RESEARCH ARTICLE

ALL-AT-ONCE APPROACH FOR OPTIMAL CONTROL OF UNSTEADY BURGERS EQUATION

F. Yılmaz\textsuperscript{a} and B. Karasözen\textsuperscript{b}\textsuperscript{*}

\textsuperscript{a}Department of Mathematics, Gazi University, 06500 Ankara, Turkey; \textsuperscript{b}Department of Mathematics and Institute of Applied Mathematics, Middle East Technical University, 06531 Ankara, Turkey

(Received 00 Month 200x; in final form 00 Month 200x)

In this paper, we apply the all-at-once method for the optimal control of unsteady Burgers equation. The all-at-once methods were applied in recent years for optimal control problems governed by linear elliptic and parabolic equations. For space discretization, we use the Galerkin finite element method. The nonlinear Burgers equation is discretized in time using the semi-implicit discretization which results an effective linearization of the optimal control problem. An a priori error analysis is developed for the state, adjoint and control variables. This all-at-once approach leads to the solution of an indefinite saddle point problem, which is usually solved iteratively using preconditioners. Because the Burgers equation is one dimensional in space, the saddle point system can be solved efficiently by direct solvers. Numerical results for distributed unconstrained and control constrained problems illustrate the performance of all-at-once approach with semi-implicit time discretization.

Keywords: Optimal control, Burgers equation, all-at-once approach, control constraints, semi-implicit discretization.

AMS Subject Classification: 49J20, 65M15, 65M60

1. Introduction

Analysis and numerical approximation of optimal control problems (OCP) for Burgers equation are important for the development of the numerical methods for the optimal control of more complicated models in fluid dynamics like Navier-Stokes equations. The distributed and boundary optimal control problems for stationary and unsteady Burgers equation are solved using SQP (sequential quadratic programming) methods, primal-dual active set and semi-smooth Newton methods [8, 18, 25–27]. Implicit Euler and Crank-Nicholson methods were compared for solving the adjoint equations arising from optimal control of unsteady Burgers equation in [13]. Distributed optimal control problems with Burgers equations are solved by simultaneous space-time discretization using COMSOL Multiphysics in [28]. In contrast to linear parabolic control problems, the optimal control problem for the Burgers equation is a non-convex problem with multiple local minima due to nonlinearity of the differential equation. Numerical methods can only compute minima close to the starting points [25].

There are two different approaches for the discretization of the OCPs: \textit{optimize then-discretize} and \textit{discretize-then-optimize}. In the \textit{optimize-then-discretize approach}, first the necessary optimality conditions are established on the continuous

\textsuperscript{*}Corresponding author. Email: bulent@metu.edu.tr

ISSN: 0233-1934 print/ISSN 1029-4945 online
© 2011 Taylor & Francis
DOI: 10.1080/0233193YYxxxxxxx
http://www.informaworld.com
level consisting of the state, adjoint and the optimality equations, and then
these equations are discretized usually by finite elements. The optimality system
consists of coupled state and adjoint equations and gradient equations arising
from the control constraints. Usually this system is integrated iteratively forward
and backward in time by gradient based algorithms. In the \textit{discretize-then-optimize}
approach, the state equation is discretized and then the optimality system for
the finite dimensional optimization problem is derived. This approach is referred
also to as the black-box approach. In other words, an existing algorithm for
the solution of the state equation is embedded into an optimization loop. The
black-box approach is easy to use because it requires no modification to an existing
partial differential equation (PDE) integrator. When a non-linear problem has to
be solved, the repeated costly solution of the state equation is needed.

In recent years, the so called all-at-once type methods were applied to PDE con-
strained OCP’s. Because all-at-once methods treat the control and state as inde-
pendent optimization variables, the optimization problem is explicitly constrained;
the state, control and adjoint state variables can be solved explicitly for all time-
steps at once by solving a large system of linear equations. All-at-once method
was applied to linear elliptic OPC’s with with distributed and boundary controls
in [15–17] and to parabolic OCP’s [19, 22] and recently to time dependent Stokes
equation [21].

In all-at-once approach, the discretization of the problem and solution via sta-
tionarity first order optimality conditions on the Lagrangian leads to a linear system
in saddle point form

\[
\begin{pmatrix}
A & BT \\
B & 0
\end{pmatrix}
x = b
\]

(1)

where \( A \in \mathbb{R}^{n \times n} \) is symmetric and positive definite or positive semi-definite and
\( B \in \mathbb{R}^{m \times n}, m < n \), is a matrix of full rank. The linear system given in (1) is well
defined and has a unique solution, if the block \( A \) is positive definite on the kernel
of \( B \) [1]. There exists methods to solve the saddle point problems (1) efficiently,
when the system matrix \( A \) is symmetric and indefinite (see [1] for a survey).

In this paper, we will apply the all-at-once method to the optimal control of
time-dependent Burgers equation. In the literature, the all-at-once methods are
applied to OCP’s with linear elliptic or parabolic PDE’s. Because the Burgers
equation is non-linear, an efficient linearization method is needed. Standard time
integrators for Burgers equation are the Cranck-Nicolson and backward Euler
method which are implicit and unconditionally stable and both methods require
solution of nonlinear equations at each time step. The semi-implicit method is
also unconditionally stable, but it does not require solution of nonlinear equations
[3, 14]. It is order of convergence in time is one like the backward Euler, whereas
the Cranck-Nicholson method is of second order. In semi-implicit method, the
diffusive part of Burgers equation is discretized implicitly, and the non-linear part
explicitly. This provides an effective linearization procedure; at each time step
a linear system of equation with the same symmetric matrix have to solved. In
practice, the linear system \( Ax = b \) usually is of sufficiently high dimension, when
a one-shot approach for time-dependent problems in two dimensional space is
considered that iterative solution methods with preconditioners are needed [19, 22].
Because we consider the unsteady Burgers equation in one dimension, the saddle
point system (1) can be solved efficiently with the direct sparse solvers of MATLAB.
The paper is organized as follows. Discretization of the problem by finite elements in space and the time discretization with the semi-implicit method will be presented in Section 2. Also the application of the all-at-once method for the unconstrained and control constrained problems are given in Section 2. In section 3, a priori error estimates are obtained for the state and control variables for the fully discretized problem. The numerical results presented in Section 4 confirm the orders of convergence obtained by the a priori error estimates in Section 3.

2. Discretization of the optimal control problem

The distributed control problem for the viscous Burgers equation with control constraints and with homogeneous Dirichlet boundary conditions can be stated as follows [6, 25]

\[
\min_{(y,u)} \ J(y,u) = \frac{1}{2} \int_0^T \int_0^1 \left( (y(x,t) - y_d(x,t))^2 + \alpha u^2(x,t) \right) \, dx dt, \tag{2}
\]

subject to \( y_t + y y_x - \nu y_{xx} = u \quad (x,t) \in (0,1) \times (0,T), \)
\[
y(0,t) = y(1,t) = 0 \quad t \in (0,T), \tag{3}
\]
\[
y(x,0) = y_0 \quad x \in (0,1),
\]

with bound constraints on the control \( u_a(t,x) \leq u(t,x) \leq u_b(t,x) \quad (x,t) \in (0,1) \times (0,T), \) where \( y(x,t), u(x,t) \) and \( y_d(x,t) \) denote the state, control and the desired state, respectively, \( \nu > 0 \) is the viscosity and \( \alpha > 0 \) is the regularization parameter.

In one-shot approach, first the state equation is discretized and the optimality conditions are derived by using the discrete Lagrangian. State and control control variables are discretized by using standard Galerkin method with linear finite elements on the interval \((0,1)\) with \(n\) uniform subdivisions with the step size \(h = 1/n\).

\[
y_h(x,t) = \sum_{j=1}^{n-1} y_j(t) \phi_j(x), \quad \text{and} \quad u_h(x,t) = \sum_{j=1}^{n-1} u_j(t) \phi_j(x). \tag{4}
\]

With the test functions \( \phi_i, i = 0, \cdots, n \) vanishing on the boundaries, the weak form of the Burgers equation becomes

\[
\frac{d}{dt} \int_0^1 \left( \sum_{j=1}^{n-1} y_j \phi_j \right) \phi_i dx + \nu \int_0^1 \frac{d}{dx} \left( \sum_{j=1}^{n-1} y_j \phi_j \right) \frac{d}{dx} \phi_i dx + \int_0^1 \frac{d}{dx} \left( \sum_{k=1}^{n-1} y_k \phi_k \right) \left( \sum_{j=1}^{n-1} y_j \phi_j \right) \phi_i dx \\
= \int_0^1 \left( \sum_{j=1}^{n-1} u_j \phi_j \right) \phi_i dx \quad i = 1, \cdots, n - 1. \tag{5}
\]

Setting \( \bar{y}(t) = (y_1(t), \cdots, y_{n-1}(t))^T \) and \( \bar{u}(t) = (u_1(t), \cdots, u_{n-1}(t))^T \) and inserting (4) in (5) we obtain the semi-discrete system

\[
M_h \frac{\bar{y}(t)}{dt} + \nu A_h \bar{y}(t) + q_h(\bar{y}(t)) = M_h \bar{u}(t) \quad t \in (0,T), \tag{6}
\]

where \( M_h \in \mathbb{R}^{(n-1) \times (n-1)} \) and \( A_h \in \mathbb{R}^{(n-1) \times (n-1)} \) are the mass and stiffness matri-
ces, respectively. And \( q_h(y(t)) \) is the nonlinear term of Burgers equation

\[
q_h(y(t)) := \frac{1}{6} \begin{pmatrix}
y_1(t) y_2(t) + y_3^2(t) \\
y_{i-1}(t) (y_{i-1}(t) + y_i(t)) + y_{i+1}(t) (y_i(t) + y_{i+1}(t)) \\
\vdots \\
-\frac{y_{n-2}(t)^2}{6} - y_n - 2 y_{n-1}(t)
\end{pmatrix} \in \mathbb{R}^{n-1}.
\]

The cost function is also discretized by linear finite elements

\[
J_h = \int_0^T \frac{1}{2} (\vec{y} - \vec{y}_d)^T M_h (\vec{y} - \vec{y}_d) dt + \int_0^T \frac{\alpha}{2} \vec{u}^T M_h \vec{u} dt.
\]

The integral (7) is approximated by trapezoidal rule and the state equation is discretized in time semi-implicit scheme ([14], pp. 411). The fully discretized optimal control problem is given by

\[
\min_{\vec{u}_1, \cdots, \vec{u}_N} \Delta t \sum_{i=1}^N \left( \frac{1}{2} (\vec{y}_i - \vec{y}_{d_i})^T M_h (\vec{y}_i - \vec{y}_{d_i}) + \frac{\alpha}{2} \vec{u}_i^T M_h \vec{u}_i \right)
\]

where \( \vec{y}_1, \cdots, \vec{y}_N \) is solution of the

\[
(M_h + \Delta t \nu A_h) \vec{y}_{i+1} - M_h \vec{y}_i + \Delta t \vec{q}_h(\vec{y}_i) = \Delta t M_h \vec{u}_{i+1} \quad i = 1, \cdots, N - 1, \\
\vec{y}(0) = \vec{y}_0,
\]

with uniform time steps \( \Delta t \).

Semi-implicit time approximation consists in evaluating the diffusive part \( y_{xx} \) at the time level \( i + 1 \), whereas the remaining parts are considered at time level \( i \). When this scheme is applied to a non-linear advection, it provides an efficient linearization.

The main motivation for dealing with the diffusive part implicitly and the advection part explicitly for diffusion-convection equations is that the semi-implicit scheme is unconditionally stable and at each step a linear system with a symmetric matrix has to be solved. Because of these the semi-implicit scheme provides an effective linearization procedure for problems with non-linear advection terms like Burgers equation. Stability and convergence analysis of the semi-implicit scheme for diffusion-convection equations with finite elements were handled in [14] and for Burgers equation with the spectral elements in [3].

When the state, control and desired state vectors at time steps 1 to \( N \) are collected in the vectors

\[
y = [\vec{y}_1, \vec{y}_2, \cdots, \vec{y}_N]^T, \quad u = [\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_N]^T, \quad y_d = [\vec{y}_{d1}, \vec{y}_{d2}, \cdots, \vec{y}_{d(N)}]^T,
\]

then the cost function (8) becomes

\[
J(y, u) = \frac{\Delta t}{2} (y - y_d)^T M_{1/2} (y - y_d)^T + \frac{\alpha \Delta t}{2} u^T M_{1/2} u,
\]
with the matrix

$$\mathcal{M}_{1/2} = \begin{pmatrix} \frac{1}{2} M_h & M_h & \cdots & \frac{1}{2} M_h \\ M_h & -\frac{1}{2} M_h & & \\ & \ddots & \ddots & \\ & & M_h & -\frac{1}{2} M_h \end{pmatrix}.$$ 

One-shot discretization of the state equation for $N$ time steps becomes

$$\begin{bmatrix} M_h + \Delta t A_h & -M_h & & \\ -M_h & M_h + \Delta t A_h & \ddots & \\ & \ddots & \ddots & -M_h \\ & & -M_h & M_h + \Delta t A_h \end{bmatrix} \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_N \end{bmatrix} - \Delta t \mathcal{M} \mathbf{u} = - \begin{bmatrix} (M_h \bar{y}_0 + \Delta t M_h q(\bar{y}_0)) \\ \Delta t M_h q(\bar{y}_1) \\ \vdots \\ \Delta t M_h q(\bar{y}_{N-1}) \end{bmatrix},$$

where $\mathcal{M} = \text{blockdiag}\{M_h, \cdots, M_h\}$.

The optimality system containing first order optimality conditions is obtained by introducing the Lagrangian [24] as

$$L(\mathbf{y}, \mathbf{u}, \mathbf{p}) := \frac{\Delta t}{2} (\mathbf{y} - \mathbf{y}_d)^T \mathcal{M}_{1/2} (\mathbf{y} - \mathbf{y}_d) + \alpha \frac{\Delta t}{2} \mathbf{u}^T \mathcal{M}_{1/2} \mathbf{u} + \mathbf{p}^T (-K \mathbf{y} + \Delta t \mathbf{M} \mathbf{u} + d),$$

with the Lagrange multiplier $\mathbf{p} = [\vec{p}_1, \vec{p}_1, \cdots, \vec{p}_N]^T$.

The optimality system can be written using the stationary conditions on the Lagrangian $L$ as

$$\begin{bmatrix} \mathcal{M} & 0 & -K^T \\ 0 & \alpha \Delta t \mathcal{M}_{1/2} & \Delta t \mathcal{M} \\ -K & \Delta t \mathcal{M} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathcal{M} \mathbf{y}_d \\ 0 \\ d \end{bmatrix}. \quad (12)$$

The first optimality condition in (12) gives the discrete adjoint equation

$$-K^T \mathbf{p} = \Delta t \mathcal{M}_{1/2} (\mathbf{y} - \mathbf{y}_d). \quad (13)$$

The matrix formulation of discrete adjoint (13) corresponds to the semi-implicit scheme as

$$-\frac{p^{n+1} - p^n}{\Delta t} - \Delta p^n = y^n - y_d.$$

As it was remarked in [19] for the backward Euler method, this scheme is not consistent when the continuous adjoint equation is discretized by the semi-implicit scheme when optimize-then-discretize approach is used. Another difference occurs at time level $t = T$. When the trapezoidal rule is used the adjoint variable $p$ is not necessarily equal to 0 at $t = T$. This obstacle can be overcame by letting $\Delta t \to 0$. However, when rectangular rule is used, the final block of the matrix $\mathcal{M}_{1/2}$ is 0 and final time condition of discrete adjoint is satisfied. Numerically, there is not big differences when we choose sufficiently small $\Delta t$. 
Similarly, for the control constrained problem the Lagrangian [24] is given as

\[
L(y, u, p, \mu_a, \mu_b) := \frac{\Delta t}{2}(y - y_d)^T M_{1/2}(y - y_d) + \frac{\alpha \Delta t}{2} u^T M_{1/2} u \\
+ p^T (-K y + \Delta t M u + d) + \mu_a^T (u_a - u) + \mu_b^T (u_b - u),
\]

(14)

where \( \mu_a \) and \( \mu_b \) represent the Lagrange multipliers for the inequality constraints on the control variable defined as

\[
\mu_a := (\alpha \Delta t M_{1/2} u + \Delta t M p)^+ \quad \text{and} \quad \mu_b := (\alpha \Delta t M_{1/2} u + \Delta t M p)^-.
\]

The discretized form of the box constraints are given as \( u_a \leq u \leq u_b \), where \( u_a \) and \( u_b \) represent the projections on the finite element space and on the discrete time interval \( t_n = n \Delta t, n = 1, \ldots, N \), with \( \mu = \mu_a - \mu_b \).

The inequality constraints can be determined by an extension of the theorem ([24], Theorem 1.4, see also [20]) about the optimality conditions for time-dependent problems as

**Theorem 2.1:** For an optimal solution \((y^*, u^*)\), there exists Lagrange multipliers \( p, \mu_a, \) and \( \mu_b \) such that

\[
\nabla_y L(y, u^*, p^*, \mu_a, \mu_b) = 0,
\]

\[
\nabla_u L(y^*, u^*, p^*, \mu_a, \mu_b) = 0,
\]

\[
\mu_a \geq 0, \quad \mu_b \geq 0,
\]

\[
\mu_a^T (u_a - u^*) = \mu_b^T (u^* - u_b) = 0.
\]

The optimality system with the control constraints are solved usually with the active set methods. This method was introduced in [2]. For a detailed discussion of active set methods we refer [2, 24]. We define the active sets as

\[
A_+ := \{i \in \{1, \ldots, N\} : (u^* - \mu)_i > u_b\}_i,
\]

\[
A_- := \{i \in \{1, \ldots, N\} : (u^* - \mu)_i < u_a\}_i,
\]

\[
I := \{1, 2, \ldots, N\} \setminus (A_+ \cup A_-).
\]

The active set method consist of solving the optimality system iteratively within the active set

\[
\begin{pmatrix}
\mathcal{M} & 0 & -\mathcal{K}^T \\
0 & \alpha \Delta t M_{1/2} \Delta t \chi_I M & 0 \\
-\mathcal{K} & \Delta t M
\end{pmatrix}
\begin{pmatrix}
y \\
u \\
p
\end{pmatrix}
= \begin{pmatrix}
\alpha \Delta t M_{1/2} (\chi_{A_+} u_a + \chi_{A_+} u_b) \\
MY_d \\
d
\end{pmatrix}.
\]

The active set method is applied for solution of control constrained elliptic OCP problem with the all-at-once method in [20] where \( \chi \) denotes the characteristic function for a given set.

3. **A priori error analysis of the optimal control problem**

In the recent years, a priori error estimates were developed for optimal control problems with elliptic and parabolic PDE's using various discretizations of the
state, adjoint and control variables. In [7] a review of various discretization concepts is given for elliptic problems with control and state constraints, among them the piecewise constant approximation of the controls and approximations avoiding explicit discretization of the controls (variational discretization). For linear parabolic equations a priori error estimates were developed using conforming finite element discretization space and discontinuous Galerkin discretization time for unconstrained [11] and for control constrained [12] problems.

There have not been a detailed research concerning a priori error analysis of control problem for Burgers equation. In [4] a priori error bounds were derived by using space-time finite elements for Burgers equation. A stability analysis for the solution of Burgers equation was considered in [5, 8, 26].

In this section, we will derive a priori error estimates for the semi-discretized and fully discretize problems. Our approach for deriving error estimates for Burgers equation is based on techniques used for the semi-implicit discretization of the linear diffusion-convection equations in [14] and for the Burgers equation with space discretization using spectral method in [3].

3.1. Error analysis for the state equation

Multiply both sides of the state equation (3) by a test function \( w \in H^1_0(\Omega) \)

\[
(y_t, w) + \nu(y_x, w_x) + (yy_x, w) = (u, w) \quad \forall w \in H^1_0(\Omega) \quad \text{a.e.} \quad t \in [0, T],
\]

we obtain the weak formulation

\[
(y_t, w) + \nu(y_x, w_x) + (yy_x, w) = (u, w) \quad \forall w \in H^1_0(\Omega),
\]

\[
y(0, \cdot) = y_0(x).
\]

where \((\cdot, \cdot)\) is the inner product in \(L^2(\Omega)\). We assume \(y_0(x) \in L^\infty(\Omega)\) and \(u(x, t) \in L^\infty(\Omega)\).

3.2. Semi-discretization

The finite element discretization of the weak form above is given with \(y^h(t) \in S_h, \ S_h \subset H^1_0(\Omega)\)

\[
(y_t^h, w) + \nu(y_x^h, w_x) + (yy_x^h, w) = (u, w) \quad \forall w \in H^1_0(\Omega),
\]

\[
y^h(0, \cdot) = y_{0h},
\]

where \(|y_{0h}|_{L^\infty(\Omega)} \leq |y_0|_{L^\infty(\Omega)}\).

We assume that the finite element polynomials of degree \(\leq p\) over any mesh have the following property

\[
\inf_{\chi \in S_h} \left\{ \|v - \chi\| + h\|\nabla(v - \chi)\| \right\} \leq Ch^s\|v\|_s
\]

for \(1 \leq s \leq p + 1 = r\) and \(v \in H^s(\Omega) \cup H^1_0(\Omega)\), where \(\| \cdot \|_s\) is the norm on \(H^s\).

Then, the problem (15) has at least a solution [Lemma 4.3,p. 52 of [9]].

**Theorem 3.1:** (Stability) The approximate solution \(y^h\) of (15) is stable for all
\[ t > 0 \text{ and for a constant } C \text{ that does not depend on } h. \]
\[
\|y^h(t)\|^2 + 2\nu \int_0^t \|
abla y^h(\Delta t)\|^2 \Delta t \leq \|y^h(0)\|^2 + \left| \int_0^t (f, y^h) \right| \Delta t,
\]
and
\[
\sup_{0 \leq t \leq T} \|y^h\| \leq C(f, y_0) \quad (17)
\]

**Proof:** Putting \( w^h = y^h \) in (15)
\[
(y^h_t, y^h) + \nu (y^h, y^h_x) + (y^h, y^h_x) = (f, y^h),
\]
and integrating over 0 to \( t \), we obtain the first result. Taking supremum of each side and applying Cauchy-Schwarz inequality on the right hand side gives the second result. \( \square \)

**Theorem 3.2:** (Convergence) Let \( y^h \) and \( y \) be solutions of (15) and (3), respectively. Then,
\[
\|y^h(t) - y(t)\| \leq C \|y_{0h} - y_0\| + h^{r-1}
\]
with \( C = C(y) \).

**Proof:** We introduce the Ritz projection \( P_1 \) into \( S_h \) as the orthogonal projection with respect to the inner product \((v_x, u_x)\)
\[
((P_1 u)_x, w_x) = (u_x, w_x) \quad \forall w \in S_h.
\]
This projection has the following properties (\[23\])
\[
\|(P_1 v - v)_x\| \leq C h^{s-1} \|v\|_s \text{ and } \|P_1 v - v\| \leq C h^s \|v\|_s
\]
for \( 1 \leq s \leq r \) and \( v \in H^s(\Omega) \cap H^1_0(\Omega) \). Let
\[
y^h - y = (y^h - P_1 y) + (P_1 y - y) = \rho^h + \rho.
\]
From the properties of the projection \( P_1 \) above, it is easy to see that the second term is bounded,
\[
\|\rho(t)\| \leq C_1(y) h^r \text{ and } \|\rho_x\| \leq C_1(y) h^{r-1}.
\]
Therefore it is sufficient to give estimates only for \( \rho^h \). We note that
\[
(v^h_t, w) + \nu (v^h_x, w_x) = (y^h_t, w) + \nu (y^h_x, w_x) - (P_1 y_t, w) - \nu (y_x, w_x)
\]
\[
= -(y^h y^h_x, w) - (P_1 y_t, w) + (y_t, w) + (y y_x, w)
\]
\[
= (y y_x - y^h y^h_x, w) + (y_t - P_1 y_t, w)
\]
\[
= -(\rho_t, w) + (y(y - y^h)_x - y^h_y(x - y^h), w).
\]
Let $w = v^h$. Then, applying the Young’s inequality we obtain

$$\frac{1}{2} \frac{d}{dt} \|v^h\|^2 + \nu \|v_x^h\|^2 = -(\rho_t, v^h) + (y(y - y^h)_x - y_x^h(y - y^h), v^h)$$

$$\leq \|\rho_t\| \|v^h\| + |(y(y - y^h)_x, v^h)| + |(y_x^h(y - y^h), v^h)|$$

$$\leq C \epsilon \|v^h\|^2 + \frac{C}{\epsilon} \|\rho_t\|^2 + |(y(y - y^h)_x, v^h)| + |(y_x^h(y - y^h), v^h)|,$$

and, $y^h - y = v^h - \rho$ implies

$$|(y(y - y^h)_x, v^h)| \leq \left( |(y \rho_x, v^h)| + |(y v_x^h, v^h)| \right) \text{ and } |(y_x^h(y - y^h), v^h)| \leq \left( |y_x^h \rho, v^h)| + |(y_x^h v^h, v^h)| \right).$$

For a trilinear term defined by $b(y, u, v) = \int_Q y u_x v \, dx \, dt$ the following estimate in [5] can be used:

$$|b(y, u, v)| \leq C \|y\|^{1/2} \|y_x\|^{1/2} \|u_x\| \|v\| \text{ and } |b(y, y, v)| \leq C \|y_x\|^{3/2} \|y\|^{1/2} \|u_x\|. $$

Then, by using Young’s inequality we obtain the following result

$$L_1 = |(y \rho_x, v^h)| \leq C \|v^h\|^{1/2} \|v_x^h\|^{1/2} \|\rho_x\| \|y\| \leq C \epsilon_4 \|v_x^h\|^2 + \frac{C}{\epsilon_1} \|\rho_x\|^2 \|y\|^2. $$

Similarly,

$$L_2 \leq C \epsilon_2 \|v_x^h\|^2 + \frac{C}{\epsilon_2} \|v^h\|^2 \|y\|^4 \text{ and } L_3 \leq C \epsilon_3 \|v_x^h\|^2 + \frac{C}{\epsilon_3} \|\rho_x\|^2 \|y_x^h\|^2.$$

Writing $y^h = v^h + P_1 y$ and using $(y_x^h y^h, y^h) = 0$ gives

$$((v^h + P_1 y)_x v^h, v^h) = (y_x^h v^h, v^h) + ((P_1 y)_x v^h, v^h) = ((P_1 y)_x v^h, v^h).$$

Using the definition of the projection operator $P_1$ and letting $w = P_1 y$ we get

$$\|P_1 y_x\| \leq C \|y_x\|. $$

Then,

$$L_4 \leq C \|v^h\|^{1/2} \|v_x^h\|^{3/2} \|y^h\| \leq C \|v^h\|^{1/2} \|v_x^h\|^{3/2} \|y_x\| \leq C \epsilon_4 \|v_x^h\|^2 + \frac{C}{\epsilon_4} \|v^h\|^2 \|y_x\|^4.$$ 

Letting $\epsilon = \epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon_4 = \frac{\nu}{\delta}$ gives

$$\frac{1}{2} \frac{d}{dt} \|v^h\|^2 + \nu \|v_x^h\|^2 \leq C \epsilon \|v_x^h\|^2 + \frac{C}{\epsilon} \|\rho_t\|^2 + L_1 + L_2 + L_3 + L_4$$

$$\leq \frac{1}{2} \frac{d}{dt} \|v^h\|^2 \leq C (\nu) \left( \|\rho_t\|^2 + \|\rho_x\|^2 (\|y\|^2 + \|y_x^h\|^2) + \|v^h\|^2 (\|y\|^4$$

$$+ \|y_x\|^4) + \|\rho\|^2 + \|v^h\|^2 \|y_x\|^4 + \|y_x^h\|^2 \|v^h\|^2).$$
Integrating over $[0, t]$, $t \leq T$ implies

$$
\|v^h(t)\|^2 \leq \|v^h(0)\|^2 + C(\nu) \int_0^t \{\|\rho_t\|^2 + \|\rho_x\|^2(\|y\|^2 + \|y_x^h\|^2) + \|v^h\|^2(\|y\|^4 + \|y_x^h\|^4) + \|\rho\|^2\} d\Delta t.
$$

Using the Cauchy-Schwarz inequality and Theorem 5.1 we obtain

$$
\int_0^T \|\rho_x\|^2(\|y\|^2 + \|y_x^h\|^2) d\Delta t \leq C\|\rho_x\|^2_{L^2(0,T;L^2(\Omega))}(\|y\|^2_{L^2(0,T;L^2(\Omega))} + \|y_x^h\|^2_{L^2(0,T;L^2(\Omega))}).
$$

By assumption $y, y_x \in L^4(0,T;L^2(\Omega))$, Gronwall’s inequality shows that

$$
\|v^h(t)\|^2 \leq C\|v^h(0)\|^2 + C(\nu) \int_0^t \{\|\rho_t\|^2 + \|\rho_x\|^2 + \|\rho\|^2\} d\Delta t.
$$

Thus,

$$
\|v^h(t)\| \leq C(\|y_{0h} - y_0\| + h^{r-1}),
$$

which completes the proof. □

### 3.3. Full discretization

Now we develop the stability and convergence analyses of full discretization using the semi-implicit scheme

$$
\begin{aligned}
\frac{1}{\Delta t} (y_h^{i+1} - y_h^i, w_h) + \nu(\nabla y_h^{i+1}, \nabla w_h) + (g^i_h \nabla y_h^i, w_h) &= (f(t_{i+1}), w_h), \\
y_h^0 &= y_{0h}, \text{ for any } w_h \in V_h.
\end{aligned}
$$

**Theorem 3.3:** (Stability) The solutions to Scheme 2 is unconditionally stable and

$$
\|y_h^i\|^2 \leq C^*(\|y_{0h}\|^2 + \frac{T C}{2\nu} \|f\|^2_{L^2(0,T;\Omega)}),
$$

where $C$ and $C^*$ are constants independent of $h, \Delta t$ and $\nu$.

**Proof:**
Leaving $w_h = y_h^{i+1}$ gives

$$
\frac{1}{\Delta t} \|y_h^{i+1}\|^2 - \frac{1}{\Delta t} \|y_h^i\|^2 + \frac{1}{\Delta t} \|y_h^{i+1} - y_h^i\|^2 + \nu \|y_h^{i+1}\|^2 \leq ((y_h^i \nabla y_h^i, y_h^{i+1}) + (f(t_{i+1}), y_h^{i+1})
$$

$$
\leq \|y_h^i \nabla y_h^i, y_h^{i+1}\| + |(f(t_{i+1}), y_h^{i+1})|.
$$
Note that, the Young inequality implies

\[ |\langle y_h^i \nabla y_h^i, y_h^{i+1} \rangle| \leq \frac{C}{2\epsilon} \|y_h^i\|^2 \|\nabla y_h^i\|^2 + \frac{\epsilon}{2}\|\nabla y_h^{i+1}\|^2 \text{ and } \| (f(t_{i+1}), y_h^{i+1}) \| \leq \frac{C}{2\epsilon} \| (f(t_{i+1})) \|^2 + \frac{\epsilon}{2}\|\nabla y_h^{i+1}\|^2. \]

It follows that

\[
\frac{1}{2\Delta t} \| y_h^{i+1} \|^2 - \frac{1}{2\Delta t} \| y_h^i \|^2 + \frac{C}{2\nu} \| y_h^{i+1} - y_h^i \|^2 + \nu \| y_h^{i+1} \|^2 \\
\leq \frac{C}{2\nu} \| y_h^i \|^2 \| \nabla y_h^i \|^2 + \frac{C}{2\nu} \| f(t_{i+1}) \|^2 + \frac{\epsilon}{2}\|\nabla y_h^{i+1}\|^2.
\]

Choosing \( \epsilon = \nu \) gives

\[
\frac{1}{2\Delta t} \| y_h^{i+1} \|^2 - \frac{1}{2\Delta t} \| y_h^i \|^2 \leq \frac{C\Delta t}{2\nu} \| y_h^i \|^2 \| \nabla y_h^i \|^2 + \frac{\Delta t C}{2\nu} \| (f(t_{i+1})) \|^2.
\]

Let now \( m \) be a fixed index, \( 1 \leq m \leq N \). Summing over \( n \) from 0 to \( m - 1 \), we find

\[
\| y_h^n \|^2 \leq \| y_{0,h} \|^2 + \frac{TC}{2\nu} \| f(t_{i+1}) \|^2 (0,T;\Omega), \quad \text{with } C^* = \exp \left( \frac{CT \sum_{n=0}^{m-1} \| \nabla y_h^i \|^2} {2\nu} \right).
\]

\[ \square \]

**Theorem 3.4:** (Convergence) Assume that \( y_0 \in H^1_0(\Omega) \) and the solution to (3) is such that \( \frac{\partial y}{\partial t} \in L^2(0,T;H^1(\Omega)) \) and \( \frac{\partial^2 y}{\partial t^2} \in L^2(0,T;L^2(\Omega)) \). Then, \( y_h^n \) satisfies

\[
\| y_h^n - y(t_n) \|^2 \leq \| (I - P_1) y(t_n) \|^2 + \exp(C^* t_n) \left\{ \| y_{h,0} - P_1 y_0 \|^2 + \frac{C}{\nu} \| (I - P_1) \partial_y (s) \|^2 \right. \\
+ \frac{C(\Delta t)^2}{\nu} \int_{t_n}^{t_{i+1}} \partial^2 \nu (s) \|^2 + \frac{(\Delta t)^2 \nu (t_{i+1})}{\nu} \| \int_{t_n}^{t_{i+1}} \partial_y (s) \|^2 \left\},
\]

where \( C^* \) only depends on \( \nabla y \) and \( \nu \).

**Proof:** We define \( \epsilon^i = y_h^i - y(t_i) \). Let \( \eta^i = y_h^i - P_1 y(t_i) \). Then, \( \epsilon^i = \eta^i + P_1 y(t_i) - y(t_i) \). Since \( \| P_1 y(t_i) - y(t_i) \| \leq \| C \| y(t_i) \| \), then also \( \eta^i \) is bounded. The fully discretized equation with the projector \( P_1 \) is given as

\[
\frac{1}{\Delta t} (\eta_{i+1} - \eta_i, w_h) + \nu (\nabla \eta_h^{i+1}, \nabla w_h) \\
= -\frac{1}{\Delta t} (P_1 y(t_{i+1}) - P_1 y(t_i), w_h) - \nu (\nabla P_1 y(t_{i+1}), \nabla w_h) - (y_h^i \nabla y_h^i, w_h) + (f(t_{i+1}), w_h).
\]

Using the definition of the projector \( P_1 \) we obtain

\[
\nu (\nabla (P_1 y(t_{i+1}), \nabla w_h)) = \nu (\nabla y(t_{i+1}), \nabla w_h) \\
= -\frac{\partial}{\partial t} (y(t_{i+1}), w_h) - (y(t_{i+1}) \nabla y(t_{i+1}), w_h) + (u(t_{i+1}), w_h).
\]
\[
\frac{1}{\Delta t} (\eta^{i+1} - \eta^i, w_h) + \nu (\nabla \eta_h^{i+1}, \nabla w_h)
\]
\[
= - \frac{1}{\Delta t} (P_1 y(t_{i+1}) - P_1 y(t_i), w_h) + (\frac{\partial}{\partial t} y(t_{i+1}), w_h) + (y(t_{i+1}) \nabla y(t_{i+1}), w_h) - (y_h^i \nabla y_h^i, w_h).
\]

The first two terms of the right hand-side can be written as
\[
= - \frac{1}{\Delta t} (P_1 y(t_{i+1}) - P_1 y(t_i), w_h) + (\frac{\partial}{\partial t} y(t_{i+1}), w_h)
\]
\[
= \frac{1}{\Delta t} \left( \int_{t_n}^{t_{i+1}} (I - P_1) \frac{\partial y}{\partial t} (s) ds, w_h \right) + \Delta t \left( \int_{t_n}^{t_{i+1}} \frac{\partial^2 y}{\partial t^2} (s) ds, w_h \right).
\]

Let \( w_h = \eta^{i+1} \). It follows that
\[
\frac{1}{2\Delta t} \|\eta^{i+1}\|^2 - \frac{1}{2\Delta t} \|\eta^i\|^2 + \nu \|\nabla \eta_h^{i+1}\|^2 \leq \frac{1}{\Delta t} \left| \left( \int_{t_n}^{t_{i+1}} (I - P_1) \frac{\partial y}{\partial t} (s) ds, \eta^{i+1} \right) \right|
\]
\[
+ \Delta t \left( \int_{t_n}^{t_{i+1}} \frac{\partial^2 y}{\partial t^2} (s) ds, \eta^{i+1} \right) + \left| \left( y(t_{i+1}) \nabla y(t_{i+1}) - y_h^i \nabla y_h^i, \eta^{i+1} \right) \right|.
\]

Adding and subtracting terms in the equations above, we achieve
\[
(y(t_{i+1}) \nabla y(t_{i+1}), \eta^{i+1}) - (y_h^i \nabla y_h^i, \eta^{i+1})
\]
\[
= -\frac{1}{2} \left( (y(t_{i+1}) - y(t_i)) y(t_{i+1}) + (y(t_i) - y_h^i) y_h^i + y(t_i)(y(t_{i+1}) - y_h^i), \nabla \eta^{i+1} \right).
\]

Also,
\[
(y(t_{i+1}) - y_h^i) = (y(t_{i+1}) - y(t_i)) + y(t_i) - y_h^i).
\]

Then, using the boundedness of \( \nabla y_h^i \) we get
\[
\left| (y(t_{i+1}) \nabla y(t_{i+1}), \eta^{i+1}) - (y_h^i \nabla y_h^i, \eta^{i+1}) \right|
\]
\[
\leq C \|\nabla \eta^{i+1}\| \left( \|\nabla y(t_{i+1})\| \|y(t_{i+1}) - y(t_i)\| + \|\nabla y_h^i\| \|y(t_i) - y_h^i\| \right)
\]
\[
+ \|\nabla y(t_i)\| \|y(t_{i+1}) - y(t_i)\| + \|\nabla y(t_i)\| \|y(t_i) - y_h^i\| \right)
\]
\[
\leq C \|\nabla \eta^{i+1}\| \|y(t_{i+1}) - y(t_i)\| + \tilde{C} \|\nabla \eta^{i+1}\| \frac{\|y(t_i) - y_h^i\|}{\eta^i}.
\]
Now, using \( y(t_{i+1}) - y(t_i) = \Delta t \int_{t_i}^{t_{i+1}} \frac{\partial y}{\partial t}(s)\,ds \)

\[
\frac{1}{2\Delta t}\|\eta^{i+1}\|^2 - \frac{1}{2\Delta t}\|\eta^i\|^2 + \nu\|\nabla\eta_h^{i+1}\|^2
\]

\[
\leq C \frac{\Delta t}{\nu} \|\int_{t_i}^{t_{i+1}} (I - P_1) \frac{\partial y}{\partial t}(s)\,ds\|^2 + \nu \frac{\Delta t}{4} \|\nabla \eta^{i+1}\|^2 + \frac{C(\Delta t)^2}{\nu} \|\int_{t_i}^{t_{i+1}} \frac{\partial^2 y}{\partial t^2}(s)\,ds\|^2 + \frac{\nu}{4}\|\nabla \eta^{i+1}\|^2
\]

\[
+ \frac{\Delta t \tilde{C}(\nabla y(t_i))}{\nu} \|\int_{t_i}^{t_{i+1}} \frac{\partial y}{\partial t}(s)\,ds\|^2 + \nu \frac{\Delta t}{4}\|\nabla \eta^{i+1}\|^2.
\]

Then,

\[
\|\eta^{i+1}\|^2 - \|\eta^i\|^2 \leq C \frac{\Delta t}{\nu} \|\int_{t_i}^{t_{i+1}} (I - P_1) \frac{\partial y}{\partial t}(s)\,ds\|^2 + \nu \frac{\Delta t}{4} \|\nabla \eta^{i+1}\|^2 + \frac{C(\Delta t)^2}{\nu} \|\int_{t_i}^{t_{i+1}} \frac{\partial^2 y}{\partial t^2}(s)\,ds\|^2
\]

\[
+ \frac{\Delta t \tilde{C}(\nabla y(t_i))}{\nu} \|\int_{t_i}^{t_{i+1}} \frac{\partial y}{\partial t}(s)\,ds\|^2 + \nu \frac{\Delta t}{4}\|\nabla \eta^{i+1}\|^2.
\]

Summing over \( n \) from 0 to \( m - 1 \) and using discrete Gronwall inequality, the desired result is obtained.

3.4. Discrete adjoint and error estimates

Discrete adjoint system (12) corresponds to a weak formulation following as

\[
\begin{cases}
\frac{1}{\Delta t}(p^i_h - p^{i+1}_h, w_h) + \nu (\nabla p^i_h, \nabla w_h) = (y(t_i) - y_d(t_i), w_h), \\
p^T_h = 0, \text{ for any } w_h \in V_h.
\end{cases}
\]

(20)

Theorem 3.5: (Stability) The solution to (20) is stable and satisfies

\[
\|p^i_h\|^2 \leq C^\ast (\|p_{T,h}\|^2 + \frac{T C}{2\nu} \|f\|^2_{L^2(0,T;\Omega)}).
\]

Proof: We let \( w_h = p^j_h \). We perform the same strategy as in Theorem 3.3 and result follows.

Before developing an error bound for the adjoint variable, we recall the continuous adjoint equation

\[
p^*_t + \nu \Delta p^* + y^* \nabla p^* = y_d - y^* \quad \text{in } Q,
\]

\[
p^*(t,0) = p^*(t,1) = 0 \quad \text{on } \Sigma,
\]

\[
p^*(T) = 0 \quad \text{in } \Omega.
\]

(21)

Theorem 3.6: (Convergence) Assume that the solution to (21) is such that \( \frac{\partial p^*}{\partial t} \in \)}
L^2(0, T; H^1_0(\Omega)) and \( \frac{\partial^2 p}{\partial t^2} \in L^2(0, T; L^2(\Omega)) \). Then \( p^n_h \), satisfies

\[
\|p^n_h - p(t_n)\|^2 \leq \|(I - P_1)p(t_n)\|^2 + \exp(T) \left\{ \frac{C}{\nu} \left( \int_0^{t_n} \|(I - P_1) \frac{\partial y}{\partial t}(s)\| ds \right)^2 \right\} + \frac{C(\Delta t)^2}{\nu} \int_0^{t_n} \left\| \frac{\partial^2 p}{\partial t^2}(s) \right\|^2 + (\Delta t)^2 \left\| y(t_i) \nabla p(t_i) \right\|^2 \},
\]

where \( C^* \) only depends on \( \nabla y \) and \( \nu \).

**Proof:** Since \( y^n_h \) is stable then for simplicity let us consider the following weak formulation

\[
\frac{1}{\Delta t}(\eta^i - \eta^{i+1}, w_h) + \nu(\nabla \eta^i_h, \nabla w_h) = \int_0^{t_i} (\partial_t^2 p(t_i), \nabla w_h) dt,
\]

Following the same procedure as in the proof of convergence theorem of state equation, for \( i = N, ..., 1 \) we get

\[
\frac{1}{\Delta t}(\eta^i - \eta^{i+1}, w_h) + \nu(\nabla \eta^i_h, \nabla w_h) = -\frac{1}{\Delta t}(P_1 p(t_i) - P_1 p(t_{i+1}), w_h) - \nu(\nabla P_1 p(t_i), \nabla w_h) + (\tilde{f}(t_i), w_h).
\]

Using the definition of \( P_1 \) and continuous adjoint equation, it follows that

\[
\frac{1}{\Delta t}(\eta^i - \eta^{i+1}, w_h) + \nu(\nabla \eta^i_h, \nabla w_h) = -\frac{1}{\Delta t}(P_1 p(t_i) - P_1 p(t_{i+1}), w_h) + (\frac{\partial}{\partial t} p(t_i), w_h) + (y(t_i) \nabla p(t_i), w_h),
\]

\[
\frac{1}{2\Delta t}||\eta^i||^2 - \frac{1}{2\Delta t}||\eta^{i+1}||^2 + \frac{1}{2\Delta t}||\eta^i - \eta^{i+1}||^2 + \nu||\nabla \eta^i_h||^2 \leq \frac{1}{\Delta t} |(\int_{t_{i+1}}^{t_i} (I - P_1) \frac{\partial p}{\partial t}(s) ds, \eta^i)| + \Delta t |(\int_{t_{i+1}}^{t_i} \frac{\partial^2 p}{\partial t^2}(s) ds, \eta^i)| + |(y(t_i) \nabla p(t_i), \eta^i)|.
\]

Considering the term \( (y(t_i) \nabla p(t_i), \eta^i) \) gives

\[
\int (y(t_i) \nabla p(t_i), \eta^i) \leq \int |(y(t_i) \nabla p(t_i), \eta^i - \eta^{i+1})| + \int |(y(t_i) \nabla p(t_i), \eta^{i+1})| \leq \frac{1}{\Delta t} \|y(t_i) \nabla p(t_i)\| \|\eta^i - \eta^{i+1}\| + \frac{\Delta t}{2} \|y(t_i) \nabla p(t_i)\|^2 + \frac{1}{2\Delta t} ||\eta^{i+1}\|^2.
\]

\[
||\eta^i||^2 - ||\eta^{i+1}||^2 + 2\Delta t \nu \|\nabla \eta^i_h\|^2 \leq 2 \int (I - P_1) \frac{\partial p}{\partial t}(s) ds, \eta^i) \| + 2(\Delta t)^2 \left( \int_{t_{i+1}}^{t_i} \frac{\partial^2 p}{\partial t^2}(s) ds, \eta^i) \right) + \Delta t \int \|y(t_i) \nabla p(t_i)\|^2 + \|\eta^{i+1}\|^2.
\]

Summation over \( N \) to 0 and using Gronwall’s lemma gives the desired result. \( \square \)
3.5. Error in the control variable

In this section, we find an error estimate for the control variable. Actually the error analysis of control variable is related to adjoint equation, which is given by the following theorem.

**Theorem 3.7**: The solutions to continuous and discretized control problem satisfy

\[
\|\bar{u} - \bar{u}_h^n\| \leq \frac{1}{\alpha} \|p - p_h^n\| + \|\bar{u} - P_1 \bar{u}\|. \tag{22}
\]

In order to prove the Theorem 3.7, we need some results related to the cost function. We recall the continuous cost function

\[
J(y, u) = \frac{1}{2} \|y - y_d\|^2_Q + \frac{\alpha}{2} \|u\|^2_Q.
\]

The reduced cost function can be stated as

\[
j(u) = J(S(u), u),
\]

where \(S\) is the solution operator. The derivative of the reduced cost function can be stated as

\[
j'(u)(\delta u) = (p, \delta u) + \alpha (u, \delta u),
\]

where \(p\) corresponds to the adjoint variable. The necessary and sufficient optimality conditions are

\[
j'(\bar{u})(\delta u - \bar{u}) \geq 0 \quad \forall \delta u \in Q, \quad \text{and} \quad j''(u)(\delta u, \delta u) \geq \alpha \|\delta u\|^2 \quad \forall \delta u \in Q,
\]

where \(\bar{u}\) denotes the optimal solution.

We can derive similar results for the discretized problem. We assume that \(S_{hn}\) is the discrete solution operator between control and state variables. Letting \(j_{hn}(u_h^n) = J_{hn}(S_{hn}(u_h^n), u_h^n)\), the optimality conditions become

\[
j_{hn}'(\bar{u}_h^n)(\delta u_h^n - \bar{u}_h^n) \geq 0 \quad \forall \delta u_h^n \in V_h \quad \text{and} \quad j_{hn}''(u_h^n)(\delta u_h^n, \delta u_h^n) \geq \alpha \|\delta u_h^n\|^2 \quad \forall \delta u_h^n \in V_h.
\]

**Lemma 3.8**: The error between the solutions of the continuous and discretized control problem satisfies

\[
\|j'(u)(r) - j_{hn}'(u)(r)\| \leq \|p(u) - p_h^n(u)\| \|r\| \quad \text{for} \quad u, r \in Q.
\]

**Lemma 3.9**: Let \(u\) be a given control. The error between the continuous state \(y\) and the discrete state \(y_{hn}^n\) can be estimated as:

\[
\|y - y_{hn}^n\| \leq C(\Delta t)^{1/2} \|u - q\| + O(h + \Delta t),
\]

where \(q \in Q\) and \(O\) denotes the order.

**Proof**: We take \(u\) instead of \(f + u\) in the Eqn. (3) and \(q\) in the place of \(f(t^{i+1})\) in (18). Then we have additional terms to Theorem 3.2. Proceeding the same steps gives the desired result. \(\square\)
Lemma 3.10: Let $q$ be a given control. The error between the continuous adjoint state $p$ and the discrete adjoint state $p_h^n$ can be estimated as:

$$\|p - p_h^n\| \leq C \Delta t \|u - q\| + O(h + \Delta t),$$

where $q \in Q$ and $O$ denotes the order and $C$ is independent of $h$ and $\Delta t$.

**Proof:** The same procedure holds as in the proof of Lemma 3.9. \hfill \square

We now come to the proof of the Theorem 3.7. Let $\bar{u}$ and $\bar{u}_h^n$ be the optimal solutions to discrete and continuous control problems, respectively. For an arbitrary $q$ we write

$$\bar{u} - \bar{u}_h^n = \bar{u} - q + q - \bar{u}_h^n.$$  

From (23) and using the optimality of $u^n_h$

$$\alpha \|q - \bar{u}_h^n\|^2 \leq j''(q)(q - \bar{u}_h^n, q - \bar{u}_h^n) = j'_h(\bar{u})(q^n_h - \bar{u}_h^n) - j'_h(\bar{u}_h^n)(q^n_h - \bar{u}_h^n),$$

we obtain

$$\alpha \|q^n_h - \bar{u}_h^n\|^2 \leq j'_h(\bar{u})(q^n_h - \bar{u}_h^n) - j'_h(\bar{u}_h^n)(q^n_h - \bar{u}_h^n) \leq \|p(\bar{u}) - p_h^n(\bar{u})\| \|q^n_h - \bar{u}_h^n\|.$$

From (24) and using the optimality of $u_h^n$

$$j'(\bar{u}_h^n)(q^n_h - \bar{u}_h^n) = 0 = j'(\bar{u})(q^n_h - \bar{u}_h^n),$$

we obtain

$$\alpha \|q^n_h - \bar{u}_h^n\|^2 \leq j'_h(\bar{u})(q^n_h - \bar{u}_h^n) - j'_h(\bar{u}_h^n)(q^n_h - \bar{u}_h^n) \leq \|p(\bar{u}) - p_h^n(\bar{u})\| \|q^n_h - \bar{u}_h^n\|.$$

Finally, $\|q^n_h - \bar{u}_h^n\| \leq \frac{1}{\alpha} \|p(\bar{u}) - p_h^n(\bar{u})\|$. We let $q = P_1(\bar{u})$ and use Lemma 3.8 to get the desired result.

**Corollary 3.11:** The solutions to continuous and discretized control problem satisfy the following estimations:

$$\|\bar{u} - p_h^n(\bar{u})\| \leq O(h + h \Delta t + \Delta t),$$

**Proof:** This corollary is a result of the projection $P_1$ property and Lemma 3.10. \hfill \square

In this section, we provide an error estimate for the control constrained case. Since (24) does not hold any more, we can not use the same argument as in the unconstrained problem. We recall that there exists an additional constraint as

$$u_a(t, x) \leq u(t, x) \leq u_b(t, x)$$

in $Q$.

This condition leads to a variational inequality as

$$j'(\bar{u})(u - \bar{u}) \geq 0$$

for all $u \in U_{ad}$, where $j(u)$ is given as $\min J(y, u) \leftrightarrow \min j(u)$. It is known that the inequality above is equivalent to

$$\bar{u} = \Pi_{Q_{ad}} \left( -\frac{1}{\alpha} \bar{p} \right),$$

where

$$\bar{u} = \Pi_{Q_{ad}} \left( -\frac{1}{\alpha} \bar{p} \right).$$
where $\Pi_{Q_{ad}}(t, x) := \max(u_a, \min(u_b, r(t, x)))$ is the projection into the admissible space $Q_{ad}$.

As in the continuous case we can deduce a projection formula as

$$\bar{u}_h^n = \Pi_{Q_{ad}}\left(-\frac{1}{\alpha}\bar{p}_h^n\right).$$

This projection $\Pi_{Q_{ad}}$ satisfies the regularity properties

$$\left\| \Pi_{Q_{ad}}\left(-\frac{1}{\alpha}\bar{p}\right) - \Pi_{Q_{ad}}\left(-\frac{1}{\alpha}\bar{p}_h^n\right) \right\|_{L^2(Q)} \leq \frac{1}{\alpha} \|\bar{p} - \bar{p}_h^n\|_{L^2(Q)}.$$  \hfill (27)

**Theorem 3.12:** Let $\bar{u}$ and $\bar{u}_h^n$ be the solutions to continuous and discrete optimal control problems respectively. Then

$$\|\bar{u} - \bar{u}_h^n\| \leq \frac{1}{\alpha} \|\bar{p} - \bar{p}_h^n\||,$$

where $\bar{p}$ and $\bar{p}_h^n$ are the corresponding continuous and discrete adjoint state variables, respectively.

**Proof:** This proof is a simple result of (26) and (27). \hfill □

### 3.6. Error bound for the cost functional

In order to measure error bound in terms of the cost functional we follow the approach in [10]

$$J(\bar{y}, \bar{u}) - J_{hn}(\bar{y}_h^n, \bar{u}_h^n) = \frac{1}{2}(\|\bar{y} - y_d\|^2 - \|\bar{y}_h^n - y_d\|^2) + \frac{\alpha}{2}(\|\bar{u}\|^2 - \|u_h^n\|^2) \quad (28)$$

Using the following identity

$$\|z\|^2 - \|v\|^2 = 2(z, z - v) - \|z - v\|^2$$

we obtain

$$\frac{1}{2}\|\bar{y} - y_d\|^2 - \|\bar{y}_h^n - y_d\|^2 \leq \{\|\bar{y}\| + \|y_d\|\}\|\bar{y} - \bar{y}_h^n\| + \frac{1}{2}\|\bar{y} - \bar{y}_h^n\|^2. \quad (29)$$

This term shows the convergence order $O(h + \Delta t)$. We add and subtract $P_1\bar{u}$ for the second term in (28).

$$(\bar{u}, \bar{u} - \bar{u}_h^n) = (\bar{u}, \bar{u} - P_1\bar{u}) + (\bar{u}, P_1\bar{u} - \bar{u}_h^n) = \|\bar{u} - P_1\bar{u}\|^2 + (\bar{u}, P_1\bar{u} - \bar{u}_h^n).$$

$$\left| (\bar{u}, \bar{u} - \bar{u}_h^n) \right| \leq \|\bar{u} - P_1\bar{u}\|^2 + \|\bar{u}\|\|P_1\bar{u} - \bar{u}_h^n\| \leq \|\bar{u} - P_1\bar{u}\|^2 + \|\bar{u}\|\|P_1\bar{u} - \bar{u}\| + \|\bar{u}\|\|\bar{u} - \bar{u}_h^n\| \quad (30)$$

Now estimates (29) and (30) gives the following result.
Corollary 3.13: Let the cost functional \( J \) and \( J_{hn} \) be the solutions to continuous and discrete problems, respectively. Then the following estimate holds:

\[
|J(\bar{y}, \bar{u}) - J_{hn}(\bar{y}_h^n, \bar{u}_h^n)| \leq \{||\bar{y}|| + ||y_d||\}||\bar{y} - \bar{y}_h^n|| + \frac{1}{2}||\bar{y} - \bar{y}_h^n||^2
+ \frac{\alpha}{2}||\bar{u} - P_1\bar{u}||^2 + ||\bar{u}|| ||P_1\bar{u} - \bar{u}|| + ||\bar{u}|| ||\bar{u} - \bar{u}_h^n||
= O(h + \Delta t)
\]

3.7. Numerical results

We carried out some numerical tests for both unconstrained and control constrained control problems of Burgers equation. Because the exact solution of the optimal control problems are unknown, we have used the cost function to show the convergence of the numerical solutions.

Run 1. (Unconstrained problem) [6] As a numerical example we have chosen the following optimal control problem in [6] with the parameters \( \alpha = 0.05, \nu = 0.01 \), with the desired state \( y_d(t, x) = y_0 \) and with the initial condition

\[
y_0 = \begin{cases} 
1 & \text{in } (0, \frac{1}{2}], \\
0 & \text{otherwise.}
\end{cases}
\]

Based on this result for the error in the control variable, estimates of optimal order for the error in the state and adjoint state variable and also in terms of the cost functional are shown. In order to compare the numerically obtained orders of convergence

\[
OC(\Delta t)_i = \frac{\ln J_{\Delta t_{i-1}} - \ln J_{\Delta t_i}}{\ln(\Delta t_{i-1}) - \ln(\Delta t_i)} \quad \text{and} \quad OC(h)_i = \frac{\ln J_{h_{i-1}} - \ln J_{h_i}}{\ln(h_{i-1}) - \ln(h_i)}
\]

with the a priori error estimates developed in the previous section, the space or time step sizes are held fixed. That is, we consider first the behavior of the error for a sequence of discretizations with decreasing size of the time steps and for fixed mesh size in space. Secondly, we examine the behavior of the error under refinement of the spatial grids for fixed time steps.

In Table 1, we present the results for a fixed space mesh \( h = 2^{-7} \). The observed order of convergence of the scheme is as expected. When \( \Delta t \to 0 \), the order of the semi-implicit scheme is one, as predicted by the a priori error estimates.

<table>
<thead>
<tr>
<th>( \Delta t )</th>
<th>( J_{\Delta t_i} )</th>
<th>( J_{\Delta t_{i-1}} - J_{\Delta t_i} )</th>
<th>order</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2^{-4} )</td>
<td>7.062e-2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>( 2^{-5} )</td>
<td>6.892e-2</td>
<td>1.706e-3</td>
<td>-</td>
</tr>
<tr>
<td>( 2^{-6} )</td>
<td>6.960e-2</td>
<td>6.794e-4</td>
<td>1.32</td>
</tr>
<tr>
<td>( 2^{-7} )</td>
<td>6.999e-2</td>
<td>3.992e-4</td>
<td>0.76</td>
</tr>
<tr>
<td>( 2^{-8} )</td>
<td>7.019e-2</td>
<td>1.916e-4</td>
<td>1.05</td>
</tr>
</tbody>
</table>

Now we fix \( \Delta t \) at \( 2^{-7} \) in order to see the order of convergence for space variable. As we expect we have first order convergency in space one.
Table 2. Unconstrained problem with fixed $\Delta t = 2^{-7}$

<table>
<thead>
<tr>
<th>$h_i$</th>
<th>$J_{h_i}$</th>
<th>$J_{h_{i-1}} - J_{h_i}$</th>
<th>order</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^{-3}$</td>
<td>5.210e-2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>6.390e-2</td>
<td>8.880e-3</td>
<td>-</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>6.744e-2</td>
<td>3.540e-3</td>
<td>1.12</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>6.901e-2</td>
<td>1.570e-3</td>
<td>1.38</td>
</tr>
<tr>
<td>$2^{-7}$</td>
<td>6.999e-2</td>
<td>9.840e-4</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Figure 1. Unconstrained problem: optimal state and control computed with $h = 2^{-7}$, $\Delta t = 2^{-6}$

In Figure 1, the optimal state and control are shown for the unconstrained problem. The solutions are similar to those in [6].

Run 2. (Control constrained problem) We choose the problem in [25]. We consider the same space as in example of unconstrained problem, with $\nu = 0.01$, $\alpha = 0.0175$ and $y_0 = 0$ and with the desired state is given as

The cost function and the observed order of the semi-implicit scheme given in Table 3 & Table 4 confirm the predicted first order convergence by the a priori error analysis.
The optimal state and control are shown in Fig. 3. The results are similar to those in [25] obtained with the SQP method.

4. Conclusions

In this paper we have applied the one-shot-approach for solving OPC with the unsteady Burgers equation and we have developed error estimates for the states, adjoints and controls. In order to apply the all-at-once method the the state equation has to linearized. It turns out the semi-implicit discretization in time is an effective linearization method for the OPC problems with Burgers equation. Because we deal here with Burgers equation in one space dimension, we were able to solve the saddle point system with direct solvers. When Burgers equation in two-dimensional is considered, the large saddle point system can only be solved iteratively with preconditioners. This will be our next task.

5. Acknowledgements

The first author would like to thank to Martin Stoll to helpful discussions at Max-Planck Institute Magdeburg. This paper was completed when the second author was visiting Fulbright scholar at Rice University. He is very grateful to Fulbright for...
the support and to the Department of Computational and Applied Mathematics for the hospitality.

References