Convergence of discontinuous time-stepping schemes for a Robin boundary control problem under minimal regularity assumptions

Konstantinos Chrysafinos*

March 21, 2012

Abstract

The minimization of the energy functional having states constrained to semi-linear parabolic PDE's is considered. The controls act on the boundary and are of Robin type. The discrete schemes under consideration are discontinuous in time but conforming in space. Stability estimates are presented at the energy norm and at arbitrary times for the state, and adjoint variables. The estimates are derived under minimal regularity assumptions and are applicable for higher order elements. Using these estimates and an appropriate compactness argument (see Walkington [49, Theorem 3.1]) for discontinuous Galerkin schemes, convergence of the discrete solution to the continuous solution is established. In addition, a discrete optimality system is derived and convergence of the corresponding discrete solutions is also demonstrated.

1 Introduction

Space-time approximations of an optimal Robin boundary control problem are examined by using discontinuous time-stepping Galerkin schemes. In particular, the optimal control problem considered here is associated to the minimization of the energy functional,

$$J(y,g) = \frac{1}{2}\int_0^T \|\nabla y\|^2_{L^2(\Omega)}dt + \frac{\alpha}{2}\int_0^T \|g\|^2_{L^2(\Gamma)}dt \tag{1.1}$$

subject to the constraints,

$$\left\{ \begin{aligned} y_t - \eta\Delta y + \phi(y) &= f && \text{in } (0,T)\times\Omega \\ y + \lambda^{-1}\eta\frac{\partial y}{\partial \mathbf{n}} &= g && \text{on } (0,T)\times\Gamma \\ y(0,x) &= y_0 && \text{in } \Omega. \end{aligned} \right. \tag{1.2}$$

[†]AMS subject classification:65M60,49J20.

Acknowledgement: This work is partially supported by a PEBE grant.

Keywords and phrases: Discontinuous Time-Stepping Schemes, Finite Element Approximations, Robin Boundary Control, Semi-linear Parabolic PDE's.

*Address: National Technical University of Athens, School of Applied Mathematical and Physical Sciences, Department of Mathematics, Zografou Campus, Athens 15780, Greece, chrysafinos@math.ntua.gr

Here, Ω denotes a bounded domain in $\mathbb{R}^2$, with Lipschitz boundary Γ, λ, η positive constants, y_0, f denote the initial data and the forcing term respectively, satisfying minimal regularity assumptions, i.e.,

$$y_0 \in L^2(\Omega), \qquad f \in L^2[0,T; H^1(\Omega)^*].$$

The boundary control g is of Robin type, and α is a penalty parameter. The nonlinear mapping ϕ is monotone and continuous satisfying certain growth conditions. Several results regarding the analysis of optimal boundary control problems can be found in [24, 34, 41, 48] (see also references within). In this work we are interested in analyzing discontinuous time-stepping schemes of arbitrary order, for the gradient minimization problem with boundary controls of Robin type. There are several structural difficulties involved in the analysis of numerical schemes of such boundary control problems.

- The minimization of the energy functional combined with the rough initial data and forcing term, severely restricts the regularity of the state y and the control g. Therefore, standard numerical techniques employed for uncontrolled parabolic equations typically fail, since they demand more regularity of y_t than anticipated.
- In addition, the associated first order necessary conditions consist of a (backward in time) adjoint equation which is coupled to the primal (forward in time) equation through an optimality condition on the boundary, and nonlinear terms (see, e.g [24, 34, 41, 48]). This leads to reduced regularity for the state and adjoint variables, and hence numerical analysis approaches based on standard 'boot-strap' techniques are not directly applicable, in particular in presence of $L^2(\Omega)$ initial data.
- If higher norms of the control g are included in the functional then a boot-strap argument can be applied to recover some additional regularity on the adjoint variable, taking into account the linear structure of the adjoint PDE, and the zero terminal data. However this approach typically leads to an optimality condition of a PDE form on the boundary which is hard to solve computationally.
- Due to the lack of regularity, the recovery of the necessary compactness of discrete schemes, via discrete version of the classical Aubin-Lions Lemma for nonlinear PDE's (see e.g. [44, 51]) is not evident.
- The parameter α effectively determines the "size" of the control g, which is needed in order to minimize the gradient. As a consequence the dependence of various stability constants upon α should be tracked. For some relevant discussions for the velocity tracking problem we refer the reader to [24].

To overcome these difficulties we analyze a classical discontinuous Galerkin scheme which is discontinuous in time and conforming in space. It is well known that discontinuous time-stepping schemes perform well for problems which satisfy low regularity properties. As we will subsequently show for the discrete control problem, the discontinuous time-stepping schemes inherit crucial regularity and stability properties of the continuous weak formulation of the underlying PDE, such as estimates under minimal regularity assumptions at arbitrary time points (see e.g. [10, 11, 49] for the uncontrolled evolutionary PDE's). Such estimates allow the use of the recently developed discrete compactness property of discontinuous time-stepping schemes under minimal regularity assumptions (see Walkington, [49, Theorem 3.1]), within the optimal control setting.

As a consequence, strong convergence in an appropriate norm is established, and hence the semi-linear terms are treated by embedding theorems. Using the above technique, we prove convergence of the discrete optimal solution to the continuous problem. In addition, a "boot-strap" argument can be rigorously applied in order to derive the associated discrete optimality system (discrete first order necessary condition) and then to prove convergence of the corresponding discrete adjoint variable, without requiring additional regularity on the time-derivatives. A novel element of the proposed methodology for the boundary control problem with semi-linear state constraints is that the time discretization step length τ can be chosen independently of the spatial discretization parameter h. In addition, the dependence upon λ, α of various stability constants is carefully tracked. The emphasis here is to avoid any exponential dependence of quantities $1/\alpha$ in various stability constants appearing in our estimates.

We note also that in the work of Meidner and Vexler in [38, 39, 40] discontinuous Galerkin schemes were analyzed for distributed optimal control problems constrained to linear parabolic PDEs and their computational effectiveness were demonstrated. A posteriori estimates for DG schemes related to linear parabolic control problems were studied by Liu and Yan in [35] and by Liu, Ma, Tang and Yan in [36]. Some a-priori estimates for linear parabolic distributed control problems with time-dependent coefficients were estiblished in [7]. In the recent work of Neitzel and Vexler in [42] first order (in time) error estimates for the controls are presented for an optimal control problem related to semi-linear parabolic PDEs. The controls are of distributed type and satisfy control constraints while the initial data belong to $H_0^1(\Omega) \cap L^\infty(\Omega)$ under weak hypothesis on semi-linear term. The controls are discretized by piecewise constants in time and space, while for the state equation, the lowest order ($k = 0$) discontinuous Galerkin (in time) combined with standard conforming finite elements (in space), are being used. Finally, in [9], symmetric error estimates for a general class of discontinuous time-stepping schemes are presented for distributed optimal control problems having states constrained to semi-linear parabolic PDEs, without control constraints.

Our technique extends the results of [8] which were developed for the minimization of tracking functional using distributed controls to the energy minimization problem with boundary controls. The approach undertaken in this work allows the rigorous derivation of the discrete first order necessary conditions, under the presence of minimal regularity assumptions. To our best knowledge the regularity / stability and convergence results within the discrete boundary control setting are also new.

1.1 Related results

For boundary controls related to nonlinear elliptic PDE's we refer the reader to the works (and references) of [3, 4, 5, 6, 14, 24, 26, 30, 28, 34, 41, 48].

Boundary controls for time-dependent problems were studied in [12, 23, 22, 27, 29, 31, 33, 37, 43, 46, 47, 48, 50]. In [50], the terminal state tracking functional is minimized using Neumann controls, while in [37] fully-discrete approximations of a Neumann boundary control problem related to homogeneous linear parabolic PDE's are presented. A Robin boundary control (of separation type) is used in [31], in order to determine the minimal time for the controlled state to reach a desired target. The state equation is a linear parabolic PDE, and the convergence of semi-discrete finite element approximations is presented. Convergence rates for a time optimal boundary control problem for a homogeneous linear parabolic PDE were given in [33] based on a semigroup approach (see also [34]). In [46, 47], nonlinear boundary controls are used to minimize a functional which can handle terminal norms and matching controls. The size of the control is limited and C^∞ smoothness on the boundary is needed. In [12], error estimates for the semi-discrete (in space) approximations of a Robin boundary control problem constrained to semilinear parabolic PDE's are presented. The state tracking functional is minimized and estimates of arbitrary order are derived under minimal regularity assumptions. The velocity tracking problem for Navier-Stokes equations using boundary controls was examined in [27] using first order (in time) discrete scheme. An optimality system (first order necessary conditions) of equations was rigorously derived and the convergence of a gradient algorithm was proven. For second order necessary condition methods for boundary control problems related to the evolutionary Navier-Stokes equations we refer the reader to [29], while second order sufficient optimality conditions where studied in [43]. In [23], a primal-dual active set strategy for the numerical solution of Neumann boundary control problems constrained to systems of semi-linear parabolic PDE's is analyzed, while in [32] a Dirichlet boundary control problem with control belonging to $L^2(\Gamma)$ is analyzed for linear evolutionary parabolic PDE's.

For results related to the discontinuous Galerkin method for the solution of parabolic PDE's (without applying controls) and its relation to adaptivity is quite extensive (see e.g. [45, 16] and references therein). Results related to semi-linear parabolic problems are presented in [15, 17, 18].

1.2 Outline

An outline of this paper follows. After introducing the necessary notation in section 2, the continuous optimal control problem is defined. In the remaining two sections we focus on the DG approximation of the Robin boundary control problem. In section 3, stability estimates for the solutions of the discretized optimal control problem are obtained at arbitrary time points. In Section 4, a discrete optimality system of equations is derived, and stability estimates on the adjoint variable at arbitrary times are presented. Then, convergence of the discrete solution of the optimality system to the solution of the continuous optimality system is shown under minimal regularity assumptions.

2 Preliminaries

2.1 Notation

We employ the standard notation for Hilbert spaces $L^2(\Omega)$, $H^s(\Omega)$, $L^2(\Gamma)$, related norms and inner products (see e.g. [19, Chapter 5]). We denote by X^* the dual of X for any Banach space X. The duality pairings of $H^1(\Omega), H^1(\Omega)^*$ and $H^{1/2}(\Gamma), H^{1/2}(\Gamma)^*$ are denoted by $\langle .,. \rangle$ and $\langle .,. \rangle_\Gamma$ respectively. Similarly we denote by $L^p[0,T;X], L^\infty[0,T;X]$ and $C[0,T;X]$

the time-space spaces, endowed with standard norms (see e.g. [19, 51]). We will use the following space for the solution (natural energy) space

$$W(0,T) = L^2[0,T;H^1(\Omega)] \cap L^\infty[0,T;L^2(\Omega)]$$

with norm

$$\|v\|^2_{W(0,T)} = \|v\|^2_{L^2[0,T;H^1(\Omega)]} + \|v\|^2_{L^\infty[0,T;L^2(\Omega)]}.$$

The bilinear form associated to our operator, is defined by

$$a(y,v) = \eta \int_\Omega \nabla y \nabla v dx \qquad \forall\, y, v \in H^1(\Omega),$$

which satisfies the standard coercivity and continuity conditions

$$a(y,y) \geq \eta \|\nabla y\|^2_{L^2(\Omega)}, \qquad a(y,v) \leq C\eta \|y\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \qquad \forall\, y, v \in H^1(\Omega).$$

For the semi-linear term we impose the following structural assumptions.

Assumption 2.1. *The semi-linear term $\phi \in \mathcal{C}^1(\mathbb{R};\mathbb{R})$, and satisfies the following monotonicity and growth properties: $p \in (1,3)$, and for all $s \in \mathbb{R}$*

$$\phi'(s) \geq 0, \qquad s\phi(s) \geq C|s|^{p+1}, \qquad \big|\phi(s)\big| \leq C|s|^p, \qquad \big|\phi'(s)\big| \leq C|s|^{p-1}.$$

Remark 2.2. *The assumption on the semi-linear term can be relaxed in some cases for the continuous optimal control problem (see e.g. [48] and references within). Here, we have imposed assumptions on ϕ that guarantee that the resulting discrete semi-linear parabolic problem, possesses $L^\infty[0,T;L^2(\Omega)]$ regularity under minimal regularity assumptions on the initial data and the forcing term.*

A suitable weak formulation of (1.2) is formulated as follows. Let $f \in L^2[0,T;H^1(\Omega)^*]$, $g \in L^2[0,T;H^{-1/2}(\Gamma)^*]$ and $y_0 \in L^2(\Omega)$. Then, we seek $y \in W(0,T)$ such that

$$\begin{aligned}(y(T),v(T)) \quad &+ \quad \int_0^T \left(-\langle y, v_t\rangle + a(y,v) + \langle \phi(y), v\rangle + \lambda \langle y, v\rangle_\Gamma\right) dt \qquad (2.1)\\ &= \quad (y_0, v(0)) + \int_0^T \left(\langle f, v\rangle + \lambda\langle g, v\rangle_\Gamma\right) dt,\end{aligned}$$

for all $v \in L^2[0,T;H^1(\Omega)] \cap H^1[0,T;H^1(\Omega)^*]$. Note that the data satisfy the minimal regularity assumptions to guarantee the existence of a unique solution in $W(0,T)$, while the Robin boundary control will be sought in the space,

$$g \in L^2[0,T;L^2(\Gamma)],$$

due to the structure of the functional. Below, we recall several useful inequalities, and the Gagliardo-Nirenberg interpolation inequality (see e.g. [2, 19]) for two dimensional domains, which will be used subsequently.

Hölder Inequality: For any measurable set E, of any dimension and for $(1/s_1)+(1/s_2)+(1/s_3) = 1$, $s_i \geq 1$,

$$\int_E f_1 f_2 f_3 dE \leq \|f_1\|_{L^{s_1}(E)} \|f_2\|_{L^{s_2}(E)} \|f_3\|_{L^{s_3}(E)}.$$

Young Inequality: For any $a, b \geq 0$, $\delta > 0$, and $s_1, s_2 > 1$

$$ab \leq \delta a^{s_1} + C(\delta) b^{s_2}, \qquad \text{with } (1/s_1) + (1/s_2) = 1.$$

Gagliardo-Nirenberg Inequality: Let $1 \leq q \leq r < \infty$. Then, for $s = 1 - (q/r)$,

$$\|u\|_{L^r(\Omega)} \leq C \|u\|^{1-s}_{L^q(\Omega)} \|u\|^s_{H^1(\Omega)}, \qquad \forall\, u \in H^1(\Omega).$$

Generalized Friedrichs Inequality: There exists $C_F > 0$ (depending only on Ω) such that:

$$\|\nabla y\|^2_{L^2(\Omega)} + \|y\|^2_{L^2(\Gamma)} \geq C_F \|y\|^2_{H^1(\Omega)}.$$

In the remaining of this section we quote several results regarding the solvability of the continuous optimal control problem.

2.2 The continuous optimal control problem

First we quote a result (see [12]) regarding the solvability of the uncontrolled weak problem (2.1) at the natural energy space.

Theorem 2.3. *Let $f \in L^2[0, T; H^1(\Omega)^*]$, $y_0 \in L^2(\Omega)$, $g \in L^2[0, T; H^{-1/2}(\Omega)]$. Then, there exists a unique solution $y \in W(0, T)$ of (2.1) which satisfies the following energy estimate*

$$\|y\|_{W(0,T)} \leq C\Big(\|f\|_{L^2[0,T;H^1(\Omega)^*]} + \|y_0\|_{L^2(\Omega)} + \|g\|_{L^2[0,T;H^{-1/2}(\Gamma)]}\Big).$$

In addition, $y_t \in L^2[0, T; H^1(\Omega)^*]$.

A well known embedding theorem (see e.g. [51]) implies that $y \in C[0, T; L^2(\Omega)]$. Therefore, the energy optimal boundary control problem (**EBCP**)can be defined as follows:

Definition 2.4. *Let $f \in L^2[0, T; H^1(\Omega)^*]$, $y_0 \in L^2(\Omega)$ be given data.*

1. *The pair (y, g) is said to be an admissible element (pair) if $y \in W(0, T)$, $g \in L^2[0, T; L^2(\Gamma)]$ satisfy (2.1). The set of admissible pairs is denoted by $\mathcal{A}_{ad}$.*
2. *The pair $(y, g) \in \mathcal{A}_{ad}$ is said to be a (locally) optimal solution if $J(y, g) \leq J(w, h)$ $\forall\,(w, h) \in \mathcal{A}_{ad}$, when $\|y - w\|_{W(0,T)} + \|g - h\|_{L^2[0,T;L^2(\Gamma)]} \leq \delta$ for $\delta > 0$ appropriately chosen.*

Below, we state the main result (see [12, Theorem 2.5]) concerning the existence of an optimal solution for the minimization of the functional (1.1).

Theorem 2.5. *Let $y_0 \in L^2(\Omega)$, $f \in L^2[0, T; (H^1(\Omega))^*]$ be given data. Then, the boundary control problem* **(EBCP)** *has solution $(y, g) \in W(0, T) \times L^2[0, T; L^2(\Gamma)]$.*

Taking into account the previous result, the corresponding optimality system can be easily derived based on well known techniques (see e.g. [12, 20, 34, 41, 48]). Below, we state the related optimality system. For problem **(EBCP)**, we seek state $y \in W(0,T)$ and adjoint $\mu \in W(0,T)$ such that all $v \in L^2[0,T;H^1(\Omega)] \cap H^1[0,T;H^1(\Omega)^*]$,

$$\begin{aligned}(y(T),v(T)) \quad &+ \quad \int_0^T (-\langle y, v_t\rangle + a(y,v) + \langle \phi(y), v\rangle + \lambda\langle y, v\rangle_\Gamma)\, dt \\ &= \quad (y_0, v(0)) + \int_0^T (\langle f, v\rangle + \lambda\langle g, v\rangle_\Gamma)\, dt,\end{aligned} \tag{2.2}$$

$$(\mu(0),v(0)) + \int_0^T \big(\langle \mu, v_t\rangle + a(\mu,v) + \langle \phi'(y)\mu, v\rangle + \lambda\langle \mu, v\rangle_\Gamma\big)\, dt = \int_0^T (\nabla y, \nabla v) dt, \tag{2.3}$$

and for all $u \in L^2[0,T;L^2(\Gamma)]$,

$$\int_0^T (\alpha g + \lambda\mu, u)_\Gamma dt = 0. \tag{2.4}$$

The optimality system consists of a forward in time parabolic equation, a backward in time parabolic equation which are coupled, through an optimality condition and nonlinear terms. It is evident that due to the presence of the gradient term at the adjoint equation the available regularity is restricted. However, the use of the $\|.\|_{L^2[0,T;L^2(\Gamma)]}$ norm in the functional is an important asset since it leads to an algebraic optimality condition on the boundary.

Remark 2.6. *Higher regularity can be obtained for the optimality system, via a bootstrap argument for more regular data, for certain type of nonlinear terms. For example, let Assumption 2.1 be satisfied with* $1 \le p \le 2$ *(or* ϕ *be Lipschitz), data* $y_0 \in H^1_0(\Omega)$*,* $f \in L^2[0,T;L^2(\Omega)]$*, and* Ω *be a domain with smooth boundary* $\Gamma \in C^2$ *or* Ω *polygonal and convex. Then, the fact that* $g \in L^2[0,T;L^2(\Gamma)]$ *and* $\phi(y) \in L^2[0,T;L^2(\Omega)]$ *(via a simple interpolation argument) implies that* $y \in L^2[0,T;H^{3/2}(\Omega)] \cap H^{1/2}[0,T;L^2(\Omega)]$*, and* $\Delta y \in L^2[0,T;H^{1/2}(\Omega)^*]$*. Therefore,* $\mu \in L^2[0,T;H^{3/2}(\Omega)] \cap H^{1/2}[0,T;L^2(\Omega)]$*, from which we deduce that* $g \in L^2[0,T;H^1(\Gamma)]$*. Alternatively, if we choose to include other norms of the control* g *in the functional we may gain some extra regularity for the control function, provided that the initial data are sufficiently regular. In the later case, the optimality condition typically involves a PDE on the boundary which is not computationally attractive. However, throughout this work we will restrict ourselves into the minimal regularity case as mentioned in the introduction.*

Remark 2.7. *We refer the reader to the book of Tröltzsch [48] (see also references within), for a comprehensive survey of various results regarding necessary and sufficient conditions for the existence of optimal control problems.*

Remark 2.8. *Several results presented in this work are excepted to hold for three-dimensional domains under more restrictive assumptions on the growth conditions of the nonlinearity. Indeed, the dimensionality of the domain, typically enters into the proofs via embedding theorems. The three-dimensional case will be investigated elsewhere.*

3 The discrete optimal control problem

Let $U_h \subset H^1(\Omega)$ be a family of finite element subspaces defined over regular triangulations of Ω, where h denotes the largest grid size for a given trianglulation, satisfying the classical approximation theory properties (see e.g. [13]).

Approximations will be constructed on a (quasi-uniform) partition $0 = t^0 < t^1 < \ldots < t^N = T$ of $[0, T]$, i.e., there exists a constant $0 < \theta < 1$ such that $\min_{n=1,..,N}(t^n - t^{n-1}) \leq \theta \max_{n=1,...,N}(t^n - t^{n-1})$. We will occasionally use the notation $\tau^n = t^n - t^{n-1}, \tau = \max_{n=1,...,N} \tau^n$ and we denote by $\mathcal{P}_k[t^{n-1}, t^n; U_h]$ the space of polynomials of degree k or less having values in U_h. We seek approximate solutions who belong to the space

$$\mathcal{U}_h = \{y_h \in L^2[0, T; H^1(\Omega)] : y_h|_{(t^{n-1},t^n]} \in \mathcal{P}_k[t^{n-1}, t^n; U_h]\}.$$

In the above definitions, we have used the following notational abbreviation, $y_{h,\tau} \equiv y_h$, $\mathcal{U}_{h,\tau} \equiv \mathcal{U}_h$ etc. We also note that throughout this work the discretization parameters τ and h can be chosen independently. By convention, the functions of $\mathcal{U}_h$ are left continuous with right limits and hence will subsequently write y^n for $y_h(t^n) = y_h(t^n_-)$, and y^n_+ for $y(t^n_+)$, while the jump at t^n, is denoted by $[y^n] = y^n_+ - y^n$. For the control variable, motivated by the optimality condition, we will use a similar discretization which allows the presence of discontinuities (in time), i.e., we define,

$$\mathcal{G}_h = \{g_h \in L^2[0, T; L^2(\Gamma)] : g_h|_{(t^{n-1},t^n]} \in \mathcal{P}_k[t^{n-1}, t^n; G_h]\}.$$

where a conforming subspace $G_h \subset L^2(\Gamma)$ is specified at each time interval $(t^{n-1}, t^n]$. The subspace G_h satisfy standard approximation properties. For suitable choices, (e.g. $G_h = U_h|_\Gamma$) we refer the reader to [21, 25] (see also references within). In the subsequent analysis, we will only need that any $u \in L^2[0, T; L^2(\Gamma)]$ can be approximated by elements $\{u_h\}_{\tau,h} \in \mathcal{G}_h$ weakly in $L^2[0, T; L^2(\Gamma)]$ norm, as $\tau, h \to 0$. The fully-discrete approximation of the (uncontrolled) constraint equation is to seek $y_h \in \mathcal{U}_h$ such that for given $y^0 \in L^2(\Omega)$, $f \in L^2[0, T; H^1(\Omega)^*]$ and $g_h \in L^2[0, T; L^2(\Gamma)]$, the following equations hold for $n = 1, ..., N$: For all $v_h \in \mathcal{P}_k[t^{n-1}, t^n; U_h]$,

$$\begin{aligned}(y^n, v^n) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(-\langle y_h, v_{ht}\rangle + a(y_h, v_h) + (\phi(y_h), v_h) + \lambda\langle y_h, v_h\rangle_\Gamma \Big) dt \qquad (3.1)\\ &= \quad (y^{n-1}, v^{n-1}_+) + \int_{t^{n-1}}^{t^n} \Big(\langle f, v_h\rangle + \lambda\langle g_h, v_h\rangle_\Gamma\Big) dt.\end{aligned}$$

The discrete admissible set $\mathcal{A}^d_{ad}$ and the associated discrete optimal control problem (**DEBCP**) are defined similar to the continuous control problem (**EBCP**).

Definition 3.1. *Suppose that the assumptions of Section 2 hold.*

- *The discrete admissible set is defined by* $\mathcal{A}^d_{ad} \equiv \{(y_h, g_h) \in \mathcal{U}_h \times \mathcal{G}_h$ *such that (3.1) holds*$\}$.
- **(DEBCP):** *We seek pair* $(y_h, g_h) \in \mathcal{A}^d_{ad}$ *such that* $J(y_h, g_h) \leq J(w_h, k_h)$ *for all* $(w_h, k_h) \in \mathcal{A}^d_{ad}$ *when* $\|y_h - w_h\|_{L^2[0,T;H^1(\Omega)]} + \|y_h - w_h\|_{L^\infty[0,T;L^2(\Omega)]} + \|g_h - k_h\|_{L^2[0,T;L^2(\Gamma)]} \leq \delta$ *for* $\delta > 0$ *appropriately chosen.*

If we denote by $\tilde{y}_h$ the solution of (3.1) without control ,without loss of generality, it is understood that the pair $(\tilde{y}_h, 0) \in \mathcal{A}^d_{ad}$, and δ are chosen in a way to guarantee that $J(y_h, g_h) \leq J(\tilde{y}_h, 0)$.

Remark 3.2. *The proof utilizes only the $L^2[0, T; L^2(\Gamma)]$ regularity for g_h. However as we will show section 4, similar to the continuous case, the given functional implies that the control variable can be implicitly computed as the discontinuous Galerkin solution of an adjoint equation through an optimality condition $g_h = -(\lambda/\alpha)\mu_h|_\Gamma$ provided that the traces of space $\mathcal{U}_h$ are in $\mathcal{G}_h$.*

Remark 3.3. *For the uncontrolled problem the existence and (local) uniqueness can be proved around the continuous solution (in an appropriate 'parabolic" cube), provided that the semi-linear term satisfies suitable continuity and monotonicity assumptions which allow the application of standard fixed point theorems (see, e.g., [1, 15, 17, 45]). For the existence of higher-order discontinuous time-stepping approximations under minimal regularity assumptions, we also refer the reader to [49].*

In order to prove the existence of a discrete optimal solution, we need to obtain stability estimates for the state equation in the natural energy norm as well as at arbitrary time points. For the later, we construct an exponential interpolant (see e.g. [11]) of functions of the form $e^{-\rho(t-t^{n-1})}y_h$, when $y_h \in \mathcal{P}_k[t^{n-1}, t^n; U_h]$. The main advantage of this approach is that the proof does not need any additional regularity, apart from the one needed to guarantee the existence of a weak solution, i.e., we do not assume that $u_t \in L^2[0, T; L^2(\Omega)]$ which is frequently used in the literature for DG approximations of parabolic PDE's (even without controls), and which is not suitable in the present boundary control setting.

3.1 Quotation of results related to an exponential interpolant

The polynomial interpolant of functions of the form $e^{-\rho(t-t^{n-1})}y$, where $y \in \mathcal{P}_k[t^{n-1}, t^m; U]$ and U is any linear space, is defined as follows (see [11, Definition 3.3]).

Definition 3.4. *Let U be a linear space, and $\rho > 0$ be given. If $u = \sum_{i=0}^k r_i(t)u_i \in \mathcal{P}_k[t^{n-1}, t^n; U]$, with $r_i \in \mathcal{P}_k[t^{n-1}, t^n]$ and $u_i \in U$, we define the exponential interpolant of u by*

$$\bar{u} = \sum_{i=0}^{k} \bar{r}_i(t)u_i$$

where $\bar{r}_i \in \mathcal{P}_k[t^{n-1}, t^n]$ is the approximation of $r_i(t)e^{-\rho(t-t^{n-1})}$ satisfying $r_i(t^{n-1}) = \bar{r}_i(t^{n-1})$ and

$$\int_{t^{n-1}}^{t^n} \bar{r}_i(t)q(t)dt = \int_{t^{n-1}}^{t^n} r_i(t)q(t)e^{-\rho(t-t^{n-1})}dt, \qquad q \in \mathcal{P}_{k-1}[t^{n-1}, t^n].$$

The following Lemma (see [11, Lemma 3.4]) asserts that the difference $u - \bar{u}$ remains small in various norms.

Lemma 3.5. *Let U and Q be linear spaces and $u \to \bar{u}$ be the map constructed in Definition 3.4, for given $\rho > 0$. If $L(.,.) : U \times Q \to \mathbb{R}$ denotes a bilinear mapping and $v \in \mathcal{P}_k[t^{n-1}, t^n; U]$*

then

$$\int_{t^{n-1}}^{t^n} L(\bar{u}(t), q(t))dt = \int_{t^{n-1}}^{t^n} L(u(t), q(t))e^{-\rho(t-t^{n-1})}dt, \qquad \forall\, q \in \mathcal{P}_{k-1}[t^{n-1}, t^n; Q].$$

If $(.,.)_U$ is a (semi) inner product on U, then there exists a constant C_k independent of $\rho > 0$, such that

$$\|u - \bar{u}\|_{L^2[t^{n-1},t^n;U]} \le C_k \rho (t^n - t^{n-1}) \|u\|_{L^2[t^{n-1},t^n;U]}.$$

Using a scaling argument, the finite dimensionality of $\mathcal{P}_k[t^{n-1}, t^n]$ and an inverse estimate, we may use the previous Lemma to bound $u - \bar{u}$ in $L^p[t^{n-1}, t^n; U]$ norm.

Lemma 3.6. *Let U be a linear space and $(.,.)_U$ be a (semi) inner product on U and let $u \to \bar{u}$ denote the exponential interpolant on $\mathcal{P}_k[t^{n-1}, t^n; U]$ constructed in Definition 3.4. Then, there exists a constant C_k depending only on k such that,*

$$\|u - \bar{u}\|_{L^p[t^{n-1},t^n;U]} \le C_k \rho \tau \|u\|_{L^p[t^{n-1},t^n;U]},$$

for all $u \in \mathcal{P}_k[t^{n-1}, t^n; U]$ and $1 \le p \le \infty$.

Proof. See [11, Lemma 3.5, 3.6]. □

3.2 Compactness properties of DG time-stepping schemes

The following result of Walkington, [49, Theorem 3.1]), will allow us to use compactness when applying DG schemes of arbitrary order. Recall that due to the presence of discontinuities the discrete time-derivative is not integrable, and hence the recovery of strong convergence by the means of a suitable Aubin-Lions Lemma (see [44]) is not evident. Below, we quote the necessary compactness result for the fully-discrete case, when conforming finite element subspaces are being used. The problems considered in [49], involve the numerical approximations of solutions $u : [0, T] \to U$ of general nonlinear evolution equations of the form

$$u_t + A(u) = f(u) \qquad u(0) = u_0 \tag{3.2}$$

where U is a Banach space and each term of the equation takes values in U^*. Here, both $A(u) = A(t, u)$ and $f(u) = f(t, u)$ may depend upon t and are allowed to be nonlinear. We assume that $U \subset H \subset U^*$ (with continuous embeddings) form the standard evolution triple, i.e., the pivot space H is a Hilbert space. The numerical schemes approximate the weak form of (3.2), i.e.,

$$\langle u_t, v\rangle + a(u, v) = \langle f(u), v\rangle, \qquad \forall\, v \in U \tag{3.3}$$

where $a : U \times U \to \mathbb{R}$ is defined by $a(u, v) = \langle A(u), v\rangle$. Recall, that for each subspace $U_h \subset U$ and partition $0 = t^0 < t^1 < ... < t^N = T$ of $[0, T]$ the DG scheme constructs a function in $\mathcal{P}_k[t^{n-1}, t^n; U_h]$ on each (t^{n-1}, t^n), which satisfies for $n = 1, ..., N$ and for all $v_h \in \mathcal{P}_k[t^{n-1}, t^n; U_h]$,

$$\int_{t^{n-1}}^{t^n} \Big((u_{ht}, v_h) + a(u_h, v_h) \Big) dt + (u_+^{n-1} - u^{n-1}, v_+^{n-1}) = \int_{t^{n-1}}^{t^n} (f(u_h), v_h) dt. \tag{3.4}$$

Here, u_h^0 is a given approximation of u_0. The following theorem [49, Theorem 3.1] establishes compactness property of the discrete approximation.

Theorem 3.7. *Let H be a Hilbert space, U be a Banach space and $U \subset H \subset U^*$ be dense and compact embeddings. Fix integer $k \geq 0$ and $1 \leq p, q < \infty$. Let $h > 0$ be the mesh parameter, and let $\{t^i\}_{i=0}^N$ denote a quasi-uniform partition of $[0,T]$. Set $F(u) \equiv f(u) - A(u)$. Assume that*

1. $u_h \in \{u_h \in L^p[0,T;U] \quad | \quad u_h|_{(t^{n-1},t^n)} \in \mathcal{P}_k[t^{n-1}, t^n; U_h]\}$ *and on each interval,*
$$\int_{t^{n-1}}^{t^n} (u_{ht}, v_h)dt + (u_+^{n-1} - u^{n-1}, v_+^{n-1}) = \int_{t^{n-1}}^{t^n} (F(u_h), v_h)dt$$
holds for every $v_h \in \mathcal{P}_k[t^{n-1}, t^n; U_h]$.
2. $\{u_h\}_{h>0}$ *is bounded in* $L^p[0,T;U]$ *and* $\{\|F(u_h)\|_{L^q[0,T;U^*]}\}_{h>0}$ *is also bounded.*

Then,

1. *If* $p > 1$ *then* $\{u_h\}_{h>0}$ *is compact in* $L^r[0,T;H]$ *for* $1 \leq r < 2p$.
2. *If* $1 \leq (1/p) + (1/q) < 2$, *and* $\sum_{i=1}^N \|[u_h]\|_H^2 < C$ *is bounded independent of h, then* $\{u_h\}_{h>0}$ *is compact in* $L^r[0,T;H]$ *for* $1 \leq r < 2/((1/p) + (1/q) - 1)$.

Remark 3.8. *The above Theorem reduces the proof of compactness for discrete time-stepping schemes to the verification of suitable embedding results for the nonlinear terms.*

3.3 The main stability estimate

First, we present the stability estimate related to the minimization of the energy functional. The stability estimate demonstrates that the DG approximations inherit the basic regularity structure of the underlying continuous weak solution. Similar to the proof of [8, Lemma 3.6] we do not assume any $\|y_t\|_{L^2[0,T;L^2(\Omega)]}$ regularity which is frequently used in the literature, even for low order schemes. The key idea in order to handle the low regularity of y_h and the Robin boundary data, is the use of the exponential interpolant of Section 3.1 (see also e.g. [11] in case of Navier-Stokes equations).

Lemma 3.9. *Suppose that $y_0 \in L^2(\Omega)$, $f \in L^2[0,T;H^1(\Omega)^*]$ are given functions, and let ϕ satisfy assumption 2.1. If $(y_h, g_h) \in \mathcal{U}_h \times L^2[0,T;G_h]$ denotes a solution of the discrete optimal control problem* **(DEBCP)**, *then*

$$\begin{aligned} J(y_h, g_h) \quad &\equiv \quad \int_0^T \|\nabla y_h\|_{L^2(\Omega)}^2 dt + (\alpha/2)\int_0^T \|g_h\|_{L^2(\Gamma)}^2 dt \\ &\leq \quad C\Big(\|y^0\|_{L^2(\Omega)}^2 + (1/\eta C_F \min\{\eta,\lambda\})\int_0^T \|f\|_{H^1(\Omega)^*}^2 dt\Big) \equiv C_{st} \end{aligned}$$

where C is a constant depending only on Ω. In addition, for all $n = 1, ..., N$

$$\begin{aligned} \|y^n\|_{L^2(\Omega)}^2 + \sum_{i=0}^{n-1} \|[y^i]\|_{L^2(\Omega)}^2 \quad &+ \quad \int_0^{t^n} \Big(C_F \min\{\eta,\lambda\}\|y_h\|_{H^1(\Omega)}^2 + \|y_h\|_{L^{p+1}(\Omega)}^{p+1} + \lambda\|y_h\|_{L^2(\Gamma)}^2\Big)dt \\ &\leq \quad D_{yst} \equiv C_{st}\max\{1, \lambda/\alpha\}. \end{aligned}$$

Let $\tau \equiv \max_{i=1,..,N} \tau_i$*, with* $\tau_i = t^i - t^{i-1}$*. If* $\tau \leq \left(1/8D_{yst}^{(p-1)/2} C_k\right)^{2/(3-p)}$*, then*

$$\|y_h\|^2_{L^\infty[0,T;L^2(\Omega)]} \leq C D_{yst}$$

where C *depends on* $(1/C_F \min\{\eta, \lambda\})$,$C_k$ *and* Ω *but not on* α, τ, h.

Proof. The first estimate can be easily derived by noting that $(\tilde{y}_h, 0)$ is an admissible pair for the discrete problem, and hence $J(y_h, g_h) \leq J(\tilde{y}_h, 0) \leq (1/2)\int_0^T \|\nabla \tilde{y}_h\|^2_{L^2(\Omega)} \leq C_{st}$. Here C_{st} is a constant independent of α. The estimate on $\|\nabla \tilde{y}_h\|_{L^2[0,T;L^2(\Omega)]}$ can be obtained identically to [10, Section 2]) since it corresponds to the stability estimate without control.

Setting $v_h = y_h$ into (3.1), using the monotonicity of ϕ and Young's and Friedrichs's inequalities, we obtain

$$\begin{aligned}
&(1/2)\|y^n\|^2_{L^2(\Omega)} + (1/2)\|[y^{n-1}]\|^2_{L^2(\Omega)} - (1/2)\|y^{n-1}\|^2_{L^2(\Omega)} \\
&+ \int_{t^{n-1}}^{t^n} \Big((C_F \min\{\eta, \lambda\}/2)\|y_h\|^2_{H^1(\Omega)} + (\lambda/4)\|y_h\|^2_{L^2(\Gamma)} + C\|y_h\|^{p+1}_{L^{p+1}(\Omega)} \Big) dt \\
&\leq \int_{t^{n-1}}^{t^n} \Big((1/C_F \min\{\eta, \lambda\})\|f\|^2_{H^1(\Omega)^*} + \lambda \|g_h\|^2_{L^2(\Gamma)} \Big) dt. \qquad (3.5)
\end{aligned}$$

Summing the resulting inequalities from $i = 1$ to n, and dropping positive terms on the left we obtain the estimate at partition points by using the previous bound on $\alpha \int_0^T \|g_h\|^2_{L^2(\Gamma)} dt$. The estimate at the energy norm follows upon summation from 1 to N. It remains to obtain a bound at arbitrary time-points. To achieve this, we will use the exponential interpolant of $e^{-\rho(t-t^{n-1})} y_h$ of Definition 3.4, denoted by $\bar{y}_h$. Then, using the definition of $\bar{y}_h$, we obtain that

$$\begin{aligned}
&\int_{t^{n-1}}^{t^n} (y_{ht}, \bar{y}_h) = \int_{t^{n-1}}^{t^n} (y_{ht}, y_h) e^{-\rho(t-t^{n-1})} \qquad (3.6) \\
&= (1/2)\|y^n\|^2_{L^2(\Omega)} e^{-\rho(t^n - t^{n-1})} - (1/2)\|y^{n-1}\|^2_{L^2(\Omega)} + (\rho/2)\int_{t^{n-1}}^{t^n} \|y_h(t)\|^2_{L^2(\Omega)} e^{-\rho(t-t^{n-1})} dt.
\end{aligned}$$

Integrating by parts (in time) (3.1), setting $v_h = \bar{y}_h$ and using (3.6) we obtain

$$\begin{aligned}
&(1/2)\|y^n\|^2_{L^2(\Omega)} e^{-\rho(t^n - t^{n-1})} + (1/2)\|[y^{n-1}]\|^2_{L^2(\Omega)} - (1/2)\|y^{n-1}\|^2_{L^2(\Omega)} \\
&+(\rho/2)\int_{t^{n-1}}^{t^n} \|y_h\|^2_{L^2(\Omega)} e^{-\rho(t-t^{n-1})} dt + \int_{t^{n-1}}^{t^n} \langle \phi(y_h), \bar{y}_h \rangle dt \\
&\leq \int_{t^{n-1}}^{t^n} \Big(|a(y_h, \bar{y}_h)| + \lambda |\langle y_h, \bar{y}_h \rangle_\Gamma| + |\langle f, \bar{y}_h \rangle| + \lambda |\langle g_h, \bar{y}_h \rangle_\Gamma| \Big) dt.
\end{aligned}$$

Using Lemma 3.5, we may bound $\bar{y}_h$ in terms of y_h in various norms. In particular, using Young's inequalities with appropriate $\delta > 0$,

$$\begin{aligned}
&\int_{t^{n-1}}^{t^n} |a(y_h, \bar{y}_h)| dt \leq C_k \eta \int_{t^{n-1}}^{t^n} \|y_h\|^2_{H^1(\Omega)} dt, \\
&\int_{t^{n-1}}^{t^n} |\langle f, \bar{y}_h \rangle| dt \leq (C_k / C_F \min\{\eta, \lambda\}) \int_{t^{n-1}}^{t^n} \|f\|^2_{H^1(\Omega)^*} dt + C_F \min\{\eta, \lambda\} \int_{t^{n-1}}^{t^n} \|y_h\|^2_{H^1(\Omega)} dt, \\
&\int_{t^{n-1}}^{t^n} \lambda |\langle g_h, \bar{y}_h \rangle_\Gamma| + \lambda |\langle y_h, \bar{y}_h \rangle_\Gamma| dt \leq \alpha \int_{t^{n-1}}^{t^n} \|g_h\|^2_{L^2(\Gamma)} dt + C_k(\lambda + (\lambda^2/\alpha)) \int_{t^{n-1}}^{t^n} \|y_h\|^2_{L^2(\Gamma)} dt.
\end{aligned}$$

Therefore, collecting the above inequalities and using standard algebra, we obtain,

$$
\begin{aligned}
&(1/2)\|y^n\|^2_{L^2(\Omega)}e^{-\rho(t^n-t^{n-1})} + (1/2)\|[y^{n-1}]\|^2_{L^2(\Omega)} - (1/2)\|y^{n-1}\|^2_{L^2(\Omega)} \\
&+(\rho/2)\int_{t^{n-1}}^{t^n} \|y_h\|^2_{L^2(\Omega)}e^{-\rho(t-t^{n-1})}dt + \int_{t^{n-1}}^{t^n}\langle\phi(y_h),\bar{y}_h\rangle dt \\
&\leq C_k\int_{t^{n-1}}^{t^n}\Big((1/C_F\min\{\eta,\lambda\})\|f\|^2_{H^1(\Omega)^*} + (\eta + C_F\min\{\eta,\lambda)\})\|y_h\|^2_{H^1(\Omega)} + \alpha\|g_h\|^2_{L^2(\Gamma)} \\
&\qquad +(\lambda+(\lambda^2/\alpha))\|y_h\|^2_{L^2(\Gamma)}\Big)dt.
\end{aligned}
$$

It remains to bound the semi-linear term. For this purpose, note first that assumption 2.1 implies that

$$
\int_{t^{n-1}}^{t^n}\langle\phi(y_h),\bar{y}_h\rangle dt \geq \int_{t^{n-1}}^{t^n}\langle\phi(y_h),\bar{y}_h - y_h\rangle dt.
$$

The growth condition and Young's inequality with $s_1 = (p+1)/p$, $s_2 = p+1$, and for $\delta > 0$ imply

$$
\int_{t^{n-1}}^{t^n}\langle\phi(y_h),\bar{y}_h - y_h\rangle dt \leq \int_{t^{n-1}}^{t^n}\|y_h\|^{p+1}_{L^{p+1}(\Omega)}dt + C(p)\int_{t^{n-1}}^{t^n}\|\bar{y}_h - y_h\|^{p+1}_{L^{p+1}(\Omega)}dt
$$

For the last term on the right hand side, the Gagliardo-Nirenberg interpolation inequality with $q = 2$, $r = p+1$, $s = 1-(2/p+1) = (p-1)/(p+1)$, states that

$$
\begin{aligned}
&\int_{t^{n-1}}^{t^n}\|\bar{y}_h - y_h\|^{p+1}_{L^{p+1}(\Omega)}dt \leq C\int_{t^{n-1}}^{t^n}\|\bar{y}_h - y_h\|^2_{L^2(\Omega)}\|\bar{y}_h - y_h\|^{p-1}_{H^1(\Omega)}dt \\
&\leq C\|\bar{y}_h - y_h\|^2_{L^\infty[t^{n-1},t^n;L^2(\Omega)]}\Big(\int_{t^{n-1}}^{t^n}1dt\Big)^{(3-p)/2}\Big(\int_{t^{n-1}}^{t^n}\|\bar{y}_h - y_h\|^2_{H^1(\Omega)}dt\Big)^{(p-1)/2} \\
&\quad \leq C_k\tau_n^2\rho^2\|y_h\|^2_{L^\infty[t^{n-1},t^n;L^2(\Omega)]}\tau_n^{(3-p)/2}\rho^{p-1}\tau_n^{p-1}\|y_h\|^{p-1}_{L^2[t^{n-1},t^n;H^1(\Omega)]}.
\end{aligned}
$$

Here we have used the generalized Hölder inequality with $s_1 = 2/(p-1) > 1$, $s_2 = 2/(3-p) > 1$ (recall $1 < p < 3$), Lemma 3.5 to bound $\bar{y}_h - y_h$ in terms of y_h, and the stability estimates at the energy norm. Hence, selecting $\rho = (1/\tau_n)$ we obtain,

$$
\begin{aligned}
&(1/2)\|y^n\|^2_{L^2(\Omega)}e^{-1} + (1/2)\|[y^{n-1}]\|^2_{L^2(\Omega)} - (1/2)\|y^{n-1}\|^2_{L^2(\Omega)} + (e^{-1}/2\tau_n)\int_{t^{n-1}}^{t^n}\|y_h\|^2_{L^2(\Omega)}dt \\
&\leq C_k\Big[\int_{t^{n-1}}^{t^n}\Big((1/C_F\min\{\eta,\lambda\})\|f\|^2_{H^1(\Omega)^*} + (\eta + C_F\min\{\eta,\lambda)\|y_h\|^2_{H^1(\Omega)} + \alpha\|g_h\|^2_{L^2(\Gamma)} \\
&\qquad +(\lambda+(\lambda^2/\alpha))\|y_h\|^2_{L^2(\Gamma)}\Big)dt + \|y_h\|^2_{L^\infty[t^{n-1},t^n;L^2(\Omega)]}\tau_n^{(3-p)/2}D_{yst}^{p-1}\Big].
\end{aligned}
$$

where we have used the previously developed estimate for the energy norm. Using the inverse estimate $\|y_h(t)\|^2_{L^2(\Omega)} \leq C_k/\tau_n\int_{t^{n-1}}^{t^n}\|y_h(s)\|^2_{L^2(\Omega)}dt$ and choosing τ_n such that to hide the last term on the left, $\tau_n^{(3-p)/2}D_{yst}^{(p-1)/2} \leq (C_k/8)$, i.e., for $\tau_n \leq \big(1/8D_{yst}^{(p-1)/2}C_k\big)^{2/3-p}$ we finally arrive to

$$
\begin{aligned}
\|y_h\|^2_{L^\infty[t^{n-1},t^n;L^2(\Omega)]} &\le C_k\Big[\|y^{n-1}\|^2_{L^2(\Omega)} \\
&+\int_{t^{n-1}}^{t^n}\Big((1/C_F\min\{\eta,\lambda\})\|f\|^2_{H^{-1}(\Omega)}+(\eta+C_F\min\{\eta,\lambda\})\|y_h\|^2_{H^1(\Omega)} \\
&\quad+\|y_h\|^{p+1}_{L^{p+1}(\Omega)}+\alpha\|g_h\|^2_{L^2(\Gamma)}+(\lambda+(\lambda^2/\alpha))\|y_h\|^2_{L^2(\Gamma)}\Big)dt\Big].
\end{aligned}
$$

The estimate now follows by using the previously derived estimates for the energy norm and at partition points. □

Remark 3.10. *We remark that the above technique does not rely on any use of discrete Gröwall Lemma, hence we avoid the exponential dependence upon $(1/\alpha)$ of the stability constant. It is also applicable in the three-dimensional case for suitable values of p, by using appropriate modifications on the embedding and interpolation inequalities.*

3.4 Convergence of the discrete optimal solution

Combining the stability estimates of Lemma 3.9 with the discrete compactness Theorem 3.7 we prove the existence of the discrete optimal solution and its convergence to the continuous one.

Theorem 3.11. *Suppose that $f \in L^2[0,T;H^1(\Omega)^*]$, $y_0 \in L^2(\Omega)$.*

1. *Let $h > 0$ and quasi-uniform partition $\{t^i\}_{i=0}^N$ of $[0,T]$ fixed, with $\tau = \max_{i=1,\dots,N}\tau_i$, $\tau_i = t^i - t^{i-1}$, satisfying the assumptions of Lemma 3.9. Then, for $\alpha > 0$, there exists solution $(y_h, g_h) \in \mathcal{U}_h \times \mathcal{G}_h$ of problem (**DEBCP**), i.e. pair (y_h, g_h) that satisfies the discrete equation (3.1) and the functional (1.1) is minimized.*

2. *Given quasi-uniform partition $\{t^i\}_{i=0}^N$ of $[0,T]$, with $\tau = \max_{i=1,\dots,N}\tau_i$, $\tau_i = t^i - t^{i-1}$ satisfying the assumptions of Lemma 3.9, let $\tau, h \to 0$. Then, for $\alpha > 0$, (y_h, g_h) converges to a (local) optimal pair (y, g) of problem (**EBCP**), in the following sense,*

$$
y_h \to y \quad \textit{weakly in } L^2[0,T;H^1(\Omega)], \quad y_h \to y \quad \textit{weakly-* in } L^\infty[0,T;L^2(\Omega)],
$$
$$
g_h \to g \quad \textit{weakly in } L^2[0,T;L^2(\Gamma)], \quad y_h \to y \quad \textit{weakly in } L^2[0,T;L^2(\Gamma)].
$$

and

$$
y_h \to y \quad \textit{strongly in } L^2[0,T;L^2(\Omega)].
$$

Proof. The proof follows the argument of [8, Theorem 4.1].
1. *(Sketch)* Let $h > 0$ and $0 = t^0 < t^1, \dots, t^N = T$ be a fixed partition of [0,T] with τ, h satisfying the assumptions of Lemma 3.9. It is easy to see that the element $(\tilde{y}_h, 0)$ and hence the discrete admissible set $\mathcal{A}^d_{ad} \neq 0$. Now let $(y_{hm}, g_{hm}) \in \mathcal{A}^d_{ad}$ be minimizing sequence where y_{hm} denotes the corresponding solution of (3.1) with right hand side g_{hm}. For example, we may extract a minimizing sequence such that $J(y_{hm}, g_{hm}) \le M$, with M be the value of the functional for an admissible element, say $(\tilde{y}_h, 0)$. Hence, the stability estimates (independent of h,τ) imply that (passing to a subsequence, if necessary), as $m \to \infty$,

$$
y_{hm} \to y_h \quad \text{weakly in } L^2[0,T;H^1(\Omega)], \quad y_{hm} \to y_h \quad \text{weakly-* in } L^\infty[0,T;L^2(\Omega)],
$$

$$g_{hm} \to g_h \quad \text{weakly in } L^2[0,T;L^2(\Gamma)] \quad y_{hm} \to y_h \quad \text{weakly in } L^2[0,T;L^2(\Gamma)]$$

The proof now follows by using standard arguments and the finite dimensionality of the subspaces. We may pass to the limit to show that $(y_h, g_h) \in \mathcal{A}^d_{ad}$ satisfy the discrete equation (3.1) (see also part *(2)*). The weak lower semi-continuity of the functional finishes the proof.

2. The stability Lemma (3.9) implies that y_h, g_h are bounded in $L^\infty[0,T;L^2(\Omega)] \cap L^2[0,T;H^1(\Omega)]$, $L^2[0,T;L^2(\Gamma)]$ respectively by constants independent of τ, h. Hence, we may extract subsequences, converging weakly to some elements (y, g) respectively in the following sense, as $h, \tau \to 0$,

$$y_h \to y \quad \text{weakly in } L^2[0,T;H^1(\Omega)], \quad y_h \to y \quad \text{weakly-* in } L^\infty[0,T;L^2(\Omega)],$$

$$g_h \to g \quad \text{weakly in } L^2[0,T;L^2(\Gamma)], \quad y_h \to y \quad \text{weakly in } L^2[0,T;L^2(\Gamma)].$$

Using the discrete compactness Theorem 3.7 we will prove the strong convergence of y_h to y in $L^2[0,T;L^2(\Omega)]$. For that purpose set $U = H^1(\Omega)$, $H = L^2(\Omega)$ and define,

$$\langle F(y), v\rangle = -a(y,v) - \langle \phi(y), v\rangle - \lambda\langle y, v\rangle_\Gamma + \langle f, v\rangle + \lambda\langle g, v\rangle_\Gamma, \qquad \forall\, y, v \in H^1(\Omega).$$

It is evident by the stability Lemma 3.9 and in particular by the estimates on y_h in $L^2[0,T;H^1(\Omega)]$, $L^\infty[0,T;L^2(\Omega)]$, $L^2[0,T;L^2(\Gamma)]$ on g_h in $L^2[0,T;L^2(\Gamma)]$, and by the assumptions on the semi-linear term, that $\{\|y_h\|_{L^2[0,T;H^1(\Omega)]}\}_{h,\tau}$, $\{\|F(y_h)\|_{L^{4/3}[0,T;H^1(\Omega)^*]}\}_{h,\tau}$ remain bounded independent of h, τ. Indeed, to bound the later term, we only need to consider the semi-linear and boundary terms. Lemma A.1 implies that $\|\phi(y_h)\|_{L^{4/3}[0,T;H^1(\Omega)^*]}$ remains bounded with constant independent of h, τ. To treat the boundary terms, recall that $\|g_h\|_{L^2[0,T;L^2(\Gamma)]}$, $\|y_h\|_{L^2[0,T;L^2(\Gamma)]}$ are bounded independent of τ, h due to Lemma 3.9. A well known trace theorem (see e.g. [21, Theorem 1.5, pp 8]), implies that

$$\int_0^T \int_\Omega |g_h||v| d\Gamma dt \le C\|g_h\|_{L^2[0,T;L^2(\Gamma)]}\|v\|_{L^2[0,T;L^2(\Gamma)]} \le C\|g_h\|_{L^2[0,T;L^2(\Gamma)]}\|v\|_{L^2[0,T;H^1(\Omega)]},$$

and hence we obtain the desired bound on $\{\|F(y_h)\|_{L^{4/3}[0,T;H^1(\Omega)^*]}\}$. Therefore, we may apply Theorem 3.7 with $p = 2, q = 4/3, r = 2$ and the strong convergence of y_h to y is proven in $L^2[0,T;L^2(\Omega)]$ norm. It remains to show that (y, g) defined as above is optimal pair for (**EBCP**). Recall that $(y_h, g_h) \in \mathcal{A}^d_{ad}$ and hence it satisfies (3.1). We need to prove that the limit (y, g) satisfies (2.1). Suppose now that we choose $v_h \in C[0,T;U_h] \cap \mathcal{U}_h$. Then, summing equations (3.1) from $n = 1$ to $n = N$, we deduce that

$$\begin{aligned}(y_h(T), v_h(T)) + \int_0^T \Big(- \langle y_h, v_{ht}\rangle + a(y_h, v_h) + \langle \phi(y_h), v_h\rangle + \lambda\langle y_h, v_h\rangle_\Gamma\Big) dt \\ = \int_0^T \Big(\langle f, v_h\rangle + \lambda\langle g_h, v_h\rangle_\Gamma\Big) dt + (y^0, v_h(0)).\end{aligned}$$

Note that we may pass the limit through the linear terms since $v_h \in C[0,T;U_h] \cap \mathcal{U}_h$ and $v_h|_\Gamma \in L^2[0,T;L^2(\Gamma)]$. The semilinear term is treated in Lemma A.1. Hence, passing to the limit we obtain equation (2.1), by a standard density argument. Then, the weak lower semi-continuity of the functional finishes the proof, after noting that any element of $\mathcal{A}_{ad}$, can be approximated by a sequence of elements of $\mathcal{A}^d_{ad}$. Indeed, for any

$(w,r) \in \mathcal{A}_{ad}$, we construct $r_h \in \mathcal{G}_h$ such that $r_h \to r$, weakly in $L^2[0,T;L^2(\Gamma)]$. Then, let $w_h \in \mathcal{U}_h$ denote the discontinuous Galerkin approximation constructed by the solution of (3.1) with Robin boundary data given by $r_h \in L^2[0,T;L^2(\Gamma)]$. Then, exactly the same arguments as above, imply $w_h \in L^\infty[0,T;L^2(\Omega)] \cap L^2[0,T;H^1(\Omega)]$ and $w_h \to w$ strongly in $L^2[0,T;L^2(\Omega)]$ (by using the compactness argument). Therefore, we deduce that $J(y,g) \le \liminf_{\tau,h\to 0} J(y_h,g_h) \le \limsup_{\tau,h\to 0} J(y_h,g_h) \le \limsup_{\tau,h\to 0} J(w_h,r_h) \le J(w,r)$, which completes our proof. □

Remark 3.12. *The key feature of the proof is that the regularity of the discontinuous Galerkin approximation mimics the one of the continuous problem under minimal regularity assumptions.*

4 The discrete optimality system

Using techniques of Calculus of Variations, and the stability estimates of Section 3, we are able to obtain the discrete adjoint equation, and an optimality condition (first order necessary condition). Here, the stability estimates at arbitrary time points under minimal regularity, will allow to establish the discrete analog of the approach of Gunzburger and Manservisi [27, Section 2]. In order to prove the availability of the first order condition we need first to establish the Gâteaux differentiability for all g_h, of the map $y_h(g_h) : L^2[0,T;L^2(\Gamma)] \to L^2[0,T;H^1(\Omega)]$, where $y_h(g_h)$ denotes the DG solution of (3.1) when $f \in L^2[0,T;H^1(\Omega)^*]$ and y^0 are given. In the subsequent Lemma, we prove the Gâteaux derivative for all directions $u_h \in L^2[0,T;L^2(\Gamma)]$.

Lemma 4.1. *Let $f \in L^2[0,T;H^1(\Omega)^*]$, $y^0 \in U_h$ denotes an approximation of $y(0) \in L^2(\Omega)$, and $g_h \in L^2[0,T;L^2(\Gamma)]$. Suppose also that the assumptions of Lemma 3.9 hold. Then, the mapping $y_h(g_h) : L^2[0,T;L^2(\Gamma)] \to L^2[0,T;H^1(\Omega)]$ has a Gâteaux derivative in every direction u_h denoted by $w_h \equiv w_h(u_h) \equiv \frac{Dy_h}{Dg_h} \cdot u_h$, which is the solution of the problem: For all $n = 1,...,N$, and for all $v_h \in \mathcal{P}_k[t^{n-1},t^n;U_h]$,*

$$\begin{aligned}(w^n,v^n) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(-\langle w_h, v_{ht}\rangle + a(w_h,v_h) + (\phi'(y_h)w_h, v_h) + \lambda\langle w_h, v_h\rangle_\Gamma \Big) dt \qquad (4.1)\\ &= \quad (w^{n-1}, v_+^{n-1}) + \int_{t^{n-1}}^{t^n} \lambda \langle u_h, v_h\rangle_\Gamma dt.\end{aligned}$$

Here $w^0 \equiv 0$. In addition $w_h \in L^\infty[0,T;L^2(\Omega)]$, and hence $w_h \in \mathcal{U}_h$.

Proof. We treat the case $3/2 < p < 3$. The case $1 \le p \le 3/2$ can be treated similarly and more easily. Let $g_h, u_h \in L^2[0,T;L^2(\Gamma)]$ be given and $s \in \mathbb{R}$ with $|s| < 1$. We denote by $Y_h \equiv Y_h(g_h + su_h)$ the discontinuous Galerkin solution of (3.1) with right hand side boundary function $g_h + su_h$, i.e., the solution of the following problem: For all $n = 1,...,N$ and for all $v_h \in \mathcal{P}_k[t^{n-1},t^n;U_h]$,

$$\begin{aligned}(Y^n,v^n) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(-\langle Y_h, v_{ht}\rangle + a(Y_h,v_h) + (\phi(Y_h), v_h) + \lambda\langle Y_h, v_h\rangle_\Gamma \Big) dt \qquad (4.2)\\ &= \quad (Y^{n-1}, v_+^{n-1}) + \int_{t^{n-1}}^{t^n} \Big(\langle f, v_h\rangle + \lambda \langle g_h + su_h, v_h\rangle_\Gamma \Big) dt.\end{aligned}$$

In addition, let $w_h \equiv w_h(u_h)$ denote the solution of (4.1), i.e., for all $n = 1, ..., N$ and for all $v_h \in \mathcal{P}_k[t^{n-1}, t^n; U_h]$,

$$\begin{aligned}(w^n, v^n) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(- \langle w_h, v_{ht}\rangle + a(w_h, v_h) + (\phi'(y_h)w_h, v_h) + \lambda\langle w_h, v_h\rangle_\Gamma \Big) dt \quad (4.3)\\ &= \quad (w^{n-1}, v_+^{n-1}) + \int_{t^{n-1}}^{t^n} \lambda\langle u_h, v_h\rangle_\Gamma dt.\end{aligned}$$

Note that using the stability estimates of Lemma 3.9 and we may easily show that $Y_h, w_h \in L^2[0,T;H^1(\Omega)] \cap L^\infty[0,T;L^2(\Omega)]$. Indeed, the estimate on Y_h is evident from Lemma 3.9 since $g_h + su_h \in L^2[0,T;L^2(\Gamma)]$ and the corresponding stability constants are independent of τ, h. Note also that the estimates are also independent of s, for any $|s| < 1$. For w_h we may work similarly. The estimate at the partition points and at the energy norm is evident. For the estimate at arbitrary time-points, we may work similarly to Lemma 3.9 using the exponential interpolant $\bar{w}_h$ of w_h (see Definition 3.4). Working identically to Lemma 3.9, we obtain,

$$\begin{aligned}&(1/2)\|w^n\|^2_{L^2(\Omega)} e^{-\rho(t^n - t^{n-1})} + (1/2)\|[w^{n-1}]\|^2_{L^2(\Omega)} - (1/2)\|w^{n-1}\|^2_{L^2(\Omega)}\\ &+(\rho/2)\int_{t^{n-1}}^{t^n} \|w_h\|^2_{L^2(\Omega)} e^{-\rho(t-t^{n-1})} dt + \int_{t^{n-1}}^{t^n} \langle \phi'(y_h)w_h, \bar{w}_h\rangle dt\\ &\leq \int_{t^{n-1}}^{t^n} \Big(|a(w_h, \bar{w}_h)| + \lambda|\langle w_h, \bar{w}_h\rangle_\Gamma| + \lambda|\langle u_h, \bar{w}_h\rangle_\Gamma| \Big) dt.\end{aligned}$$

For the nonlinear term, adding and subtracting w_h and using the monotonicity of ϕ, we deduce,

$$\int_{t^{n-1}}^{t^n} (\phi'(y_h)w_h, \bar{w}_h) dt \geq \int_{t^{n-1}}^{t^n} (\phi'(y_h)w_h, \bar{w}_h - w_h) dt.$$

For the later term, the growth condition on ϕ', Hölder's inequality (with $s_1 = 2/(p-1), s_2 = s_3 = 4/(3-p)$), and the continuous embedding $H^1(\Omega) \subset L^{4/(3-p)}(\Omega)$, imply

$$\begin{aligned}&\Big|\int_{t^{n-1}}^{t^n} (\phi'(y_h)w_h, \bar{w}_h - w_h) dt\Big| \leq C \int_{t^{n-1}}^{t^n} \|y_h\|^{p-1}_{L^2(\Omega)} \|w_h\|_{L^{4/(3-p)}(\Omega)} \|\bar{w}_h - w_h\|_{L^{4/(3-p)}(\Omega)} dt\\ &\leq C\|y_h\|^{p-1}_{L^\infty[t^{n-1},t^n;L^2(\Omega)]} \int_{t^{n-1}}^{t^n} \|w_h\|_{H^1(\Omega)} \|\bar{w}_h - w_h\|_{H^1(\Omega)} dt\\ &\leq CD_{yst}^{(p-1)/2} \|w_h\|_{L^2[t^{n-1},t^n;H^1(\Omega)]} \|\bar{w}_h - w_h\|_{L^2[t^{n-1},t^n;H^1(\Omega)]}\\ &\leq CD_{yst}^{(p-1)/2} \rho\tau_n \|w_h\|^2_{L^2[t^{n-1},t^n;H^1(\Omega)]},\end{aligned}$$

where we have used the stability estimate on $\|y_h\|^2_{L^\infty[0,T;L^2(\Omega)]} \leq D_{yst}$, and Lemma 3.5 to bound $\bar{w}_h - w_h$ in terms of w_h. Thus, the estimate on $\|w_h\|_{L^\infty[0,T;L^2(\Omega)]}$ independent of τ, h and s, follows as in Lemma 3.9.

To complete the proof of the Lemma, we need to prove the following result:

$$\lim_{s\to 0} \Big(\frac{\|Y_h - y_h - sw_h\|_{L^2[0,T;H^1(\Omega)]}}{|s|} \Big) = 0. \quad (4.4)$$

Set $\tilde{y}_h = Y_h - y_h - sw_h$ and note that $\tilde{y}_h^0 \equiv 0$. In addition, using (3.1)-(4.1)-(4.2), we have that $\tilde{y}_h$ satisfies the following equation: For all $n = 1, ..., N$ and for all $v_h \in \mathcal{P}_k[t^{n-1}, t^n; U_h]$,

$$\begin{aligned}(\tilde{y}^n, v^n) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(- \langle \tilde{y}_h, v_{ht} \rangle + a(\tilde{y}_h, v_h) \Big) dt \\ &+ \quad \int_{t^{n-1}}^{t^n} \Big((\phi(Y_h) - \phi(y_h) - s\phi'(y_h)w_h, v_h) + \lambda \langle \tilde{y}_h, v_h \rangle_\Gamma \Big) dt \\ &= \quad (\tilde{y}^{n-1}, v_+^{n-1})\end{aligned} \tag{4.5}$$

Setting $v_h = \tilde{y}_h$ into (4.6) and using the standard algebra we deduce,

$$\begin{aligned}&(1/2)\|\tilde{y}^n\|^2_{L^2(\Omega)} + (1/2)\|[\tilde{y}^{n-1}]\|^2_{L^2(\Omega)} - (1/2)\|\tilde{y}^{n-1}\|^2_{L^2(\Omega)} \\ &+ \int_{t^{n-1}}^{t^n} \Big((C_F \min\{\eta, \lambda\}/2)\|\tilde{y}_h\|^2_{H^1(\Omega)} + (\lambda/2)\|\tilde{y}_h\|^2_{L^2(\Gamma)} \Big) dt + \int_{t^{n-1}}^{t^n} (\phi(Y_h) - \phi(y_h), \tilde{y}_h) dt \\ &\leq C \int_{t^{n-1}}^{t^n} \big| (s\phi'(y_h)w_h, \tilde{y}_h) \big| dt.\end{aligned} \tag{4.6}$$

where C is an algebraic constant. It remains to treat the semi-linear terms. For the last term on the right hand side of (4.6), we use the growth condition on ϕ', the generalized Hölder inequality, the interpolation inequality $\|.\|^2_{L^4(\Omega)} \leq C\|.\|_{L^2(\Omega)}\|.\|_{H^1(\Omega)}$, the continuous embedding $H^1(\Omega) \subset L^4(\Omega)$, and Young's inequality, to obtain

$$\begin{aligned}&\int_{t^{n-1}}^{t^n} \big| (s\phi'(y_h)w_h, \tilde{y}_h) \big| dt \leq C|s| \int_{t^{n-1}}^{t^n} \int_\Omega |y_h|^{p-1}|w_h||\tilde{y}_h| dx dt \\ &\leq C|s| \int_{t^{n-1}}^{t^n} \||y_h|^{p-1}\|_{L^2(\Omega)} \|w_h\|_{L^4(\Omega)} \|\tilde{y}_h\|_{L^4(\Omega)} dt \\ &\leq C|s|^2 \int_{t^{n-1}}^{t^n} \||y_h|^{p-1}\|^2_{L^2(\Omega)} \|w_h\|_{L^2(\Omega)} \|w_h\|_{H^1(\Omega)} dt + \int_{t^{n-1}}^{t^n} (C_F \min\{\eta, \lambda\}/4)\|\tilde{y}_h\|^2_{H^1(\Omega)} dt.\end{aligned}$$

The last term can be hidden on the left hand-side of (4.6), while the stability estimate on $w_h, y_h \in L^\infty[0, T; L^2(\Omega)] \cap L^2[0, T; H^1(\Omega)]$, the embedding $L^4(\Omega) \subset L^{(2p-2)}(\Omega)$, the interpolation inequality $\|.\|^2_{L^4(\Omega)} \leq C\|.\|_{L^2(\Omega)}\|.\|_{H^1(\Omega)}$, and Young's inequality, imply that

$$\begin{aligned}&\int_{t^{n-1}}^{t^n} \||y_h|^{p-1}\|^2_{L^2(\Omega)} \|w_h\|_{L^2(\Omega)} \|w_h\|_{H^1(\Omega)} dt \leq C(D_{wst}) \int_{t^{n-1}}^{t^n} \|y_h\|^{p-1}_{L^{2(p-1)}(\Omega)} \|w_h\|_{H^1(\Omega)} dt \\ &\leq C(D_{wst}) \int_{t^{n-1}}^{t^n} \|y_h\|^{p-1}_{L^4(\Omega)} \|w_h\|_{H^1(\Omega)} dt \leq C(D_{wst}) \int_{t^{n-1}}^{t^n} \Big(\|w_h\|^2_{H^1(\Omega)} + \|y_h\|^{p-1}_{L^2(\Omega)} \|y_h\|^{p-1}_{H^1(\Omega)} \Big) dt \\ &\leq C(D_{yst}, D_{wst}) \int_{t^{n-1}}^{t^n} \Big(\|w_h\|^2_{H^1(\Omega)} + \|y_h\|^2_{H^1(\Omega)} \Big) dt.\end{aligned}$$

Here the constant C depends on the stability constants of y_h, w_h. For the remaining semi-linear term of equation (4.6), we observe that the monotonicity of ϕ implies that,

$$\int_{t^{n-1}}^{t^n} (\phi(Y_h) - \phi(y_h), Y_h - y_h - sw_h) dt \geq \int_{t^{n-1}}^{t^n} (\phi(Y_h) - \phi(y_h), -sw_h) dt.$$

Using the mean value theorem and the growth condition on ϕ' we may bound the last term as follows:

$$\Big|\int_{t^{n-1}}^{t^n}(\phi(Y_h)-\phi(y_h),-sw_h)dt\Big| \le C\int_{t^{n-1}}^{t^n}\int_\Omega |s||w_h||Y_h-y_h|\big(|Y_h|^{p-1}+|y_h|^{p-1}\big)dxdt$$
$$= C\int_{t^{n-1}}^{t^n}\int_\Omega |s||w_h||\tilde{y}_h+sw_h|\big(|Y_h|^{p-1}+|y_h|^{p-1}\big)dxdt.$$

Here we have also used the definition of $\tilde{y}_h = Y_h - y_h - sw_h$. It remains to bound the resulting four integrals. Using similar considerations as above, we deduce that

$$\int_{t^{n-1}}^{t^n}\int_\Omega |s||w_h||\tilde{y}_h||Y_h|^{p-1}dxdt \le (C_F\min\{\eta,\lambda\}/8)\int_{t^{n-1}}^{t^n}\|\tilde{y}_h\|^2_{H^1(\Omega)}dt$$
$$+C(D_{Yst},D_{wst})|s|^2\int_{t^{n-1}}^{t^n}\Big(\|w_h\|^2_{H^1(\Omega)}+\|Y_h\|^2_{H^1(\Omega)}\Big)dt$$

where the constant C now depends on the stability constants of Y_h, w_h. The integral $\int_{t^{n-1}}^{t^n}\int_\Omega |s||w_h||\tilde{w}_h||y_h|^{p-1}dxdt$ can be bounded using exactly similar arguments. Finally, the integral $\int_{t^{n-1}}^{t^n}\int_\Omega |s|^2|w_h|^2\big(|Y_h|^{p-1}+|y_h|^{p-1}\big)dxdt$ can be bounded similarly and more easily. Therefore, collecting the above inequalities into (4.6) we arrive to,

$$(1/2)\|\tilde{y}^n\|^2_{L^2(\Omega)} + (1/2)\|[\tilde{y}^{n-1}]\|^2_{L^2(\Omega)} - (1/2)\|\tilde{y}^{n-1}\|^2_{L^2(\Omega)}$$
$$+\int_{t^{n-1}}^{t^n}\Big((C_F\min\{\eta,\lambda\}/16)\|\tilde{y}_h\|^2_{H^1(\Omega)}+(\lambda/2)\|\tilde{y}_h\|^2_{L^2(\Gamma)}\Big)dt$$
$$\le \tilde{C}|s|^2\int_{t^{n-1}}^{t^n}\Big(\|y_h\|^2_{H^1(\Omega)}+\|Y_h\|^2_{H^1(\Omega)}+\|w_h\|^2_{H^1(\Omega)}\Big)dt.$$

Here $\tilde{C}$ denotes a constant that depends on the stability constants of Lemma 3.9, as well as the corresponding stability constants of Y_h, w_h. The conclusion follows by summing the inequalities after noting that $\tilde{y}^0 = 0$. □

Using the Gâteaux differentiability property of Lemma 4.4 we may derive the discrete first order necessary condition.

Lemma 4.2. *Suppose that $f \in L^2[0,T;H^1(\Omega)^*]$, $y(0) \in L^2(\Omega)$, $y_h \in \mathcal{U}_h$, and $u_h \in L^2[0,T;G_h]$ and let $w_h \equiv w_h(u_h)$ be defined by (4.1). Suppose also that the assumptions of Lemma 3.9 are also satisfied. Then for every $p_h \in L^2[0,T;H^1(\Omega)]$ we have*

$$\int_0^T(\nabla p_h,\nabla w_h)dt = \int_0^T \lambda\langle u_h,\mu_h\rangle_\Gamma dt, \tag{4.7}$$

where μ_h is the solution of the following problem: For all $n = 1,...,N$ and for all $v_h \in \mathcal{P}_k[t^{n-1},t^n;U_h]$,

$$\begin{aligned} -(\mu^n_+,v^n) \quad &+ \int_{t^{n-1}}^{t^n}\Big(\langle\mu_h,v_{ht}\rangle + a(v_h,\mu_h)+\langle v_h,\phi'(y_h)\mu_h\rangle+\lambda\langle v_h,\mu_h\rangle_\Gamma\Big)dt \\ &= \; -(\mu^{n-1}_+,v^{n-1}_+)+\int_{t^{n-1}}^{t^n}(\nabla p_h,\nabla v_h)dt. \end{aligned} \tag{4.8}$$

and $\mu_h(T,x)=0$. In addition, $\mu_h \in L^\infty[0,T;L^2(\Omega)]$.

Proof. First note that for $p_h \in L^2[0,T;H^1(\Omega)]$, and $\mu_h(T,x) = 0$ the solution μ_h of (4.8) exists by adjusting the techniques of Lemmas 3.9 and 4.1, and hence $\mu_h \in \mathcal{U}_h$. Setting $v_h = \mu_h$ into (4.1) and integrating by parts with respect to time, we obtain,

$$\begin{aligned}(w_+^{n-1}, \mu_+^{n-1}) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(\langle w_{ht}, \mu_h \rangle + a(w_h, \mu_h) + (\phi'(y_h) w_h, \mu_h) + \lambda \langle w_h, \mu_h \rangle_\Gamma \Big) dt \\ &= \quad (w^{n-1}, \mu_+^{n-1}) + \int_{t^{n-1}}^{t^n} \lambda \langle u_h, \mu_h \rangle_\Gamma dt.\end{aligned} \tag{4.9}$$

Setting now $v_h = w_h$ into (4.8), we obtain,

$$\begin{aligned}-(\mu_+^n, w^n) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(\langle \mu_h, w_{ht} \rangle + a(w_h, \mu_h) + \langle w_h, \phi'(y_h) \mu_h \rangle + \lambda \langle w_h, \mu_h \rangle_\Gamma \Big) dt \\ &= \quad -(\mu_+^{n-1}, w_+^{n-1}) + \int_{t^{n-1}}^{t^n} (\nabla p_h, \nabla w_h) dt.\end{aligned} \tag{4.10}$$

Substituting (4.9) into (4.10), and using standard algebra,

$$\int_{t^{n-1}}^{t^n} (\nabla p_h, \nabla w_h) dt = -(\mu_+^n, w^n) + (w^{n-1}, \mu_+^{n-1}) + \int_{t^{n-1}}^{t^n} \lambda \langle u_h, \mu_h \rangle_\Gamma dt$$

Summing the above equalities and noting that $\mu_+^N = 0$, $w^0 = 0$ we obtain (4.7). $\square$

Now, recall that if (y_h, g_h) is the discrete optimal solution then the Gâteaux derivative vanishes. The following Lemma gives an explicit formula of the first order necessary conditions.

Lemma 4.3. *Let the assumptions of Lemma 4.2 hold. Then, there exists function $\mu_h \in \mathcal{U}_h$ such that, for all $n = 1, ..., N$ and for all $v_h \in \mathcal{P}_k[t^{n-1}, t^n; U_h]$,*

$$\begin{aligned}-(\mu_+^n, v^n) \quad &+ \quad \int_{t^{n-1}}^{t^n} \Big(\langle \mu_h, v_{ht} \rangle + a(v_h, \mu_h) + \langle v_h, \phi'(y_h) \mu_h \rangle + \lambda \langle v_h, \mu_h \rangle_\Gamma \Big) dt \\ &= \quad -(\mu_+^{n-1}, v_+^{n-1}) + \int_{t^{n-1}}^{t^n} (\nabla y_h, \nabla v_h) dt.\end{aligned} \tag{4.11}$$

where $\mu_h(T,x) = 0$, and for all $u_h \in L^2[0,T;G_h]$,

$$\int_0^T \langle \lambda \mu_h + \alpha g_h, u_h \rangle_\Gamma dt = 0. \tag{4.12}$$

Proof. We compute the Gâteaux derivative of the functional $J(y_h, g_h)$ in the arbitrary direction $u_h \in L^2[0,T;G_h]$, i.e.,

$$\begin{aligned}\frac{DJ(y_h, g_h)}{Dg_h} \cdot u_h \quad &= \quad \int_0^T \langle \nabla y_h, \nabla w_h \rangle dt + \alpha \int_0^T \langle g_h, u_h \rangle_\Gamma dt, \\ &= \quad \int_0^T \langle \lambda \mu_h + \alpha g_h, u_h \rangle dt \equiv 0.\end{aligned}$$

Here we have used Lemma 4.2, to replace the first integral. $\square$

Therefore, we have proven that the discrete optimality system consisting from the state and adjoint equations (3.1)- (4.11) respectively and the optimality condition (4.12).

We close this section by stating an a-priori estimate at arbitrary time-points for the solution of (4.11) when the right hand side belongs only in $L^2[0,T;H^1(\Omega)^*]$. The proof follows the arguments of Lemmas 3.9, and 4.1 suitably modified to handle the backwards in time pde (see also [8]).

Lemma 4.4. *Suppose that $y_0 \in L^2(\Omega)$, $f \in L^2[0,T;H^1(\Omega)^*]$ are given functions, let ϕ satisfy the growth condition assumptions 2.1. If (y_h, g_h) denote the solution of problem* (**DEBCP**) *and (y_h, μ_h, g_h) satisfy (3.1)-(4.11)-(4.12) then*

$$\|\mu_+^0\|^2_{L^2(\Omega)} + \sum_{i=1}^N \|[\mu^i]\|^2_{L^2(\Omega)} + \int_0^T C_F \min\{\eta,\lambda\}\|\mu_h\|^2_{H^1(\Omega)}dt$$
$$+ \int_0^T \Big(\eta\|\nabla\mu_h\|^2_{L^2(\Omega)} + \lambda\|\mu_h\|^2_{L^2(\Gamma)}\Big)dt \le C_{st}/\eta$$

and for $n = 1, ..., N$

$$\|\mu_+^{n-1}\|^2_{L^2(\Omega)} \le C_{st}/\eta,$$

where C_{st} is defined in Lemma 3.9. Let $\tau \equiv \max_{i=1,..,n}\tau_i$, with $\tau_i = t^i - t^{i-1}$ satisfy the assumption of Lemma 3.9. Then

$$\|\mu_h\|^2_{L^\infty[0,T;L^2(\Omega)]} \le C\Big((C_{st}/\eta) + D_{yst}^{(p-1)/2}\Big) \equiv D_{\mu st},$$

where C does not depend on α, τ, h, but only on $C_c/\eta, C_k$, Ω and D_{yst} denotes the constant of Lemma 3.9.

Theorem 4.5. *Suppose that $f \in L^2[0,T;H^1(\Omega)^*]$, and $y_0 \in L^2(\Omega)$. Given quasi-uniform partition $\{t^i\}_{i=0}^N$ of $[0,T]$, with $\tau_i = t^i - t^{i-1}$ and $\tau = \max_{i=1,...,N}\tau_i$, satisfying the assumptions of Lemma 3.9, 4.4, let $\tau, h \to 0$. In addition, suppose that $\phi \in C^2(\mathbb{R};\mathbb{R})$ with $|\phi''(s)| \le C|s|^{p-2}$ for $2 < p < 3$ or ϕ' be uniformly continuous. Then, for any $\alpha > 0$ in addition to 3.11, there exists $\mu \in L^\infty[0,T;L^2(\Omega)] \cap L^2[0,T;H^1(\Omega)]$ such that*

$$\mu_h \to \mu \quad \text{weakly in } L^2[0,T;H^1(\Omega)], \quad \mu_h \to \mu \quad \text{weakly-* in } L^\infty[0,T;L^2(\Omega)]$$

and

$$\mu_h \to \mu \quad \text{strongly in } L^2[0,T;L^2(\Omega)].$$

Furthermore, (y, g, μ) which satisfy (2.2)-(2.3)-(2.4).

Proof. (*Sketch:*) The proof follows similarly to Theorem 3.11. Recall that y_h, g_h are bounded in $L^\infty[0,T;L^2(\Omega)] \cap L^2[0,T;H^1(\Omega)]$ and $L^2[0,T;L^2(\Omega)]$ by constants independent of τ, h and converge to some elements (y, g) as stated in 3.11. Similarly μ_h is bounded in $L^\infty[0,T;L^2(\Omega)] \cap L^2[0,T;L^2(\Omega)]$ by a constant independent of τ, h. Hence, we may extract subsequence, converging to some element μ, as follows,

$$\mu_h \to \mu \quad \text{weakly in } L^2[0,T;H^1(\Omega)], \quad \mu_h \to \mu \quad \text{weakly-* in } L^\infty[0,T;L^2(\Omega)],$$

while an application of the discrete compactness Theorem 3.7 guarantees the strong convergence of μ_h to μ in $L^2[0,T;L^2(\Omega)]$. Indeed, Lemma A.1 implies that $\|\phi'(y_h)\mu_h\|_{L^{4/3}[0,T;H^1(\Omega)^*]}$. The rest of the terms can be treated easily. The proof is completed after noting that we may pass the limit into equations (4.11)-(4.12), similar to the proof of the discrete case of Theorem 3.11, with the help of Lemma A.1. $\square$

5 Conclusion

We have proved basic stability and convergence properties for a general class of discontinuous time-stepping schemes for a Robin boundary control problem for semilinear parabolic pdes, with rough initial data and forcing term. The underlying stability properties demonstrate that the discrete state, adjoint and control variables exhibit similar regularity properties to the continuous optimal control problem, which is an important asset in the analysis and implementation of numerical schemes. The emphasis was on the minimal regularity assumptions on the data, but error estimates and computational issues will be also considered in a future work.

A Bounds on semi-linear terms

Lemma A.1. *Let ϕ satisfy the assumptions 2.1, and let $y_h, \mu_h \in L^\infty[0,T;L^2(\Omega)] \cap L^2[0,T;H^1(\Omega)]$ with bounds independent of τ, h. The following statements hold:*

- $\|\phi(y_h)\|_{L^{4/3}[0,T;H^1(\Omega)^*]}$ *is bounded independent of τ, h.*
- $\|\phi'(y_h)\mu_h\|_{L^{4/3}[0,T;H^1(\Omega)^*]}$ *is bounded independent of τ, h.*
- *If $v_h \in C[0,T;U_h]$, $y, \mu \in W(0,T)$ then,*
$$\int_0^T \Big|\langle \phi(y) - \phi(y_h), v_h\rangle\Big| dt \le C_{y,y_h} \|y - y_h\|^2_{L^2[0,T;L^2(\Omega)]}.$$
- *In addition, if $\phi \in C^2(\mathbb{R};\mathbb{R})$ with $|\phi''(s)| \le C|s|^{p-2}$ for $2 < p < 3$ or ϕ' be uniformly continuous then,*
$$\int_0^T \Big|\langle \phi'(y)\mu - \phi'(y_h)\mu_h, v_h\rangle\Big| dt \le C_{y,\mu,y_h,\mu_h}\Big(\|\mu - \mu_h\|^2_{L^2[0,T;L^2(\Omega)]} + \|y - y_h\|^2_{L^2[0,T;L^2(\Omega)]}\Big).$$

Here $C_{y,y_h}, C_{y,y_h,\mu,\mu_h}$ are constants depending upon $\|y\|_{W(0,T)}, \|\mu\|_{W(0,T)}$ and the stability constants of Lemmas 3.9, 4.4.

Proof. For the first part of this Lemma, we treat the case $(3/2) \le p < 3$. The case $1 \le 1 < (3/2)$ can be treated similarly and more easily. Let $v \in L^4[0,T;H^1(\Omega)]$. Using Hölder's inequalities, the embedding $H^1(\Omega) \subset L^4(\Omega)$,

$$\begin{aligned}\int_0^T \int_\Omega |y_h|^p v dx dt &\le C \int_0^T \|y_h\|^p_{L^{4p/3}(\Omega)} \|v\|_{L^4(\Omega)} dt \\ &\le C\Big(\int_0^T \|y_h\|^{4p/3}_{L^{4p/3}(\Omega)}\Big)^{3/4}\Big(\int_0^T \|v\|^4_{H^1(\Omega)}\Big)^{1/4}.\end{aligned}$$

The Gagliardo-Nirenberg interpolation inequality with $r = 2, q = (4p/3)$ and $s = 1 - \frac{2}{(4p/3)} \equiv \frac{2p-3}{2p}$, $1 - s = 3/2p$, implies (note that $(4p/3) \geq 2$ for $p \leq (3/2)$)

$$\|y_h\|_{L^{4p/3}(\Omega)} \leq C\|\hat{y}_h\|_{L^2(\Omega)}^{1-s}\|\hat{y}_h\|_{H^1(\Omega)}^{s}$$

and hence,

$$\int_0^T \int_\Omega |y_h|^p v dx dt \quad \leq \quad C\Big(\int_0^T \|y_h\|_{L^2(\Omega)}^2 \|y_h\|_{H^1(\Omega)}^{\frac{2}{3}\times(2p-3)}\Big)^{3/4} \|v\|_{L^4[0,T;H^1(\Omega)]} \leq C$$

where at the last step, we have used the stability bounds for y_h and $\frac{2}{3} \times (2p-3) < 2$, for $1 < p < 3$. For the second statement, we obtain

$$\begin{aligned}
&\int_0^T \langle \phi'(y_h)\mu_h, v_h\rangle dt \leq C\|\mu_h\|_{L^4[0,T;L^4(\Omega)]}\||y_h|^{p-1}\|_{L^2[0,T;L^2(\Omega)]}\|v_h\|_{L^4[0,T;L^4(\Omega)]} \\
&\leq C\|\mu_h\|_{L^\infty[0,T;L^2(\Omega)]}^{1/2}\|\mu_h\|_{L^2[0,T;H^1(\Omega)]}^{1/2}\||y_h|^{p-1}\|_{L^2[0,T;L^2(\Omega)]}\|v_h\|_{L^4[0,T;H^1(\Omega)]}.
\end{aligned}$$

Note that using the embedding $L^4(\Omega) \subset L^{2(p-1)}(\Omega)$ (recall that $3/2 \leq p < 3$), and the interpolation inequality $\|.\|_{L^4(\Omega)}^2 \leq C\|.\|_{L^2(\Omega)}\|.\|_{H^1(\Omega)}$, we may show that $\||y_h|^{p-1}\|_{L^2[0,T;L^2(\Omega)]} < \infty$ since y_h remains bounded (with constant independent of h, τ) in $L^\infty[0,T;L^2(\Omega)] \cap L^2[0,T;H^1(\Omega)]$. The second part, recall that by the mean value Theorem, the growth condition on ϕ', and the generalized Hölder's inequality with $(1/2) + (1/s_1) + (1/s_2) = 1$, we obtain that,

$$\begin{aligned}
\int_0^T \Big|\langle \phi(y) - \phi(y_h), v_h\rangle\Big| dt \quad &\leq \quad C\int_0^T \|y_h - y\|_{L^2(\Omega)}\big(\||y_h|^{p-1}\|_{L^{s_1}(\Omega)} + \||y|^{p-1}\|_{L^{s_1}(\Omega)}\big)\|v_h\|_{L^{s_2}(\Omega)} dt \\
&\equiv \quad \mathbf{I}_1 + \mathbf{I}_2.
\end{aligned}$$

We treat the first integral. Choosing $s_1 = 2 + \epsilon$, with $\epsilon > 0$ small enough, and using the continuous embedding of $H^1(\Omega) \subset L^{s_2}(\Omega)$, we deduce,

$$\begin{aligned}
\mathbf{I}_1 \quad &\equiv \quad \int_0^T \|y_h - y\|_{L^2(\Omega)}\||y_h|^{p-1}\|_{L^{2+\epsilon}(\Omega)}\|v_h\|_{H^1(\Omega)} dt \\
&\leq \quad C\|v_h\|_{C[0,T;H^1(\Omega)]}\|y_h - y\|_{L^2[0,T;L^2(\Omega)]}\||y_h|^{p-1}\|_{L^2[0,T;L^{2+\epsilon}(\Omega)]}.
\end{aligned}$$

It remains to prove that the last norm is bounded independent of τ, h. Note that $\||y_h|^{p-1}\|_{L^{2+\epsilon}(\Omega)} \leq C\|y_h\|_{L^{(p-1)(2+\epsilon)}(\Omega)}^{p-1}$. Hence, using the Gagliardo-Nirenberg inequality with $r = 2$, $q = (p-1)(2+\epsilon)$, and $s = 1 - 2/(p-1)(2+\epsilon)$, we obtain that

$$\int_0^T \|y_h\|_{L^{(p-1)(2+\epsilon)}(\Omega)}^{2(p-1)} dt \leq \int_0^T \|y_h\|_{L^2(\Omega)}^{4/(2+\epsilon)}\|y_h\|_{H^1(\Omega)}^{(2(p-1)(2+\epsilon)-4)/(2+\epsilon)} dt.$$

The proof is completed after noting that $\|y_h\|_{L^\infty[0,T;L^2(\Omega)]}$ is bounded independent of h, τ, and noting that as $\epsilon \to 0$, $(2(p-1)(2+\epsilon) - 4)/(2+\epsilon) \to 2p - 4 < 2$ for $p < 3$. The last statement can be proved similarly, after noting that

$$\int_0^T \langle \phi'(y)\mu - \phi'(y_h)\mu_h, v_h\rangle dt = \int_0^T \langle(\phi'(y) - \phi'(y_h))\mu, v_h\rangle dt + \int_0^T \langle \phi'(y_h)(\mu - \mu_h), v_h\rangle dt.$$

$\square$

References

[1] G. Akrivis and C. Makridakis, Galerkin time-stepping methods for nonlinear parabolic equations, *ESAIM: Math. Model. and Numer. Anal.*,**38** (2004), pp. 261-289.

[2] H. Brezis, Analyse fonctionnelle. Theorie et applications (French), *Masson*, Paris, 1983.

[3] E. Casas and M. Mateos, Error estimates for the numerical approximation of Neumann control problems, *Comput. Optim. and Appl.*, **39** (2008), pp. 265-295.

[4] E. Casas, M. Mateos, and J.-P. Raymond, Penalization of Dirichlet boundary control problems, *ESAIM Cont. Optim. and Calc. Var.*, **15** (2009), pp 782-809.

[5] E. Casas, M. Mateos and F. Tröltzsch, Error estimates for the numerical approximation of boundary semilinear elliptic control problem, *Comput. Optim. and Appl.*, **31** (2005), pp. 193-219.

[6] E. Casas, and J.-P. Raymond, Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equation, *SIAM J. Control and Optim.*, **45** (No 5) (2006), pp. 1586-1611.

[7] K. Chrysafinos, Analysis and finite element approximations for distributed optimal control problems for implicit parabolic PDE's, *J. Comput. Appl. Math.*, **231** (2009), pp. 327-348.

[8] K. Chrysafinos, Convergence of discontinous Galerkin approximations for an optimal control problem associated to semilinear parabolic PDE's., *ESAIM M2AN.*, **44** (2010), pp 189-206.

[9] K. Chrysafinos, and E. Karatzas, Symmetric error estimates for discontinuous Galerkin approximations for an optimal control problem associated to semilinear parabolic pdes, *Discr. Cont. Dynam. Sys., Ser B.*, **17** (2012), pp 1473-1506.

[10] K. Chrysafinos and N.J. Walkington, Error estimates for the discontinuous Galerkin methods for parabolic equations, *SIAM J. Numer. Anal.*, **44** (No 1) (2006), pp 349-366.

[11] K. Chrysafinos and N.J. Walkington, Discontinuous Galerkin approximations for the Stokes and Navier-Stokes equations, *Math. Comp.*, **272** (2010), pp. 2135-2167.

[12] K. Chrysafinos, M.D. Gunzburger and L.S. Hou, Semidiscrete approximations of optimal Robin boundary control problems constrained by semilinear parabolic PDE, *J. Math. Anal. Appl.*, **323** (2006), pp 891-912.

[13] P.G. Ciarlet, The finite element method for elliptic problems, *SIAM Classics in Applied Math*, 2002.

[14] J.-C. de los Reyes, and K. Kunisch, A semi-smooth Newton method for control constrained boundary control of the Navie-Stokes equtions, *Nonlinear Anal.*, **62** (2005), pp. 1289-1316.

[15] D. Estep and S. Larsson, The discontinuous Galerkin method for semilinear parabolic equations, *RAIRO Modél. Math. Anal. Numér.*, **27** (1993), pp 35-54.

[16] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. I. A linear model problem, *SIAM J. Numer. Anal.*, **28** (1991), pp. 43-77.

[17] K. Ericksson and C. Johnson, Adaptive finite element methods for parabolic problems IV: Nonlinear problems, *SIAM J. Numer. Anal.*, **32** (6) (1995), pp. 1729-1749.

[18] K. Eriksson, C. Johnson and V. Thomée, Time discretization of parabolic problems by the discontinuous Galerkin method, *RAIRO Modél. Math. Anal. Numér.*, **29** (1985), pp 611-643.

[19] L. Evans, Partial Differential Equations, AMS, Providence RI, 1998.

[20] A. Fursikov, Optimal control of distributed systems. Theory and applications, AMS, Providence, 2000.

[21] V. Girault and P-A. Raviart, Finite Element Methods for Navier-Stokes, Springer-Verlag, New York, 1986.

[22] W. Gong and N. Yan, A posteriori error estimates for boundary control problems governed by parabolic partial differential equations, *J. Comp. Math.*, **27** (2009), pp. 68-88.

[23] R. Griesse and S. Volkwein, A primal-dual active set strategy for optimal boundary control of a nonlinear reaction-diffusion system, *SIAM J. on Control and Optim.*, **44** (2) (2005), p.467-494.

[24] M. D. Gunzburger Perspectives in flow control and optimization, SIAM, Advances in Design and Control, Philadelphia, 2003.

[25] M.D. Gunzburger and L.S. Hou, Treating inhomogeneous essential boundary conditions in finite element methods and the calculation of boundary stresses, *SIAM J. Numer. Anal.*, **29** (2), pp. 390-424.

[26] M. D. Gunzburger, L. S. Hou and T. Svobodny, Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with Dirichlet controls, *RAIRO Model. Math. Anal. Numer.*, **25** (1991), pp. 711-748.

[27] M. D. Gunzburger and S. Manservisi, The velocity tracking problem for Navier-Stokes flows with boundary control, *SIAM J. Control and Optim.*, **39** (2000), pp. 594-634.

[28] M. Hinze, A variational discretization concept in control onstrained optimization: The linear-quadratic case, *Comput. Optim. Appl.*, **30** (2005), pp. 45-61.

[29] M. Hinze and K. Kunisch, Second order methods for boundary control of the instationary Navier-Stokes system, *ZAMM Z. Angew. Math. Mech.*, **84** (2004), pp 171-187.

[30] L.S. Hou and S. Ravindran, A penalized Neumann control approach for solving an optimal Dirichlet control problem for the Navier-Stokes equations, *SIAM J. Control Optim.*, **36** (1998), pp. 1795-1814.

[31] G. Knowles, Finite element approximation of parabolic time optimal control problems, *SIAM J.Control and Optim.*, **20** (1982), pp. 414-427.

[32] K. Kunisch and B. Vexler, Constrained Dirichlet boundary control in L^2 for a class of evolution equations, *SIAM J. on Control and Optim.*, **46** (5) (2007), pp. 1726 - 1753.

[33] I. Lasiecka, Rietz-Galerkin approximation of the time optimal boundary control problem for parabolic systems with Dirichlet boundary conditions, *SIAM J. Control and Optim.*, **22** (1984), pp. 477-500.

[34] I. Lasiecka and R. Triggiani, Control theory for partial differential equations, Cambridge University press, 2000.

[35] W.-B. Liu and N. Yan, A posteriori error estimates for optimal control problems governed by parabolic equations, *Numer. Math.*, **93** (2003), pp.497-521.

[36] W.-B.Liu, H.-P. Ma, T. Tang and N. Yan, A posteriori error estimates for DG time-stepping method for optimal control problems governed by parabolic equations, *SIAM J. Numer. Anal.*, **42** no 3 (2004), pp. 1032-1061.

[37] K. Malanowski, Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal-control problems, *Appl. Math. Optim.*, **8** (1981), pp. 69-95.

[38] D. Meidner and B. Vexler, Adaptive space-time finite element methods for parabolic optimization problems, *SIAM J. Control and Optim.*, **46** (2007), pp. 116-142.

[39] D. Meidner and B. Vexler, A priori error estimates for space-time finite element discretization of parabolic optimal control problems. Part I: Problems without control constraints, *SIAM J. on Control. and Optim.*, **47** (3) (2008), pp. 1150 - 1177.

[40] D. Meidner and B. Vexler, A priori error estimates for space-time finite element discretization of parabolic optimal control problems. Part II: Problem with control constraints, *SIAM J. on Control and Optim.*, **47** (3) (2008), pp. 1301 - 1329.

[41] P. Neittaanmaki and D. Tiba, Optimal control of nonlinear parabolic systems. Theory, algorithms and applications. M. Dekker, New York, 1994.

[42] I. Neitzel and B. Vexler, A priori error estimates for space-time finite element discretization of semilinear parabolic optimal control problems, *Numer. Math.*, **120** (2011), 345-386.

[43] A. Rösch and F. Tröltzsch, Sufficient second-order optimality conditions for a parabolic optimal control problem with pointwise control-state constraints, *SIAM J. Control Optim.*, **42** (1) (2003), pp. 138-154.

[44] R. Temam, Navier-Stokes equations, North Holand, Amesterdam, 1977.

[45] V. Thomée, Galerkin finite element methods for parabolic problems, Spinger-Verlag, Berlin, 1997

[46] F. Tröltzsch, Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems, *International Series of Numerical Mathematics*, **111** (1993), pp 57-68.

[47] F. Tröltzsch, Semidiscrete Ritz-Galerkin approximation of nonlinear parabolic boundary control problems- Strong convergence of optimal controls, *Appl. Math. Optim.*, **29** (1994), pp 309-329.

[48] F. Tröltzsch, Optimal control of partial differential equations: Theory, methods and applications, *Graduate Studies in Mathematics, AMS*, **112**, Providence 2010.

[49] N. J. Walkington, Compactness properties of the DG and CG time stepping schemes for parabolic equations, *SIAM J. Numer. Anal.*, **47** No 6 (2010), pp 4680-4710.

[50] R. Winther, Initial value methods for parabolic control problems, *Math. Comp.*, **34** (1980), pp. 115-125

[51] E. Zeidler, Nonlinear functional analysis and its applications, II/B Nonlinear monotone operators, Springer-Verlag, New York, 1990.