + \max_{i=1}^d\|\lambda_i\|\right)\|\phi\|_\mu^2 + \|M\|\|\phi\|_\mu^2 \leq 2\omega\|\phi\|_\mu^2. \tag{28}$$

We have

$$\begin{aligned}
&\sum_{i=1}^m \lambda_i(0)|\phi_i(0)|^2 e^{-\mu L} \\
&\leq\; 2e^{-\mu L}\left(\left(\sum_{i=1}^m\sum_{j=1}^m \lambda_i(0)(K_{ij}^{00})^2\right)\sum_{i=1}^m |\phi_j(l)|^2 + \left(\sum_{i=1}^m\sum_{j=m+1}^d \lambda_i(0)(K_{ij}^{01})^2\right)\sum_{i=m+1}^d |\phi_j|^2\right)
\end{aligned} \tag{29}$$

and

$$\begin{aligned}
&\sum_{i=m+1}^{d} |\lambda_i(L)||\phi_i(L)|^2 e^{-\mu L} \\
&\le 2e^{-\mu L}\left(\left(\sum_{i=m+1}^{d}\sum_{j=1}^{m}|\lambda_i(L)|(K_{ij}^{10})^2\right)\sum_{j=1}^{m}|\phi_j|^2\right. \\
&\quad \left.+\left(\sum_{i=m+1}^{d}\sum_{j=m+1}^{d}|\lambda_i(L)|(K_{ij}^{11})^2\right)\sum_{i=m+1}^{d}|\phi_j(0)|^2+\sum_{i=m+1}^{d}|\lambda_i(L)|b_i^2 v_i^2\right).
\end{aligned} \tag{30}$$

If we choose μ large enough, we can compensate the quadratic terms in $\phi_i(L), i = 1, \ldots, m$, and $\phi_i(0), i = m+1, \ldots, d$, in (29) and (30) for the estimation of (27) to obtain

$$B < -\delta\left(\sum_{i=1} m|\phi_i(L)|^2 + \sum_{i=m+1}^{d}|\phi(0)|^2\right) + e^{-\mu L}\sum_{i=m+1}^{d}|\lambda_i(L)|b_i^2 v_i^2. \tag{31}$$

Then, putting (31) and (28) together, we arrive at

$$(-\Lambda\partial_x\phi - M\phi, \phi)_\mu \le e^{-\mu L}\sum_{i=m+1}^{d}|\lambda_i(L)|^2 b_i^2 v_i^2$$

for ϕ satisfying the boundary conditions in (2). If we define $v := e^{-\omega t}y$, we get

$$\begin{aligned}
\partial_t v + \Lambda\partial_x v + Mv &= e^{\omega(\kappa-1)t}f(v) + B_d u, \\
\begin{pmatrix} v^+(t,0) \\ v^-(t,L)\end{pmatrix} &= K\begin{pmatrix} v^+(t,L) \\ v^-(t,0)\end{pmatrix} + B_b\tilde{v}, \\
v(0) &= y_0
\end{aligned} \tag{32}$$

with $\tilde{v}(t) := e^{-\omega}tv$. If we now multiply (32) by v w.r.t. the weighted inner product, use (31), and the Property (iv) in (24), we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|v\|_\mu^2 + |(f(v), v)_\mu| \le (B_d u, v)_\mu + e^{-\mu L}|\Lambda^-(L)||bv|. \tag{33}$$

Finally, we use the fact that the weighted norm and the standard norm are equivalent in order to arrive at the assertion of the lemma. □

We can now integrate (33) w.r.t. time and use the Gronwall lemma to infer that if we approximate the initial data and the controls by

$$y_0^n \to y_0 \text{ in } D(\mathcal{A}), \tag{34}$$

$$u^n \to u \text{ in } \mathcal{H}, \tag{35}$$

$$v^n \to v \text{ in } L^2(0,T)^{d-m} \tag{36}$$

and if we associate to these data the strong solution y^n, we obtain

$$(y^n)_n \text{ is bounded in } \mathcal{H},$$
$$(y^n)_n \text{ is bounded in } L^{\kappa+1}(Q_T)^d,$$

and we can then extract subsequences, still denoted by $(y^n)_n$, such that

$$y^n \rightharpoonup y \text{ weakly in } \mathcal{H},$$
$$y^n \rightharpoonup y \text{ weakly in } L^{\kappa+1}(Q_T)^d.$$

However, this alone does not imply $f(y^n) \to f(y)$ in $L^{\frac{\kappa+1}{\kappa}}(Q_T)$. Notice that with $q = (\kappa+1)/\kappa$, we have $p = \kappa + 1$. We now recall Lemma 1.3 from Lions (1971).

Lemma 2 *Let $O \subset \mathbb{R}_+ \times \mathbb{R}^d$ be an open and bounded set and let $(g^n)_n, g \in L^q(O)$, $1 < q < \infty$, such that*

$$\|g^n\|_{L^q} \leq C, \; g^n \to g \quad a.e. \text{ in } O.$$

Then $g^n \rightharpoonup g$ weakly in $L^q(O)$ as $n \to \infty$.

In our situation, $g^n := \|y^n\|^\kappa$, $q = \frac{\kappa+1}{\kappa}$, and

$$\int_0^L |g^n|^q \,\mathrm{d}x = \int_0^L \|y^n\|^{\kappa+1} \,\mathrm{d}x \leq C$$

holds. What is needed is that $g^n \to g$ a.e. in Q_T. We follow the same arguments as in Prodi (2010) (see pages 209–211) to show that we can indeed extract a strongly convergent subsequence, i.e.,

$$y^n \to y \quad \text{in } \mathcal{H},$$

which then implies that $f(y^n) \rightharpoonup f(y)$ in $L^{\frac{\kappa+1}{\kappa}}(Q_T)$. This, finally, allows us to pass to the limit in (26). We can therefore conclude with the following theorem.

Theorem 1 *Let the sequences $(y_0^n)_n \subset D(\mathcal{A})$, $(u^n)_n \in \mathcal{H}$, and $(v^n)_n$ in $L^2(0,T;\mathbb{R}^{d-m})$ be given such that* (34) *holds. Let $(y^n)_n$ be the corresponding strong solution of* (2). *Furthermore, let $\phi \in D(\mathcal{A}^*)$ be such that $\mathcal{A}^*\phi = -w \in L^{\kappa+1}$ holds. Then, we can pass to the limit in* (26) *and, hence, y is a weak solution of* (2).

With this result, we can now state our existence theorem for the optimal control problem.

Theorem 2 *The optimal control problem* (4) *can be re-framed in a reduced form as*

$$\min_w \quad J(w, y(w)) \quad s.t. \quad u \in U_{\mathrm{ad}},$$

where $y(w)$ is the mapping from the controls into y given by (26). *This problem admits a unique solution.*

Theorem 3 *For unconstrained controls $u \in L^\infty(0,T;\mathcal{H})$ and $v \in L^\infty(0,T)^{d-m}$ there exists a unique solution $(\bar{u}, \bar{v}, \bar{y})$ of problem* (4). *Moreover, there exists an adjoint state $\bar{p}$ such that $(\bar{y}, \bar{p})$ satisfies the optimality system* (2), (5), (8), (7) *for unconstrained controls and* (9) *in case of constrained controls.*

Remark 5 *1) An analogous result holds for the time-decomposed semilinear problem* (17). *2) Higher-order regularity of the boundary controls, as with* (18)*, has to be reflected in the cost function. For the pure final-value problem, the one where $\mu = 0$ in* (3)*, the control appears in the cost function only as a first-order derivative $\frac{\mathrm{d}}{\mathrm{d}t}v$; see* (20). *Thus, we may then replace the original control v by its derivative $\tilde{v} := \frac{\mathrm{d}}{\mathrm{d}t}v$, which is an $L^\infty(0,T)^{d-m}$-control.*

5. Convergence for the case of unconstrained controls

We now derive the proof of convergence for the described setting. We recall the decomposed optimality system (16), (7), (8), (14), (15) and introduce the errors

$$\tilde{y}_k^n := y_k^n - y_k, \quad \tilde{p}_k^n := p_k^n - p_k, \quad \tilde{u}_k^n := u_k^n - u_k, \quad \tilde{v}_k^n := v_k^n - v_k.$$

Then, y_k^{n+1} and p_k^{n+1} solve the semilinear problem

$$\partial_t \tilde{y}_k^{n+1} + \Lambda \partial_x \tilde{y}_k^{n+1} + M \tilde{y}_k^{n+1} = f_k(y_k^n) - f_k(y_k) + B_d \tilde{u}_k^n, \quad (t,x) \in Q_k, \tag{37a}$$

$$\partial_t \tilde{p}_k^n + \Lambda \partial_x \tilde{p}_k^n + (\Lambda_x - M^T + \mathcal{D}_y f(y_k^{n+1})^T)\tilde{p}_k^{n+1} \tag{37b}$$

$$-(\mathcal{D}_y f_k(y_k^{n+1})^\top - \mathcal{D}_y f_k(y_k)^\top)p_k = \mu \tilde{y}_k^{n+1}, \quad (t,x) \in Q_k, \tag{37c}$$

$$\begin{pmatrix} y_k^{+,n}(t,0) \\ y^{-,n}(t,L) \end{pmatrix} = K \begin{pmatrix} y^{+,n}(t,L) \\ y^{-,n}(t,0) \end{pmatrix} + B_b v_k^n(t), \quad t \in I_k, \tag{37d}$$

$$\begin{pmatrix} p^{+,n}(t,L) \\ p^{-,n}(t,0) \end{pmatrix} = \tilde{K} \begin{pmatrix} p^{+,n}(t,0) \\ p^{-,n}(t,L) \end{pmatrix}, \quad t \in I_k \tag{37e}$$

and

$$\begin{aligned} \tilde{y}_k^{n+1}(T_{k+1}) + \beta \tilde{p}_k^{n+1}(T_{k+1}) &= \tilde{\phi}_{k,k+1}^n, \qquad k = 0, \dots, K-1, \\ \tilde{y}_k^{n+1}(T_k) - \beta \tilde{p}_k^{n+1}(T_k) &= \tilde{\phi}_{k,k-1}^n, \qquad k = 1, \dots, K, \end{aligned} \tag{38}$$

together with (15).

$$\nu \tilde{u}_k^n = \tilde{p}_k^n, \quad \mu \tilde{v}_{ki}^n = |\Lambda_i(t,1)| B_b^T \tilde{p}_{ki}^n(t,1). \tag{39}$$

In order to turn (37) into a semilinear problem just in $\tilde{y}_k^n$, $\tilde{p}_k^n$, we may rewrite

$$f_k(t, y_k^n) = f_k(t, \tilde{y}_k^n + y_k), \quad \mathcal{D}_y f_k(y_k^n) = \mathcal{D}_y f_k(\tilde{y}_k^n + y_k).$$

Now, to compensate for the nonlinear terms, we make the following assumption.

Assumption 1 *There exists a constant $L > 0$ such that for each $k = 0, \dots, K$ the following holds:*

1. *The functions f_k, $\mathcal{D}_y f_k$ are bounded by L.*
2. *The following pointwise estimates are valid:*

$$\|(\mathcal{D}_y f_k(t, y_k^n(t)) - \mathcal{D}_y f_k(t, y_k(t))\tilde{y}_k^n(t)\|_{L^2} \leq L\|\tilde{y}_k^n(t)\|_{L^2},$$
$$\|f_k(t, y_k^n(t)) - f_k(t, y_k(t)) - \mathcal{D}_y f_k(t, y_k^n(t))\tilde{y}_k^n(t)\|_{L^2} \leq L\|\tilde{y}_k^n(t)\|_{L^2}.$$

Remark 6 *If we do not assume the boundedness of the derivatives of the nonlinear terms, we may use the "Stampacchia-trick" in the sense that we first extend the corresponding derivatives outside a given ball by constants and then show that for small enough data the solutions stay small and, hence, the extensions are not active. This procedure, however, would substantially extend the arguments and the length of this article as error estimates for the state and its traces would be in order. For the general concept however, see Casas and Fernández (1989). As this is very closely related to the issue of existence of optimal controls, we defer the analysis to a forthcoming publication. As of now, we therefore leave it to the reader to assess the validity of the remark.*

The following arguments follow the spirit of Lagnese and Leugering (2003), where the linear wave equation (second-order in space and time) is considered. We introduce

$$X := \left((\phi_{k,k+1})_{k=0}^{K-1}, (\phi_{k,k-1})_{k=1}^{K} \right) \in \mathcal{X} = L^2(0,1)^{2Kd},$$
$$\|X\|^2 = \sum_{k=0}^{K-1} \|\phi_{k,k+1}\|^2 + \sum_{k=1}^{K} \|\phi_{k,k-1}\|^2. \tag{40}$$

For given iteration histories $\phi_{k,k+1}$ and $\phi_{k,k-1}$, i.e., for given X, we consider the unique solution y_k^n, p_k^n of (37), (38) and then define $T : \mathcal{X} \to \mathcal{X}$ by

$$TX := \left((y_{k+1}(T_{k+1}) + \beta p_{k+1}(T_{k+1}))_{k=0}^{K-1}, (y_{k-1}(T_k) - \beta p_{k-1}(T_k))_{k=1}^{K} \right). \tag{41}$$

In order to consider the fixed-point iteration

$$X^{n+1} = (1-\varepsilon)TX^n + \varepsilon X^n, \quad \varepsilon \in [0,1), \tag{42}$$

we will show that the operator $T : \mathcal{X} \to \mathcal{X}$ is non-expansive. Note that even though the definition of T involves the time traces of the state and the adjoint linearly, the mapping is, in fact, nonlinear, according to the nonlinear term in the problem formulation. As X is a fixed point,

$$X^n - X = \left(\left(\tilde{y}_k^n(T_{k+1}) + \beta\tilde{p}_k^n(T_{k+1})\right)_{k=0}^{K-1}, \left(\tilde{y}_k^n(T_k) - \beta\tilde{p}_k^n(T_k)\right)_{k=1}^{K} \right),$$
$$TX^n - X = TX^n - TX = \left(\left(\tilde{y}_{k+1}^n(T_{k+1}) + \beta\tilde{p}_{k+1}^n(T_{k+1})\right)_{k=0}^{K-1}, \left(\tilde{y}_{k-1}^n(T_k) - \beta\tilde{p}_{k-1}^n(T_k)\right)_{k=1}^{K} \right) \tag{43}$$

holds at iteration n. Moreover, we define the energies

$$\mathcal{E}_k^n(t) := \|\tilde{y}_k^n(t)\|^2 + \beta^2\|\tilde{p}_k^n\|^2, \quad \mathcal{E}^n := \sum_{k=0}^{K-1} \mathcal{E}_k^n(T_{k+1}) + \mathcal{E}_{k+1}^n(T_{k+1}). \tag{44}$$

Now,

$$
\begin{aligned}
\|\tilde{X}^n\|^2 &= \sum_{k=0}^{K-1} \|\tilde{y}_k^n(T_{k+1}) + \beta \tilde{p}_k^n(T_{k+1})\|^2 + \sum_{k=1}^{K} \|\tilde{y}_k^n(T_k) - \beta \tilde{p}_k^n(T_k)\|^2 \\
&= \sum_{k=0}^{K-1} \left(\|\tilde{y}_k^n(T_{k+1})\|^2 + 2\beta \tilde{y}_k^n(T_{k+1}) \tilde{p}_k^n(T_{k+1}) + \beta^2 \|\tilde{p}_k^n(T_{k+1})\|^2 \right) \\
&\quad + \sum_{k=1}^{K} \left(\|\tilde{y}_k^n(T_k)\|^2 - 2\beta \tilde{y}_k^n(T_k) \tilde{p}_k^n(T_k) + \beta^2 \|\tilde{p}_k^n(T_k)\|^2 \right) \\
&= \mathcal{E}^n + 2\beta \left(\sum_{k=0}^{K-1} \tilde{y}_k^n(T_{k+1}) \tilde{p}_k^n(T_{k+1}) - \sum_{k=1}^{K} \tilde{y}_k^n(T_k) \tilde{p}_k^n(T_k) \right) \\
&=: \mathcal{E}^n + \mathcal{F}^n
\end{aligned}
\tag{45}
$$

holds. Similarly, we have

$$\|TX^n - TX\|^2 = \mathcal{E}^n - \mathcal{F}^n$$

and, hence,

$$\|TX^n - TX\|^2 = \|X^n - X\| - 2\mathcal{F}^n.$$

In fact, given any pair $X_1, X_2 \in \mathcal{X}$, we have $\|TX_1 - TX_2\| = \|X_1 - X_2\| - 2\mathcal{F}$ by the same arguments. Therefore, T is non-expansive if $\mathcal{F} \geq 0$. For the sake of brevity, it is enough to show this with the iteration errors. We proceed with relations concerning the global errors $\mathcal{E}$ and the term $\mathcal{F}$.

Proposition 1 (Lemma 2 in Lagnese and Leugering, 2003) *For any $\varepsilon \in [0,1)$ and $n \in \mathbb{N}$, we have*

1. $\mathcal{E}^{n+1} + \mathcal{F}^{n+1} \leq \mathcal{E}^n - (1 - 2\varepsilon)\mathcal{F}^n$,
2. $\mathcal{E}^{n+1} + \sum_{l=1}^{n+1} c_l(\varepsilon) F^l \leq \mathcal{E}^1$ *with* $c_1(\varepsilon) = 1 - 2\varepsilon$, $c_{n+1}(\varepsilon) = 1$, *and* $c_l(\varepsilon) = 2(1-\varepsilon)$ *for* $l = 2, \dots, n$.

Proof As the proof is on the level of relations between $\mathcal{E}^{n+1}, \mathcal{F}^{n+1}, \mathcal{E}^n, \mathcal{F}^n$ only, we refer to Lagnese and Leugering (2003). □

In order to utilize the crucial Property 2 of Proposition 1, we need to establish the positiveness of $\mathcal{F}^n$, where $\mathcal{F}^n$ is defined in (45). To this end, we now multiply (37a) by $\tilde{p}_k^n$ and the integration by parts then leads to

$$
\begin{aligned}
0 &= \int_{T_k}^{T_{k+1}} \int_0^L \left(\partial_t \tilde{y}_k^n + \Lambda \partial_x \tilde{y}_k^n + M \tilde{y}_k^n - (f_k(y_k^n) - f_k(y_k)) - B_d \tilde{u}_k^n\right) \tilde{p}_k^n \,\mathrm{d}x \,\mathrm{d}t \\
&= \int_0^L \left(\tilde{y}_k^n(T_{k+1}) \tilde{p}_k^n(T_{k+1}) - \tilde{y}_k^n(T_k) \tilde{p}_k^n(T_k)\right) \,\mathrm{d}x \\
&\quad - \int_{T_k}^{T_{k+1}} B_b^T |\Lambda^-(L)| \tilde{p}_k^n(t, L) \tilde{v}_k^n \,\mathrm{d}t - \int_{T_k}^{T_{k+1}} \int_0^L \tilde{u}_k^n B_d^T \tilde{p}_k^n \,\mathrm{d}x \,\mathrm{d}t \\
&\quad - \int_{T_k}^{T_{k+1}} \int_0^L \tilde{y}_k^n \left(\partial_t \tilde{p}_k^n + \Lambda \partial_x \tilde{p}_k^n + (\Lambda_x - M^T) \tilde{p}_k^n\right) \,\mathrm{d}x \,\mathrm{d}t \\
&\quad - \int_{T_k}^{T_{k+1}} \int_0^L (f_k(t, y_k^n) - f_k(t, y_k)) \tilde{p}_k^n \,\mathrm{d}x \,\mathrm{d}t \\
&= \int_0^L \left(\tilde{y}_k^n(T_{k+1}) \tilde{p}_k^n(T_{k+1}) - \tilde{y}_k^n(T_k) \tilde{p}_k^n(T_k)\right) \,\mathrm{d}x \\
&\quad - \int_{T_k}^{T_{k+1}} B_b^T |\Lambda^-(L)| \tilde{p}_k^n(t, L) \tilde{v}_k^n \,\mathrm{d}t - \int_{T_k}^{T_{k+1}} \int_0^L \tilde{u}_k^n B_d^T \tilde{p}_k^n \,\mathrm{d}x \,\mathrm{d}t \\
&\quad - \mu \int_{T_k}^{T_{k+1}} \int_0^L \tilde{y}_k^n \tilde{y}_k^n \,\mathrm{d}x \,\mathrm{d}t \\
&\quad + \int_{T_k}^{T_{k+1}} \int_0^L \tilde{y}_k^n \left(\left(\mathcal{D}_y f_k(t, y_k^n)^\top\right) \tilde{p}_k^n + \left(\mathcal{D}_y f_k(t, y_k^n)^\top - \mathcal{D}_y f_k(t, y_k)^\top\right) p_k\right) \\
&\quad\quad - (f_k(t, y_k^n) - f_k(t, y_k)) \tilde{p}_k^n \,\mathrm{d}x \,\mathrm{d}t.
\end{aligned} \tag{46}
$$

We obtain

$$
\begin{aligned}
&\int_0^L \left(\tilde{y}_k^n(T_{k+1}) \tilde{p}_k^n(T_{k+1}) - \tilde{y}_k^n(T_k) \tilde{p}_k^n(T_k)\right) \,\mathrm{d}x \\
&= \int_{T_k}^{T_{k+1}} B_b^T |\Lambda^-(L)| \tilde{p}_k^n(t, L) \tilde{v}_k^n \,\mathrm{d}t + \int_{T_k}^{T_{k+1}} \int_0^L \tilde{u}_k^n B_d^T \tilde{p}_k^n \,\mathrm{d}x \,\mathrm{d}t \\
&\quad + \mu \int_{T_k}^{T_{k+1}} \int_0^L \tilde{y}_k^n \tilde{y}_k^n \,\mathrm{d}x \,\mathrm{d}t \\
&\quad - \int_{T_k}^{T_{k+1}} \int_0^L \tilde{y}_k^n \left(\left(\mathcal{D}_y f_k(t, y_k^n)^\top\right) \tilde{p}_k^n + \left(\mathcal{D}_y f_k(t, y_k^n)^\top - \mathcal{D}_y f_k(t, y_k)^\top\right) p_k\right) \\
&\quad\quad - (f_k(t, y_k^n) - f_k(t, y_k)) \tilde{p}_k^n \,\mathrm{d}x \,\mathrm{d}t.
\end{aligned} \tag{47}
$$

Up to now, we have not made use of the assumption that the controls are unconstrained. If we do, we obtain from (47) that

$$
\begin{aligned}
&\int_0^L \left(\tilde{y}_k^n(T_{k+1})\tilde{p}_k^n(T_{k+1}) - \tilde{y}_k^n(T_k)\tilde{p}_k^n(T_k)\right)\,\mathrm{d}x \\
&= \frac{1}{\rho}\int_{T_k}^{T_{k+1}} |B_b^T|\Lambda^-(L)|\tilde{p}_k^n(t,L)|^2\,\mathrm{d}t + \frac{1}{\nu}\int_{T_k}^{T_{k+1}}\int_0^L |B_d^T\tilde{p}_k^n|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \mu\int_{T_k}^{T_{k+1}}\int_0^L \tilde{y}_k^n\tilde{y}_k^n\,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_{T_k}^{T_{k+1}}\int_0^L \tilde{y}_k^n\left(\left(\mathcal{D}_y f_k(t,y_k^n)^\top\right)\tilde{p}_k^n + \left(\mathcal{D}_y f_k(t,y_k^n)^\top - \mathcal{D}_y f_k(t,y_k)^\top\right)p_k\right) \\
&\qquad - (f_k(t,y_k^n) - f_k(t,y_k))\tilde{p}_k^n\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
\tag{48}
$$

holds. Obviously, if there is no nonlinearity, the left-hand side of (48) is positive. In the case of appearance of the nonlinear term f, we need to absorb the last integral in (48) into the quadratic terms. For this, we will resort to (1). We derive from (48) the following expression for $\mathcal{F}$:

$$
\begin{aligned}
\frac{1}{2\beta}\mathcal{F} &= \sum_{k=0}^K \int_0^L \left(\tilde{y}_k^n(T_{k+1})\tilde{p}_k^n(T_{k+1}) - \tilde{y}_k^n(T_k)\tilde{p}_k^n(T_k)\right)\,\mathrm{d}x - \int_{T_k}^{T_{k+1}} \tilde{y}_K^n(T_{K+1})\tilde{p}_K^n(T_{K+1})\,\mathrm{d}t \\
&= \kappa\int_0^L |\tilde{y}_k^n(T)|\,\mathrm{d}x + \mu\sum_{k=0}^K\int_{T_k}^{T_{k+1}}\int_0^L \tilde{y}_k^n\tilde{y}_k^n\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \frac{1}{\rho}\sum_{k=0}^K\int_{T_k}^{T_{k+1}} |B_b^T|\Lambda^-(L)|\tilde{p}_k^n(t,L)|^2\,\mathrm{d}t + \frac{1}{\nu}\sum_{k=0}^K\int_{T_k}^{T_{k+1}}\int_0^L |B_d^T\tilde{p}_k^n|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\quad - \sum_{k=0}^K\int_{T_k}^{T_{k+1}}\int_0^L \tilde{y}_k^n\left(\left(\mathcal{D}_y f_k(t,y_k^n)^\top\right)\tilde{p}_k^n + \left(\mathcal{D}_y f_k(t,y_k^n)^\top - \mathcal{D}_y f_k(t,y_k)^\top\right)p_k\right) \\
&\qquad - (f_k(t,y_k^n) - f_k(t,y_k))\tilde{p}_k^n\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
\tag{49}
$$

In order to compensate the last term of (49), we invoke Assumption 1. We have

$$
\begin{aligned}
&\sum_{k=0}^K\int_{T_k}^{T_{k+1}}\int_0^L \left(\left(f_k(t,y_k^n) - f_k(t,y_k) - \mathcal{D}_y f_k(t,y_k^n)\tilde{y}_k^n\right)\tilde{p}_k^n\right. \\
&\qquad\qquad \left. - \left(\mathcal{D}_y f_k(t,y_k^n) - \mathcal{D}_y f_k(t,y_k)\right)\tilde{y}_k^n p_k\ \right)\,\mathrm{d}x\,\mathrm{d}t \\
&\le \sum_{k=0}^K\int_0^T\int_0^L \left(\sigma\|\tilde{y}_k^n\|^2 + \eta\|\tilde{p}_k^n\|^2\right)\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
\tag{50}
$$

for some suitable numbers $\sigma, \eta > 0$ that depend on L and the L^∞-norm of p, i.e., of the adjoint of the underlying global optimality system, which is also equal to νu. The estimation in (49), together with (50), leads to

$$
\begin{aligned}
\mathcal{F}^l \geq\ & 2\beta \left\{ \kappa \int_0^L |\tilde{y}_k^n(T)| \,\mathrm{d}x + \sum_{k=0}^{K} \left(\int_{T_k}^{T_{k+1}} \int_0^L \mu \|\tilde{y}_k^l\|^2 + \frac{1}{\nu} \|\tilde{p}_k^l\|^2 \,\mathrm{d}x \,\mathrm{d}t \right.\right. \\
& \left.\left. + \frac{1}{\rho} \int_{T_k}^{T_{k+1}} |B_b^T| \Lambda^-(L) |\tilde{p}_k^l(t,L)^2| \,\mathrm{d}t \, - \int_{T_k}^{T_{k+1}} \int_0^L \rho \|\tilde{y}_k^l\|^2 + \eta \|\tilde{p}_k^l\|^2 \,\mathrm{d}x \,\mathrm{d}t \right) \right\} \\
\geq\ & 2\beta \left\{ \kappa \int_0^L |\tilde{y}_k^n(T)| \,\mathrm{d}x + \sum_{k=0}^{K} \left(\int_{T_k}^{T_{k+1}} \int_0^L \left((\mu - \sigma) \|\tilde{y}_k^l\|^2 + \left(\frac{1}{\nu} - \eta \right) \|\tilde{p}_k^l\|^2 \right) \mathrm{d}x \,\mathrm{d}t \right.\right. \\
& \left.\left. + \frac{1}{\rho} \int_{T_k}^{T_{k+1}} |B_b^T| \Lambda^-(L) |\tilde{p}_k^l(t,L)^2| \,\mathrm{d}t \right) \right\}.
\end{aligned}
$$

Assume now that the parameters κ, μ are chosen sufficiently large and ρ, ν sufficiently small, such that $\delta_1 := \mu - \sigma > 0$ and $\delta_2 := \frac{1}{\nu} - \eta > 0$ holds. Then, Proposition 1 (2) yields

$$
\begin{aligned}
\mathcal{E}^{n+1} + \sum_{l=1}^{n+1} c_l(\varepsilon) & \left\{ \kappa \int_0^L |\tilde{y}_k^n(T)| \,\mathrm{d}x + \sum_{k=0}^{K} \left(\int_{T_k}^{T_{k+1}} \int_0^L \delta_1 \|\tilde{y}_k^l\|^2 + (\delta_2) \|\tilde{p}_k^l\|^2 \,\mathrm{d}x \,\mathrm{d}t \right.\right. \\
& \left.\left. + \frac{1}{\rho} \int_{T_k}^{T_{k+1}} |B_b^T| \Lambda(L) |\tilde{p}_{ki}^l(t,1)|^2 \,\mathrm{d}t \right) \right\} \leq \mathcal{E}^1, \quad n = 1, 2, \ldots
\end{aligned}
\tag{51}
$$

and, in turn, (51) provides

$$
\begin{aligned}
& \mathcal{E}^n \text{ is bounded}, \\
& \tilde{y}_k^l \to 0, \ \tilde{p}_k^l \to 0 \text{ in } L^2\left(I_k; L^2(0,1)\right), \\
& \tilde{p}_{ki}^l(t,L) \to 0, \ i = m+1, \ldots, , d \ \tilde{y}_k^l(T,\cdot), \quad \tilde{p}_K(T,\cdot) \to 0 \text{ in } L^2(0,L)^d
\end{aligned}
\tag{52}
$$

as $l \to \infty$ and for $\varepsilon \in [0,1)$. From (52), it is now clear that $\tilde{y}_k^n(T_{k+1})$, $\tilde{y}_k^n(T_k)$, $\tilde{p}_k^n(T_{k+1})$, and $\tilde{p}_k^n(T_k)$ converge to zero. In other words, the transmission conditions are satisfied in the limit. We know, however, that $\mathcal{E}^n$ is bounded and, hence, that there is a weakly convergent subsequence. Notice, though, that on a subsequence it is not necessarily clear that with index k also $k+1$ is a part of that subsequence and, thus, the iteration cannot be used.

On the other hand, we see from (52) that $\tilde{u}_k^n \to 0$ and $\tilde{v}_k^n \to 0$ in $L^2(0,T;L^2(0,1))^d$ or $L^2(0,T)^m$, respectively. Due to the continuity of the nonlinear functions f_k and g^1, we obtain vanishing right-hand sides in the state and adjoint equations and homogeneous boundary conditions in the limit on the entire sequence. Due to the uniqueness of the solution of the optimality system, the initial and final data $\tilde{y}_k^n(T_k)$, $\tilde{y}_k^n(T_{k+1})$, $\tilde{p}_k^n(T_k)$, and $\tilde{p}_k^n(T_{k+1})$ converge to zero in $L^2(0,1)^d$. As the functions $y_k(\cdot)$ and $p_k(\cdot)$ satisfy Conditions (13), in the limit, the transmission conditions hold. This is true even for $\varepsilon = 0$.

Theorem 4 *The solutions* (y_k^n, p_k^n) *of* (16), (7), (8), (14), (15) *with* $\varepsilon \in [0,1)$ *converge to* (2), (5), (39) *strongly in* $L^2\left(0,T;L^2(0,1)\right)$ *to* (y_k, p_k)*, which is the solution of* (17)*, for* $k = 0, \ldots, K$.

Remark 7 *For $\varepsilon \in (0,1)$, we can derive the convergence of the initial and final data for y_k^n, p_k^n at T_k, T_{k+1} directly. To this end, we consider the term*

$$\begin{aligned}\langle TX^n, X^n\rangle = &\sum_{k=0}^{K-1}\left(\tilde{y}_k^n(T_{k+1}) + \beta\tilde{p}_k^n(T_{k+1})\right)\left(\tilde{y}_{k+1}^n(T_{k+1}) + \beta\tilde{p}_{k+1}(T_{k+1})\right)\\ &+\sum_{k=1}^{K}\left(\tilde{y}_k^n(T_k) - \beta\tilde{p}_k^n(T_k)\right)\left(\tilde{y}_{k-1}^n(T_k) - \beta\tilde{p}_{k-1}^n(T_k)\right)\\ =& 2\sum_{k=0}^{K-1}\left(\tilde{y}_k^n(T_{k+1})\tilde{y}_{k+1}^n(T_{k+1}) + \beta^2\tilde{p}_k^n(T_{k+1})\tilde{p}_{k+1}^n(T_{k+1})\right).\end{aligned}$$

Thus,

$$\begin{aligned}&\frac{1}{2}\|TX^n - X^n\|^2 = \mathcal{E}^n - \langle TX^n, X^n\rangle\\ &=\sum_{k=0}^{K-1}\left(\|\tilde{y}_k^n(T_{k+1}) - \tilde{y}_{k+1}^n(T_{k+1})\|^2 + \beta^2\|\tilde{p}_k^n(T_{k+1}) - \tilde{p}_{k+1}^n(T_{k+1})\|^2\right).\end{aligned}$$

On the other hand

$$\|T_\varepsilon^n X^1 - T_\varepsilon^{n-1}X^1\| = (1-\varepsilon)\|TX^n - X^n\| \to 0$$

holds according to Schaefer's fixed-point theorem (Schaefer, 1957) for $\varepsilon \in (0,1)$. This directly shows the desired convergence.

Remark 8 *We see from the proof of Theorem 4 that with the given nonlinearity f_k, we need the distributed control u_k to compensate for the appearance of the nonlinearity in the estimates. We also need the tracking term with $\kappa > 0$ being distributed over space and time. On the other hand, in this setting, we obtain stronger convergence results than in Lagnese and Leugering (2003, 2004).*

Remark 9 *Distributed control with full access to the state are typically hard to implement in practice. Here, observers can be useful as they can provide estimates of the states also based upon pointwise-in-space measurements. See, e.g., Gugat, Giesselmann and Kunkel (2021) for a recent work on nodal observers for networked semilinear systems.*

For boundary controls, on the other hand, full access is a not critical issue. As our convergence proofs reveal, full access to the state is, however, essential to compensate for the distributed nonlinear term. In this respect, we add that one may replace the distributed control by yet another virtual control, however, at the expense of introducing an approximation to the adjoint variable appearing in the optimality system on the decomposed level. We do not have the space to elaborate on that variant here in detail but rather sketch the idea and refer to a further publication. Indeed, assume we have an approximation $\hat{p}_{ki}^n$ of the true global adjoint variable $p_k := p|_{Q_k}$ such that $\hat{p}_{ki}^n \to p_{ki}$.

Then, we can write

$$\partial_t y_k^n + \Lambda \partial_x y_k^n + M y_k^n = f_k\,(t, y_k) + \frac{1}{\sigma} p_k^n - \frac{1}{\sigma} \hat{p}_k^n, \qquad (t,x) \in Q_k,$$
$$\partial_t p_k^n + \Lambda \partial_x p_k^n = \kappa (y_k^n - y_d) - (\mathcal{D}_y f_k^\top + \partial_x A - M^T) p_k^n, \qquad (t,x) \in Q_k,$$

on the level of Q_k and n with the same initial and boundary conditions as in (16)*. We can interpret the appearance of $\frac{1}{\sigma} p_k^n$ as the result of an unconstrained distributed control, where $\frac{1}{\sigma} \hat{p}_k^n$ is seen as an external input. Obviously, if there is convergence, then, in the limit, p_k and its approximation do no longer appear in the state equation—therefore, the notion of a virtual control. Then, if we consider the errors, we have*

$$\partial_t \tilde{y}_k^n + \Lambda \partial_x \tilde{y}_k^n = f_k(y_k^n) - f_k(y_k) + \frac{1}{\sigma} \tilde{p}_k^n + \frac{1}{\sigma}(p_k - \hat{p}_k^n), \ (t,x) \in Q_k,$$
$$\partial_t \tilde{p}_k^n + \partial_x A \tilde{p}_k^n - M^T \tilde{p}_k^n = \kappa \tilde{y}_k^n - (\mathcal{D}_y f(y_k^n)^\top) \tilde{p}_k^n - (\mathcal{D}_y f_k(y_k^n)^\top - \mathcal{D}_y f_k(y_k)^\top) p_k,$$
$$(t,x) \in Q_k.$$

The additional virtual load, which is indeed the error between the original adjoint variable and its numerical approximation $\frac{1}{\sigma}(p_k - \hat{p}_k^n)$ then accumulates in the crucial energy inequality (51)*. The assumed boundedness of the accumulated error*

$$\sum_{\ell=1}^{n+1} \sum_{k=0}^{K} \sum_{i=1}^{d} \frac{1}{\sigma} \int_{T_k}^{T_{k+1}} \int_0^L (\hat{p}_{ki}^\ell - p_{ki})^2 \,\mathrm{d}x\,\mathrm{d}t$$

then provides the same result as in Theorem 4. It is then the question of how to design the approximate solutions in terms of the actual interface values to obtain a coarse-grain-small-grain error analysis. This procedure results in a different concept of non-exact time-domain decomposition, which is beyond the scope of this article. We remark that the approach by Benamou (1996) also uses the idea of adding $\frac{1}{\sigma} p_k^n - \frac{1}{\sigma} p_k^{n-1}$. However, this will not be sufficient to compensate for the nonlinear terms.

6. Convergence in the presence of control constraints and linear dynamics

In this section, we consider pointwise constraints on u and v, i.e.,

$$u(t) \in U_{\mathrm{ad}}^{\mathrm{d}}, \quad v(t) \in U_{\mathrm{ad}}^{\mathrm{b}} \quad \text{a.e. in } (0,T). \tag{53}$$

However, we do not take into account nonlinearities. It turns out that the interaction of the control bounds with the bounds on the nonlinearities is rather complicated and not fully explored up to now. Nevertheless, we provide the convergence proof as also this extension is new in the context of optimal control for linear hyperbolic systems. We stay in the context of problem (4), together with (53). We notice that this has already been considered in Krug et al. (2021) for simpler boundary conditions. We recall the optimality conditions (7) and (8). Moreover, we also have

$$\int_{T_k}^{T_{k+1}} \int_0^L (\nu u_k^n - p_k^n)(\hat{u}_k - u_k^n) \,\mathrm{d}x\,\mathrm{d}t \geq 0 \quad \text{for all } \hat{u}_k \in U_{\mathrm{ad}}^{\mathrm{d}}.$$

For the corresponding errors we have (with $u_k \in U_{\mathrm{ad}}^{\mathrm{d}}$)

$$
\begin{aligned}
0 &\le \int_{T_k}^{T_{k+1}} \int_0^L (\nu\tilde{u}_k^n - \tilde{p}_k^n + \nu u_k - p_k)(u_k - u_k^n)\,\mathrm{d}x\,\mathrm{d}t \\
&= \int_{T_k}^{T_{k+1}} \int_0^L (\nu\tilde{u}_k^n - \tilde{p}_k^n)(-\tilde{u}_k^n)\,\mathrm{d}x\,\mathrm{d}t - \int_{T_k}^{T_{k+1}} \int_0^L (\nu u_k - p_k)(u_k^n - u_k)\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
$$

and, therefore,

$$
\int_{T_k}^{T_{k+1}} \int_0^L (\nu\tilde{u}_k^n - \tilde{p}_k^n)(-\tilde{u}_k^n)\,\mathrm{d}x\,\mathrm{d}t \ge \int_{T_k}^{T_{k+1}} \int_0^L (\nu u_k - p_k)(u_k^n - u_k)\,\mathrm{d}x\,\mathrm{d}t \ge 0.
$$

Hence,

$$
\int_{T_k}^{T_{k+1}} \int_0^L \tilde{u}_k^n \tilde{p}_k^n\,\mathrm{d}x\,\mathrm{d}t \ge \nu \int_{T_k}^{T_{k+1}} \int_0^L \tilde{u}_k^n \tilde{u}_k^n\,\mathrm{d}x\,\mathrm{d}t \tag{54}
$$

holds. By the same argument, we obtain

$$
\int_{T_k}^{T_{k+1}} B_b^T |\Lambda^-(L)| \tilde{p}_k^n(t,1) \tilde{v}_{ki}^n(t)\,\mathrm{d}t \ge \mu \int_{T_k}^{T_{k+1}} |\tilde{v}_k^n(t)|^2\,\mathrm{d}t, \tag{55}
$$

where (u_k^n, v_k^n, p_k^n) and $(\tilde{u}_k^n, \tilde{v}_k^n, \tilde{p}_k^n)$ solve (16) and (37), respectively. We recall (46).

Then, according to (45) and (47), we obtain

$$
\begin{aligned}
\mathcal{F}^n = 2\beta \Bigg\{ \kappa \int_0^L |\tilde{y}_K^n|^2(T)\,\mathrm{d}x + \sum_{k=0}^K \Bigg(&\int_{T_k}^{T_{k+1}} \tilde{v}_{ki}^n(t) B_b^T |\Lambda^-(L)| \tilde{p}_k^n(t,L)\,\mathrm{d}t \\
&+ \int_{T_k}^{T_{k+1}} \int_0^L \tilde{u}_k^n \tilde{p}_k^n\,\mathrm{d}x\,\mathrm{d}t + \mu \int_{T_k}^{T_{k+1}} \int_0^L \|\tilde{y}_k^n\|^2\,\mathrm{d}x\,\mathrm{d}t \Bigg) \Bigg\}.
\end{aligned}
$$

Next, using (54) and (55), we can estimate $\mathcal{F}$ from below by

$$
\begin{aligned}
\mathcal{F}^n \ge 2\beta \Bigg\{ \kappa \int_0^L |\tilde{y}_K^n|(T)^2\,\mathrm{d}x \sum_{k=0}^K \Bigg(&\int_{T_k}^{T_{k+1}} \int_0^L \mu \|\tilde{y}_k^n\|^2\,\mathrm{d}x\,\mathrm{d}t \\
&+ \int_{T_k}^{T_{k+1}} \int_0^L \nu \|\tilde{u}_k^n\|^2\,\mathrm{d}x\,\mathrm{d}t + \int_{T_k}^{T_{k+1}} \rho |\tilde{v}_k^n|^2\,\mathrm{d}t \Bigg) \Bigg\}.
\end{aligned}
$$

The definitions in (40)–(44) stay unchanged also in the case under consideration. Thus, we arrive at the conclusion that

$\mathcal{E}^n$ is bounded,

$$
\begin{aligned}
\tilde{y}_k^n &\to 0 \text{ in } L^2\left(I_k, L^2(0,L)^d\right), \\
\tilde{p}_k^n &\to 0 \text{ in } L^2\left(I_k, L^2(0,L)^d\right), \\
\tilde{v}_{ki}^n &\to 0 \text{ in } L(I_k)^{d-m}, \\
\tilde{y}_K^n(T) &\to 0 \text{ in } L^2(0,L)^d, \\
\tilde{p}_K^n(T) &\to 0 \text{ in } L^2(0,L)^d,
\end{aligned}
\tag{56}
$$

as $n \to \infty$.

Theorem 5 *Suppose that the controls u and v satisfy the pointwise constraints $u(t) \in U_{\text{ad}}^{\text{d}}$, $v(t) \in U_{\text{ad}}^{\text{b}}$, where $U_{\text{ad}}^{\text{d}} \subset L^2(0,1)^d$ as well as $U_{\text{ad}}^{\text{b}} \subset \mathbb{R}^m$ are convex and closed. Further, let the iterates be defined as solutions (y_k^n, p_k^n) of* (16)*,* (14)*,* (15) *with* (54)*,* (55)*. Then, these iterates converge in the sense of* (6) *to the corresponding solutions of* (2)*,* (5)–(8)*.*

7. Conclusions

We have considered mixed two-point initial-boundary value problems for semilinear hyperbolic systems with distributed and boundary controls. We provided a detailed well-posedness analysis for both strong and weak solutions. This enabled us to rigorously investigate a generic tracking-type optimal control problem and we derived optimality conditions for constrained and unconstrained controls. We then provided an iterative time-domain decomposition in the spirit of Krug et al. (2021) and proved convergence for unconstrained controls and, in the linear case, for constrained controls. The numerical analysis of the iterative time-domain decomposition method derived here is beyond the scope of this article. This and a posteriori error estimates will be subject to a forthcoming publication. In particular, applications in the context of networks with cycles are envisioned.

From the definition of the optimal control problem (4) it can be expected that the optimal state shows a turnpike structure, see, e.g., Gugat (2021). This would allow to choose longer—and therefore less—time-intervals in the decomposition of $[0, T]$ in a neighborhood of $T/2$. This, however, exceeds the scope of this work and could possibly motivate further research in the future.

Acknowledgments

The authors thank the Deutsche Forschungsgemeinschaft for their support within projects A05, B07, and B08 of the Sonderforschungsbereich / Transregio 154 "Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks".

References

Barker, A. T. and Stoll, M. (2015) Domain decomposition in time for PDE-constrained optimization. In: *Computer Physics Communications* 197, 136–143. doi: 10.1016/ j.cpc.2015.08.025.

Bastin, G. and Coron, J.-M. (2016) *Stability and Boundary Stabilization of 1-D Hyperbolic Systems*, **88**. *Progress in Nonlinear Differential Equations and their Applications. Subseries in Control*. Birkhäuser/Springer, doi: 10.1007/978-3-319-32062-5.

Benamou, J.-D. (1996) A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations. *SIAM Journal on Numerical Analysis* **33**(6), 2401–2416. doi: 10. 1137 / S0036142994267102.

Casas, E. and Fernández, L. A. (1989) A GreenâĂŹs formula for quasilinear elliptic operators. *Journal of Mathematical Analysis and Applications* **142**(1), 62–73. doi: 10.1016/0022- 247X(89)90164-9.

Gander, M. J. and Kwok, F. (2016) Schwarz methods for the time-parallel solution of parabolic control problems. *Domain Decomposition Methods in Science and Engineering* **XXII**. *Lecture Notes in Computational Science and Engineering*, 104. Springer, Cham, 207–216. doi: 10.1007/978-3-319-18827-0_19.

Gander, M. J., Kwok, F. and Salomon, J. (2020) ParaOpt: a parareal algorithm for optimality systems. *SIAM Journal on Scientific Computing* **42**(5), A2773–A2802. doi: 10.1137/ 19M1292291.

Gander, M. J. and Vandewalle, S. (2007) Analysis of the parareal time-parallel time-integration method. *SIAM Journal on Scientific Computing* **29**(2), 556–578. doi: 10.1137/ 05064607X.

Glowinski, R. and Le Tallec, P. (1990) Augmented Lagrangian interpretation of the nonoverlapping Schwarz alternating method. In: *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (Houston, TX, 1989)*. SIAM, Philadelphia, PA, 224–231.

Gugat, M. (2021) On the turnpike property with interior decay for optimal control problems. *Mathematics of Control, Signals, and Systems* **33**(2), 237–258. doi: 10.1007/ s00498-021-00280-4.

Gugat, M., Giesselmann, J. and Kunkel, T. (2021) Exponential synchronization of a nodal observer for a semilinear model for the flow in gas networks. *IMA Journal of Mathematical Control and Information* (Sept. 2021). doi: 10.1093/imamci/dnab 029.

Hante, F. M., Leugering, G., Martin, A., Schewe, L. and Schmidt, M. (2017) Challenges in Optimal Control Problems for Gas and Fluid Flow in Networks of Pipes and Canals: From Modeling to Industrial Applications. In: P. Manchanda, R. Lozi and A. H. Siddiqi, eds., *Industrial Mathematics and Complex Systems: Emerging Mathematical Models, Methods and Algorithms. Industrial and Applied Mathematics*, 77–122. doi: 10.1007/978-981-10-3758-0_5.

Heinkenschloss, M. (2005) A time-domain decomposition iterative method for the solution of distributed linear quadratic optimal control problems. *Journal of Com-*

putational and Applied Mathematics **173**(1), 169–198. doi: 10.1016/j.cam.2004.03.005.

Krug, R., Leugering, G., Martin, A., Schmidt, M. and Weninger, D. (2021) Time-Domain Decomposition for Optimal Control Problems Governed by Semilinear Hyperbolic Systems. *SIAM Journal on Control and Optimization*. url: http://www. optimization-online. org/DB_HTML/2020/11/8132.html. Forthcoming.

Lagnese, J. E. and Leugering, G. (2002) Time domain decomposition in final value optimal control of the Maxwell system. *ESAIM: Control, Optimisation and Calculus of Variations* 8, 775–799. doi: 10.1051/cocv:2002042.

Lagnese, J. E. and Leugering, G. (2003) Time-domain decomposition of optimal control problems for the wave equation. *Systems & Control Letters* **48**(3-4), 229–242. doi: 10.1016/ S0167-6911(02)00268-2.

Lagnese, J. E. and Leugering, G. (2004) Domain decomposition methods in optimal control of partial differential equations. *International Series of Numerical Mathematics* **148**. Birkhäuser Verlag, Basel. doi: 10.1007/978-3-0348-7885-2.

Leugering, G., Martin, A., Schmidt, M. and Sirvent, M. (2017) Nonoverlapping domain decomposition for optimal control problems governed by semilinear models for gas flow in networks. *Control and Cybernetics* **46**(3), 191–225.

Leugering, G. and Rodriguez, C. (2020) *Boundary feedback stabilization for the intrinsic geometrically exact beam model*. Technical report. url: https://arxiv.org/abs/1912.02543.

Leugering, G. and Schmidt, J. G. (2002) On the modelling and stabilization of flows in networks of open canals. *SIAM Journal on Control and Optimization* **41**(1), 164–180. doi: 10.1137/S0363012900375664.

Lions, J.-L. (1990) On the Schwarz alternating method. III. A variant for nonoverlapping subdomains. In: *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (Houston, TX, 1989)*. SIAM, Philadelphia, PA, 202–223.

Lions, J.-L. (1971) *Optimal Control of Systems Governed by Partial Differential Equations. Die Grundlehren der mathematischen Wissenschaften*, **Band 170**. Translated from French by S. K. Mitter. Springer-Verlag, New York-Berlin.

Lions, J.-L, Maday, Y. and Turinici, G. (2001) Résolution d'EDP par un schéma en temps "pararéel". *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics* **332**(7), 661–668. doi: 10.1016/S0764-4442(00)01793-6.

Maday, Y., Salomon, J. and Turinici, G. (2006) Monotonic time-discretized schemes in quantum control. *Numerische Mathematik* **103**(2), 323–338. doi: 10.1007/s00211- 006-0678-x.

Maday, Y. and Turinici, G. (2002) A parareal in time procedure for the control of partial differential equations. *Comptes Rendus Mathématique. Académie des Sciences*. Paris **335**(4), 387–392. doi: 10.1016/S1631-073X(02)02467-6.

Prodi, G., ed. (2010) *Problems in Non-Linear Analysis. Centro Internazionale Matematico Estivo (C.I.M.E.)* **55** *Summer Schools*. Springer, Heidelberg; Fondazione C.I.M.E., Florence.

ROUBICEK, T. (2005) *Nonlinear Partial Differential Equations with Applications. International Series of Numerical Mathematics*, **153**. Birkhäuser Verlag, Basel. doi: 10. 1007/978-3-0348-0513-1.

SCHAEFER, H. (1957) Über die Methode sukzessiver Approximationen. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 59, 131–140.

ULBRICH, S. (2007) Generalized SQP methods with "parareal" time-domain decomposition for time-dependent PDE-constrained optimization. In: *Real-time PDE-constrained Optimization*. *Computational Science & Engineering* 3, SIAM, 145–168. doi: 10 . 1137 / 1 . 9780898718935.ch7.

VRABIE, I. I. (2003) *C_0-Semigroups and Applications. North-Holland Mathematics Studies* **191**. North-Holland Publishing Co., Amsterdam.

WU, S.-L. AND HUANG, T.-Z. (2018) A fast second-order parareal solver for fractional optimal control problems. *Journal of Vibration and Control* **24**(15), 3418–3433. doi: 10.1177/ 1077546317705557.

WU, S.-L. AND LIU, J. (2020) A parallel-in-time block-circulant preconditioner for optimal control of wave equations. *SIAM Journal on Scientific Computing* **42**(3), A1510–A1540. doi: 10.1137/19M1289613.