4-1-2010

Analysis of Block Parareal Preconditioners for Parabolic Optimal Control Problems

Tarek P. Mathew

Marcus Sarkis
Worcester Polytechnic Institute, msarkis@wpi.edu

C. E. Schaerer

Suggested Citation

Mathew, T. P., Sarkis, M., & Schaerer, C. E. (2010). Analysis of Block Parareal Preconditioners for Parabolic Optimal Control Problems. *SIAM Journal on Scientific Computing, 32*(3), 1180-1200. http://dx.doi.org/10.1137/080717481

ANALYSIS OF BLOCK PARAREAL PRECONDITIONERS FOR PARABOLIC OPTIMAL CONTROL PROBLEMS

T. P. MATHEW*, M. SARKIS*†, AND C. E. SCHAERER*

Abstract. In this paper, we describe block matrix algorithms for the iterative solution of *large scale* linear-quadratic optimal control problems arising from the optimal control of parabolic partial differential equations over a finite control horizon. We describe three iterative algorithms. The first algorithm employs a CG method for solving a symmetric positive definite reduced linear system involving only the unknown control variables. This system can be solved using the CG method, but requires double iteration. The second algorithm is designed to avoid double iteration by introducing an auxiliary variable. It yields a symmetric indefinite system and a positive definite block preconditioner. The third algorithm uses a symmetric positive definite block diagonal preconditioner for the saddle point system and is based on the parareal algorithm. Theoretical results show that the preconditioned algorithm has optimal convergence properties and parallel scalability. Numerical experiments are provided to confirm the theoretical results.

Key words. Parabolic PDE, preconditioners, parareal, multiple shooting, control problems, Krylov methods

AMS subject classifications. 49J20, 65F10, 65M55

1. Introduction.

Let (t_0, t_f) denote a time interval, let $\Omega \subset \mathbb{R}^2$ be a convex polygonal domain of size of order $O(1)$ and let $\mathcal{A}$ be a coercive map from a Hilbert space $L^2(t_o, t_f; Y)$ to $L^2(t_o, t_f; Y')$, where $Y = H_0^1(\Omega)$ and $Y' = H^{-1}(\Omega)$, i.e., the dual of Y with respect to the pivot space $H = L^2(\Omega)$. Denote the state variable space as $\mathcal{Y} = \{z \in L^2(t_o, t_f; Y) : z_t \in L^2(t_o, t_f; Y')\}$. It can be shown that $\mathcal{Y} \subset \mathcal{C}^0([t_o, t_f]; H)$; see [15]. Given $y_o \in H$, we consider the following state equation on (t_0, t_f) with $z \in \mathcal{Y}$:

$$\begin{cases} z_t + \mathcal{A}z &= \mathcal{B}v \quad \text{for} \quad t_o < t < t_f, \\ z(0) &= y_o. \end{cases} \tag{1.1}$$

The distributed control v belongs to an admissible space $\mathcal{U} = L^2(t_o, t_f; U)$, where in our application $U = L^2(\Omega)$, and $\mathcal{B}$ is an operator in $\mathcal{L}(\mathcal{U}, L^2(t_o, t_f; H))$. It can be shown that the problem (1.1) is well-posed, see [15], and we indicate the dependence of z on $v \in \mathcal{U}$ using the notation $z(v)$. Given the target functions $\tilde{y}$ in $L^2(t_o, t_f; H)$, and the parameters $q > 0$, $r > 0$ and $s \geq 0$, we shall employ the following cost function which we associate with the state equation (1.1):

$$\begin{aligned} J(z(v), v) := \frac{q}{2} \int_{t_0}^{t_f} \|z(v)(t,\cdot) - \tilde{y}(t,\cdot)\|^2_{L^2(\Omega)}\, dt + \frac{r}{2} \int_{t_0}^{t_f} \|v(t,\cdot)\|^2_{L^2(\Omega)}\, dt \\ + \frac{s}{2} \|z(v)(t_f,\cdot) - \tilde{y}(t_f,\cdot)\|^2_{L^2(\Omega)}. \end{aligned} \tag{1.2}$$

For simplicity, we assume that $y_o \in Y$ and $\tilde{y} \in L^2(t_o, t_f; Y)$, and normalize $q = 1$. The optimal control problem for equation (1.1) consists of finding a controller $u \in \mathcal{U}$ which *minimizes* the cost function (1.2):

$$J(y, u) \; := \; \min_{v \in \mathcal{U}} J(z(v), v). \tag{1.3}$$

*Instituto de Matemática Pura e Aplicada - IMPA, E. D. Castorina 110, 22460-320, Rio de Janeiro, RJ, Brazil. (msarkis@impa.br, cschaer@fluid.impa.br, tmathew@poonithara.org). C. E. S. is currently at Universidad Nacional de Asunción, Politécnica, SL, Paraguay. This research was supported partially by CNPq: 500075/2002-6.

†Worcester Polytechnic Institute, Worcester, MA 01609, USA. (msarkis@wpi.edu).

Since $0 < r$ (regularization parameter) and $0 \leq q, s$, the optimal control problem (1.3) is well-posed (see [15]).

In this paper, we describe block matrix iterative algorithms for solving an "all at once" discretization [5, 11, 12, 22, 23] of the optimal control problem (1.3). The problem (1.3) seeks a control function u (optimal forcing term in the sense of the cost function (1.2)), such that the solution $z(v)$ to the parabolic equation closely matches the given "tracking" function $\tilde{y}$ on the interval (t_o, t_f). We also require that $z(v)(t_f, \cdot)$ be close to $\tilde{y}(t_f, \cdot)$ by introducing the term $\frac{s}{2} \|z(v)(t_f, \cdot) - \tilde{y}(t_f, \cdot)\|^2_{L^2(\Omega)}$ in the cost function. The constrained minimization problem (1.3) minimizes (1.2) subject to the constraint (1.1), where the *quadratic* cost functional (1.2), is a square norm of the difference between the solution to the parabolic equation and the tracking function with appropriate *regularization*, while the *linear* constraint (1.1) is a parabolic equation involving the state variable.

Our discussion is organized as follows. In § 2, we describe the finite dimensional linear-quadratic optimal control problem obtained using an *all at once* discretization of the control problem, where the spatial discretization is obtained by the finite element method, and the temporal discretization by the θ-scheme. The discretization of the functional is obtained by the finite element method [5, 6, 20]. This transforms the optimal control problem into a large algebraic constrained minimization problem, where optimality conditions yield a large saddle point system involving the state variables $\mathbf{y}$, the control variables $\mathbf{u}$ and the Lagrange multipliers $\mathbf{p}$, see [2, 1, 3].

In § 3, we describe the algorithms. Using a reduction approach employed in [21, 22], we obtain a symmetric positive definite reduced system for the unknown *control variables* (with low dimension in realistic situations). We refer this algorithm as "the reduction to $\mathbf{u}$ algorithm" and a preconditioner is described. We prove that the rate of convergence is independent of the spatial discretization parameter h. If the parameter $s = 0$, the rate of convergence depends only on the parameter r. If $s > 0$, we prove that the rate of convergence depends on the time parameter τ and the parameter r (see expression (3.17)). As a result, the Preconditioned Conjugate Gradient method (PCG) can be used to solve (3.1), but double iterations are required. To overcome this drawback, we introduce an auxiliary variable resulting in a symmetric indefinite ill-conditioned system on the auxiliary and control variables. For this expanded system we employ a symmetric positive definite block diagonal preconditioner [25]. We also prove that under the same conditions of the first preconditioner ($s = 0$), this preconditioner is independent of the h and τ parameters when MINRES acceleration is used. Results analogous to the first preconditioner are obtained if $s \neq 0$; see Theorem 4.4.

In § 4, we present a saddle point preconditioner based on the *parareal* algorithm [16, 19, 24, 8, 10, 26] and iterative shooting methods [12, 27, 4]. This preconditioner yields a rate of convergence independent of the mesh size h with a dependence on τ. If the parareal preconditioner is used as an approximate solver in the context of the reduction to $\mathbf{u}$ algorithm, and enough number of parareal iterations is performed, then the algorithm depends on τ for $s = 0$, and τ and r for $s \neq 0$. Finally, in § 5, numerical tests confirm the theoretical results and show that the parareal preconditioner yields scalability when the number of subdomains is increased.

2. The discretization and the saddle point system . To discretize the state equation (1.1) in space, we apply the finite element method to its weak formulation for each fixed $t \in (t_o, t_f)$. We choose a quasi-uniform triangulation $\mathcal{T}_h(\Omega)$ of Ω, and employ the $\mathcal{P}_1$ conforming finite element space $Y_h \subset Y$ for $z(t, \cdot)$, and the $\mathcal{P}_0$ finite element space $U_h \subset U$ for approximating $v(t,\cdot)$. Let $\{\phi_j\}_{j=1}^{\hat{q}}$ and $\{\psi_j\}_{j=1}^{\hat{p}}$ denote the standard basis functions for Y_h and U_h, respectively. Throughout the paper we use the same notation $z \in Y_h$ and $z \in \mathbb{R}^{\hat{q}}$, or $v \in U_h$ and $v \in \mathbb{R}^{\hat{p}}$, to denote both a finite element function in space and its corresponding vector representation, and to indicate their time dependence, we denote $\underline{z}$ and $\underline{v}$, respectively.

A discretization in space of the continuous time linear-quadratic optimal control problem will seek to minimize the following quadratic functional:

$$J_h(\underline{z},\underline{v}) := \frac{1}{2}\int_{t_o}^{t_f} (\underline{z}-\underline{\tilde{y}})^T(t) M_h (\underline{z}-\underline{\tilde{y}})(t)\,dt + \frac{r}{2}\int_{t_o}^{t_f} \underline{v}^T(t) R_h \underline{v}(t)\,dt \\ + \frac{s}{2}(\underline{z}(t_f)-\underline{\tilde{y}}(t_f))^T M_h(\underline{z}(t_f)-\underline{\tilde{y}}(t_f)) \tag{2.1}$$

subject to the *constraint* that $\underline{z}$ satisfies the discrete equation of state:

$$M_h \dot{\underline{z}} + A_h\,\underline{z} = B_h \underline{v}, \quad \text{for} \quad t_o < t < t_f; \quad \text{and} \quad \underline{z}(t_o) = y_o^h. \tag{2.2}$$

Here $(\underline{z}-\underline{\tilde{y}}^h)(t)$ and $(\underline{z}(t_f)-\underline{\tilde{y}}(t_f))$ denote the tracking and the final error. Here $\underline{\tilde{y}}^h(t)$ and y_0^h belong to Y_h and are approximations to $\underline{\tilde{y}}(t)$ and y_o (for instance, use $L^2(\Omega)$-projections into Y_h). The matrices $M_h, A_h \in \mathbb{R}^{\hat{q}\times\hat{q}}$, $B_h \in \mathbb{R}^{\hat{q}\times\hat{p}}$ and $R_h \in \mathbb{R}^{\hat{p}\times\hat{p}}$ have entries $(M_h)_{ij} := (\phi_i,\phi_j)$, $(A_h)_{ij} := (\phi_i, \mathcal{A}\phi_j)$, and $(B_h)_{ij} := (\phi_i, \mathcal{B}\psi_j)$ and $(R_h)_{ij} := (\psi_i,\psi_j)$, where $(\cdot,\cdot)$ denotes the $L^2(\Omega)$ inner product.

To obtain a temporal discretization of (2.1) and (2.2), we partition $[t_o,\, t_f]$ into $\hat{l}$ equal sub-intervals with time step size $\tau = (t_f - t_o)/\hat{l}$. We denote $t_l = t_o + l\,\tau$ for $0 \le l \le \hat{l}$. Associated with this partition, we assume that the state variable $\underline{z}$ is continuous in $[t_o, t_f]$ and linear in each sub-interval $[t_{l-1}, t_l]$, $1 \le l \le \hat{l}$ with associated basis functions $\{\vartheta_l\}_{l=0}^{\hat{l}}$. Denoting $z_l \in \mathbb{R}^{\hat{q}}$ as the nodal representation of $\underline{z}(t_l)$, we have $\underline{z}(t) = \sum_{l=0}^{\hat{l}} z_l \vartheta_l(t)$. The control variable $\underline{v}$ is assumed to be time discontinuous and constant in each sub-interval (t_{l-1}, t_l) with basis functions $\{\chi_l\}_{l=1}^{\hat{l}}$. Denoting $v_l \in \mathbb{R}^{\hat{p}}$ as the nodal representation of $\underline{v}(t_l - (\tau/2))$, yields $\underline{v}(t) = \sum_{l=1}^{\hat{l}} v_l \chi_l(t)$.

The corresponding discretization of the expression (2.1) yields:

$$J_h^\tau(\mathbf{z},\mathbf{v}) = \frac{1}{2}\,(\mathbf{z}-\tilde{\mathbf{y}})^T \mathbf{K}(\mathbf{z}-\tilde{\mathbf{y}}) + \frac{1}{2}\,\mathbf{v}^T\mathbf{G}\mathbf{v} + (\mathbf{z}-\tilde{\mathbf{y}})^T\mathbf{g}. \tag{2.3}$$

The block vectors $\mathbf{z} := [z_1^T, \ldots, z_{\hat{l}}^T]^T \in \mathbb{R}^{\hat{l}\hat{q}}$ and $\mathbf{v} := [v_1^T, \ldots, v_{\hat{l}}^T]^T \in \mathbb{R}^{\hat{l}\hat{p}}$ denote the state and control variables, respectively, at all the discrete times. The discrete target is $\tilde{\mathbf{y}} := [\tilde{y}_1^T, \ldots, \tilde{y}_{\hat{l}}^T]^T \in \mathbb{R}^{\hat{l}\hat{q}}$ with target error $e_l = (z_l - \tilde{y}_l^h)$ for $0 \le l \le \hat{l}$. The matrix $\mathbf{K} := \mathbf{Z} + \mathbf{\Gamma}$ with $\mathbf{Z}, \mathbf{\Gamma} \in \mathbb{R}^{(\hat{l}\hat{q})\times(\hat{l}\hat{q})}$. The matrix $\mathbf{\Gamma} = s\,diag(0,0,...,M_h)$ and $\mathbf{Z} = D_\tau \otimes M_h$, $D_\tau \in \mathbb{R}^{\hat{l}\times\hat{l}}$ with entries $(D_\tau)_{ij} := \int_{t_o}^{t_f} \vartheta_i(t)\vartheta_j(t)dt$, for $1 \le i,j \le \hat{l}$, where $\otimes$ stands for the Kronecker product. The matrix $\mathbf{G} = r\tau I_{\hat{l}} \otimes R_h \in \mathbb{R}^{(\hat{l}\hat{p})\times(\hat{l}\hat{p})}$, and $I_{\hat{l}} \in \mathbb{R}^{\hat{l}\times\hat{l}}$ is an identity matrix. The vector $\mathbf{g} = (g_1^T, 0^T, \ldots, 0^T)^T$, where $g_1 = \frac{\tau}{6} M_h e_0$. Note that g_1 does not necessarily vanish because it is not assumed that $y_0^h = \tilde{y}_0^h$.

Employing the θ-scheme discretization in time, the equation (2.2) takes the form:

$$F_1 \underline{z}_{i+1} = F_0\,\underline{z}_i + \tau B_h \underline{v}_i, \quad \text{for} \quad t_o < t < t_f; \quad \text{and} \quad \underline{z}(t_o) = y_o^h. \tag{2.4}$$

where $F_0, F_1 \in \mathbb{R}^{\hat{q}\times\hat{q}}$ are (fixed) matrices given by $F_0 := M_h - (1-\theta)\tau A_h$ and $F_1 := M_h + \theta\tau A_h$. Using a full discretization in time, equation (2.2) has matrix form:

$$\mathbf{E}\,\mathbf{z} + \mathbf{N}\,\mathbf{v} = \mathbf{f}, \tag{2.5}$$

where the input vector is $\mathbf{f} := [(F_0 y_0^h)^T, 0^T, ..., 0^T]^T \in \mathbb{R}^{\hat{l}\hat{q}}$. The block lower bidiagonal matrix $\mathbf{E} \in \mathbb{R}^{(\hat{l}\hat{q})\times(\hat{l}\hat{q})}$ is given by

$$\mathbf{E} = \begin{bmatrix} F_1 & & & \\ -F_0 & F_1 & & \\ & \ddots & \ddots & \\ & & -F_0 & F_1 \end{bmatrix}, \tag{2.6}$$

The block diagonal matrix $\mathbf{N} \in \mathbb{R}^{(\hat{l}\hat{q})\times(\hat{l}\hat{p})}$ is given by $\mathbf{N} = -\tau I_{\hat{l}} \otimes B_h$. The Lagrangian functional $\mathcal{L}_h^\tau(\mathbf{z}, \mathbf{v}, \mathbf{q})$ for minimizing (2.3) subject to constraint (2.5) is:

$$\mathcal{L}_h^\tau(\mathbf{z}, \mathbf{v}, \mathbf{q}) = J_h^\tau(\mathbf{z}, \mathbf{v}) + \mathbf{q}^T(\mathbf{E}\mathbf{z} + \mathbf{N}\mathbf{v} - \mathbf{f}). \tag{2.7}$$

To obtain a discrete saddle point formulation of (2.7), we apply optimality conditions for $\mathcal{L}_h^\tau(\cdot,\cdot,\cdot)$. This yields the symmetric indefinite linear system:

$$\begin{bmatrix} \mathbf{K} & \mathbf{0} & \mathbf{E}^T \\ \mathbf{0} & \mathbf{G} & \mathbf{N}^T \\ \mathbf{E} & \mathbf{N} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{K}\tilde{\mathbf{y}} - \mathbf{g} \\ \mathbf{0} \\ \mathbf{f} \end{bmatrix}, \tag{2.8}$$

where $\tilde{\mathbf{y}} := [(\tilde{y}_1^h)^T, \ldots, (\tilde{y}_{\hat{l}}^h)^T]^T \in \mathbb{R}^{\hat{l}\hat{q}}$. Next, we study the condition number of $\mathbf{E}\mathbf{E}^T$, where $\mathbf{E}$ is the evolution matrix.

THEOREM 2.1. *Let A_h be a $\hat{q}\times\hat{q}$ symmetric positive definite matrix. Let λ_q for $1 \le q \le \hat{q}$ denote the generalized eigenvalues of matrix A_h with respect to M_h. Let the evolution matrix $\mathbf{E}$ be defined by* (2.5) *with matrices F_0 and F_1 defined as:*

$$F_0 := M_h - (1-\theta)\tau A_h \ \ \text{and} \ \ F_1 := M_h + \theta\tau A_h, \tag{2.9}$$

respectively. Then, for $\theta \ge \frac{1}{2}$, the scheme (2.4) *will be stable for all $\tau > 0$, while for $\theta < \frac{1}{2}$, it will be stable only if $\tau \le 2/\left((1-2\theta)\rho_{max}\right)$. The following bound:*

$$\text{cond}(\mathbf{E}\mathbf{E}^T) \le \frac{4\left(1+\theta\tau\rho_{\max}\right)^2}{(\tau\rho_{\min})^2} \tag{2.10}$$

will hold, where $\rho_{\max} := \max_q |\lambda_q|$ and $\rho_{\min} := \min_q |\lambda_q|$.

Proof. *Part 1* (stability condition). Consider the θ-scheme for equation (2.2) given by:

$$\underline{z}_{l+1} = \Phi\, \underline{z}_l + \tau\, F_1^{-1}\, B_h\, \underline{v}_l, \tag{2.11}$$

where $\Phi \in \mathbb{R}^{\hat{q}\times\hat{q}}$ is the marching matrix given by

$$\Phi := (M_h + \theta\tau A_h)^{-1}(M_h - (1-\theta)\tau A_h). \tag{2.12}$$

Consider $V_h := [v_1, ..., v_{\hat{q}}]$ and $\Lambda_h := diag\{\lambda_1, ..., \lambda_{\hat{q}}\}$ as the generalized eigenvectors and eigenvalues of A_h with respect to M_h, i.e., $A_h V_h = M_h V_h \Lambda_h$, with $V_h^{-1} M_h V_h = I$. Then the stability condition for (2.11) is given by

$$|\,(1+\theta\tau\lambda_q)^{-1}(1-(1-\theta)\tau\lambda_q)\,| \le 1\,, \tag{2.13}$$

or equivalently,

$$\begin{cases} \ \ 1-(1-\theta)\tau\lambda_q & \le & 1+\theta\tau\lambda_q \\ -1+(1-\theta)\tau\lambda_q & \le & 1+\theta\tau\lambda_q. \end{cases} \tag{2.14}$$

From (2.14), we obtain $0 \le \tau\lambda_q$ and $\tau(1-2\theta)|\lambda_q| \le 2$ since $0 < \lambda_q$. In the case $\theta \ge 1/2$, there is no restriction on τ, consequently the marching scheme is unconditionally stable. On the other hand, if $\theta < 1/2$ then $0 < (1-2\theta)$ and in order for the scheme to be stable it is necessary that $\tau \le 2/((1-2\theta)\rho_{\max})$, where $\rho_{\max} = \max_q |\lambda_q|$. In this case, the marching scheme is conditionally stable.

Part 2 (estimation of cond($\mathbf{E}\mathbf{E}^T$)). We shall diagonalize the blocks of $\mathbf{E}\mathbf{E}^T$:

$$\mathbf{E}\mathbf{E}^T = \begin{bmatrix} F_1F_1^T & -F_1F_0^T & & & \\ -F_0F_1^T & F_0F_0^T + F_1F_1^T & -F_1F_0^T & & \\ & -F_0F_1^T & F_0F_0^T + F_1F_1^T & -F_1F_0^T & \\ & & \ddots & \ddots & \ddots \\ & & & -F_0F_1^T & F_0F_0^T + F_1F_1^T \end{bmatrix}. \tag{2.15}$$

Notice that $V_h^{-1}M_hV_h = I$ and $V_h^{-1}M_h^{-1}A_hV_h = \Lambda_h$ imply that F_0 and F_1 are diagonalized by V_h, yielding $\Lambda_0 = V_h^{-1}F_0V_h = V_h^{-1}(M_h - (1-\theta)\tau A_h)V_h$ and matrix $\Lambda_1 = V_h^{-1}F_1V_h = V_h^{-1}(M_h + \theta\tau A_h)V_h$. If $\boldsymbol{\mathcal{V}}_h := \text{blockdiag}(V_h, \ldots, V_h)$, then the block matrix $\boldsymbol{\mathcal{V}}_h^{-1}\mathbf{E}\mathbf{E}^T\boldsymbol{\mathcal{V}}_h$ will have blocks which are diagonal matrices:

$$\boldsymbol{\mathcal{V}}_h^{-1}\mathbf{E}\mathbf{E}^T\boldsymbol{\mathcal{V}}_h = \begin{bmatrix} \Lambda_1^2 & -\Lambda_0\Lambda_1 & & & \\ -\Lambda_0\Lambda_1 & \Lambda_0^2 + \Lambda_1^2 & -\Lambda_1\Lambda_0 & & \\ & -\Lambda_0\Lambda_1 & \Lambda_0^2 + \Lambda_1^2 & -\Lambda_1\Lambda_0 & \\ & & \ddots & \ddots & \ddots \\ & & & -\Lambda_0\Lambda_1 & \Lambda_0^2 + \Lambda_1^2 \end{bmatrix}. \tag{2.16}$$

Next, we permute the rows and columns of the block tridiagonal matrix (2.16) using a permutation matrix $\mathbf{\Pi}$, so that $\mathbf{\Theta} := \mathbf{\Pi}(\boldsymbol{\mathcal{V}}_h^{-1}\mathbf{E}\mathbf{E}^T\boldsymbol{\mathcal{V}}_h)\mathbf{\Pi}^T = \text{blockdiag}(\Theta_1, \ldots, \Theta_{\hat{l}})$ where each block submatrix Θ_l is a tridiagonal matrix with entries:

$$\Theta_l := (\mathbf{\Pi}(\boldsymbol{\mathcal{V}}_h^{-1}\mathbf{E}\mathbf{E}^T\boldsymbol{\mathcal{V}}_h)\mathbf{\Pi}^T)_l = \begin{bmatrix} a_q^2 & -a_qb_q & & & \\ -a_qb_q & a_q^2 + b_q^2 & -a_qb_q & & \\ & -a_qb_q & a_q^2 + b_q^2 & -a_qb_q & \\ & & \ddots & \ddots & \ddots \\ & & & -a_qb_q & a_q^2 + b_q^2 \end{bmatrix}, \tag{2.17}$$

where $b_q := (1 - (1-\theta)\tau\lambda_q)$ and $a_q := (1 + \theta\tau\lambda_q)$. Let $\mu(\Theta_l)$ denote an eigenvalue of submatrix Θ_l (and hence also of $\mathbf{\Theta}$). Then, Gershgorin's Theorem [7, 28] yields:

$$|\ \mu(\Theta_l) - a_q^2\ | \le |\ a_qb_q\ | \quad \text{or} \quad |\ \mu(\Theta_l) - a_q^2 - b_q^2\ | \le 2\ |\ a_qb_q\ | \tag{2.18}$$

Using condition (2.13), we guarantee stability when $|\ b_q\ | \le |\ a_q\ |$ obtaining

$$\mu(\Theta_l) \le \max\left(|\ a_q\ |\ (|\ a_q\ | + |\ b_q\ |)\,, (|\ a_q\ | + |\ b_q\ |)^2\right) \le \max 4\ |\ a_q\ |^2 \tag{2.19}$$

and

$$\mu(\Theta_l) \ge \min\left((|\ a_q\ |^2 - |\ a_q\ ||\ b_q\ |)\,, (|\ a_q\ | - |\ b_q\ |)^2\right) \ge \min\,(|\ a_q\ | - |\ b_q\ |)^2 \tag{2.20}$$

An upper bound for $\mu(\Theta_l)$ from (2.19) is given by $\mu(\Theta_l) \le 4\,(1 + \theta\tau\rho_{\max})^2$, where $\rho_{\max} = \max_q |\lambda_q|$. To obtain a lower bound for $\mu(\Theta_l)$, from (2.20) we employ the notation $\rho_{\min} := \min_q |\lambda_q|$ obtaining $\mu(\Theta_l) \geq (\tau\rho_{\min})^2$. Therefore, the condition number of the matrix $\mathbf{E}\mathbf{E}^T$ will satisfy the bound:

$$\mathrm{cond}(\mathbf{E}\mathbf{E}^T) \leq 4\left(\frac{1+\theta\tau\rho_{\max}}{\tau\rho_{\min}}\right)^2. \tag{2.21}$$

This completes the proof. □

Remark 1. For finite element discretizations on a polygonal domain of size $O(1)$, the generalized eigenvalues λ_q will satisfy the bounds $\alpha_1 \leq |\lambda_q| \leq \alpha_2 h^{-2}$. Then, using (2.21) we obtain:

$$\mathrm{cond}(\mathbf{E}\mathbf{E}^T) \approx \left(\frac{1+\theta\tau\alpha_2 h^{-2}}{\tau\alpha_1}\right)^2. \tag{2.22}$$

Thus, the matrix $\mathbf{E}\mathbf{E}^T$ will be ill-conditioned with a condition number that grows as $O(h^{-4})$ depending on τ and h. If the system (2.8) is solved using the Uzawa's method, it is necessary to solve $-(\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T + \mathbf{N}\mathbf{G}^{-1}\mathbf{N}^T)\mathbf{p} = \mathbf{f} - \mathbf{E}\hat{\mathbf{y}} + \mathbf{E}\mathbf{K}^{-1}\mathbf{g}$, where the matrix $\mathbf{S} := (\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T + \mathbf{N}\mathbf{G}^{-1}\mathbf{N}^T)$ is the Schur complement of the system (2.8) with respect to the Lagrange multiplier $\mathbf{p}$.

Next, we analyze the condition number of the matrix $\mathbf{S}$. Notice that due to the positive semi-definiteness of matrix M_h in (2.1), we obtain in the sense of quadratic forms that $\mathbf{K}^{-1} = (\mathbf{Z}+\mathbf{\Gamma})^{-\mathbf{1}} \leq \mathbf{Z}^{-\mathbf{1}}$ and apply it in the following estimate for the condition number of the Schur complement $\mathbf{S}$.

LEMMA 2.2. *Let the upper and lower bound for the singular values of $\mathbf{E}\mathbf{E}^T$ be given by $4(1+\theta\tau\rho_{max})^2$ and $(\tau\rho_{min})^2$, respectively. Let $\mathbf{K} = \mathbf{Z}+\mathbf{\Gamma}$, and suppose that the mass matrices $\mathbf{Z}$, $\mathbf{G}$, $\mathbf{N}$, and $\mathbf{\Gamma}$ satisfy:*

$$c_1\,\tau\,\mathbf{y}^T\mathbf{y} \leq \quad \mathbf{y}^T\mathbf{Z}\mathbf{y} \quad \leq c_2\,\tau\,\mathbf{y}^T\mathbf{y} \tag{2.23}$$
$$c_3\,r\,\tau\,h^d\mathbf{u}^T\mathbf{u} \leq \quad \mathbf{u}^T\mathbf{G}\mathbf{u} \quad \leq c_4\,r\,\tau\,h^d\mathbf{u}^T\mathbf{u}, \tag{2.24}$$
$$c_5\,\tau^2\,h^d\,\mathbf{p}^T\mathbf{p} \leq \mathbf{p}^T\mathbf{N}\mathbf{N}^T\mathbf{p} \leq c_6\,\tau^2\,h^d\,\mathbf{p}^T\mathbf{p} \quad \textit{and} \tag{2.25}$$
$$0 \leq \quad \mathbf{y}^T\mathbf{\Gamma}\mathbf{y} \quad \leq c_7\,s\,\mathbf{y}^T\mathbf{y}, \tag{2.26}$$

where the constants c_i are independent of r, s, h and τ. Then, the condition number of matrix $\mathbf{S}$ is bounded by:

$$\mathrm{cond}(\mathbf{S}) \leq \left(\frac{c_4\,r\,(c_5\,\tau + c_7\,s)}{c_1\,\tau\,c_3\,r}\right)\left(\frac{4\,c_3\,r\,(1+\rho_{\max}\ \tau\,\theta)^2 + c_6\,\tau^2\,c_1}{c_4\,r\,(\tau\,\rho_{\min})^2 + c_5\,\tau\,(c_2\,\tau + c_7\,s)}\right), \tag{2.27}$$

where $\mathbf{S} := \mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T + \mathbf{N}\mathbf{G}^{-1}\mathbf{N}^T$ denotes the Schur complement.

Proof. Using the upper and lower bounds for $\mathbf{K}$, $\mathbf{E}\mathbf{E}^T$, $\mathbf{N}\mathbf{N}^T$ and $\mathbf{G}$ we obtain: *Upper bound*:

$$\mathbf{p}^T\mathbf{S}\mathbf{p} = \mathbf{p}^T\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T\mathbf{p} + \mathbf{p}^T\mathbf{N}\mathbf{G}^{-1}\mathbf{N}^T\mathbf{p} \tag{2.28}$$
$$\leq \mathbf{p}^T\mathbf{E}\mathbf{Z}^{-1}\mathbf{E}^T\mathbf{p} + \mathbf{p}^T\mathbf{N}\mathbf{G}^{-1}\mathbf{N}^T\mathbf{p} \tag{2.29}$$
$$\leq \frac{1}{c_1\,\tau}\mathbf{p}^T\mathbf{E}\mathbf{E}^T\mathbf{p} + \frac{1}{c_3\,r\,\tau\,h^d}\mathbf{p}^T\mathbf{N}\mathbf{N}^T\mathbf{p} \tag{2.30}$$
$$\leq \left(\frac{4}{c_1\,\tau}(1+\tau\,\theta\,\rho_{\max})^2 + \frac{c_6\,\tau^2 h^d}{c_3\,r\,\tau\,h^d}\right)\mathbf{p}^T\mathbf{p} \tag{2.31}$$
$$= \left(\frac{4}{c_1\,\tau}(1+\tau\,\theta\,\rho_{\max})^2 + \frac{c_6\,\tau}{c_3\,r}\right)\mathbf{p}^T\mathbf{p}. \tag{2.32}$$

Lower bound:

$$\mathbf{p}^T\mathbf{S}\mathbf{p} \geq \frac{1}{(c_2\,\tau + c_7)}\mathbf{p}^T\mathbf{E}\mathbf{E}^T\mathbf{p} + \frac{1}{c_4\,r\,\tau\,h^d}\mathbf{p}^T\mathbf{N}\mathbf{N}^T\mathbf{p} \tag{2.33}$$

$$\geq \left(\frac{(\tau\,\rho_{\min})^2}{(c_2\,\tau + c_7\,s)} + \frac{c_5\,\tau^2\,h^d}{c_4\,r\,\tau\,h^d}\right)\mathbf{p}^T\mathbf{p} \tag{2.34}$$

$$= \left(\frac{(\tau\,\rho_{\min})^2}{(c_2\,\tau + c_7\,s)} + \frac{c_5\,\tau}{c_4\,r}\right)\mathbf{p}^T\mathbf{p}. \tag{2.35}$$

Therefore, the condition number of matrix $\mathbf{S}$ can be estimated by:

$$\mathrm{cond}(\mathbf{S}) \leq \frac{c_4\,r\,(c_5\,\tau + c_7\,s)}{c_1\,\tau\,c_3\,r}\,\frac{4\,c_3\,r\,(1+\theta\tau\rho_{\max}\,)^2 + c_6\,\tau^2\,c_1}{c_4\,r\,(\tau\,\rho_{\min})^2 + c_5\,\tau\,(c_2\,\tau + c_7\,s)}. \tag{2.36}$$

□

Remark 2. The estimate given in (2.36) shows that matrix $\mathbf{S}$ is ill-conditioned. Indeed, using that $c_i = O(1)$, the expression (2.36) reduces to:

$$\mathrm{cond}(\mathbf{S}) \approx \left(\frac{\tau + s}{\tau}\right)\left(\frac{r\,(1+\theta\tau h^{-2})^2 + \tau^2}{r\tau^2 + \tau^2 + s\tau}\right). \tag{2.37}$$

Choosing $\theta = 1$ and using the reasonable assumptions:

$$ch^2 \leq \tau,\; r \leq O(1)$$

yields that $\mathrm{cond}(\mathbf{S}) \approx O(rh^{-4})$.

3. The reduced system for u. We shall now describe an algorithm to solve the saddle point system (2.8) based on the solution of a reduced Schur complement system for the control variable $\mathbf{u}$. Solving the first and third block rows in (2.8) yield $\mathbf{p} = -\mathbf{E}^{-T}\mathbf{K}\mathbf{y} + \mathbf{E}^{-T}\mathbf{K}\tilde{\mathbf{y}} - \mathbf{E}^T\mathbf{g}$ and $\mathbf{y} = -\mathbf{E}^{-1}\mathbf{N}\mathbf{u} + \mathbf{E}^{-1}\mathbf{f}$, respectively. System (2.8) can then be reduced to the following Schur complement system for $\mathbf{u}$:

$$(\mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} = \mathbf{b} - \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{g}, \tag{3.1}$$

where $\mathbf{b} := \mathbf{N}^T\mathbf{E}^{-T}\left(\mathbf{K}\mathbf{E}^{-1}\mathbf{f} - \mathbf{K}\tilde{\mathbf{y}} + \mathbf{g}\right)$. If $\mathbf{G} > 0$, then $(\mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})$ is symmetric and positive definite. In the next theorem, we show that matrix $\mathbf{G}$ is spectrally equivalent to the matrix $(\mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})$.

THEOREM 3.1. *Let the matrices* $\mathbf{Z}$, $\mathbf{G}$, $\mathbf{N}$ *and* $\mathbf{\Gamma}$ *satisfy the bounds described in* (2.23) − (2.26). *Let* $\mathbf{K} = \mathbf{Z} + \mathbf{\Gamma}$. *Let the matrix* $\mathbf{E}$ *be such that the upper and lower bounds for* $\mathbf{E}\mathbf{E}^T$ *are given by* $4(1+\theta\tau\rho_{\max})^2$ *and* $(\tau\rho_{\min})^2$, *respectively. Then, there exist* $\mu_{\min} > 0$ *and* $\mu_{\max} > 0$, *independent of* h *and* $\mathbf{u}$, *such that:*

$$\mu_{\min}\,\left(\mathbf{u}^T\mathbf{G}\mathbf{u}\right) \leq \mathbf{u}^T(\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} \leq \mu_{\max}\,\left(\mathbf{u}^T\mathbf{G}\mathbf{u}\right) \tag{3.2}$$

and

$$\mathbf{u}^T\mathbf{G}\mathbf{u} \leq \mathbf{u}^T(\mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} \leq (1+\mu_{\max})\,\mathbf{u}^T\mathbf{G}\mathbf{u}. \tag{3.3}$$

where $\mu_{\max} = \frac{(c_2\,\tau + c_7\,s)\,c_6}{(\rho_{min})^2\,c_3\,r\,\tau}$ *and* $\mu_{\min} = \frac{c_1\,c_5\,\tau^2}{4\,(1+\tau\,\rho_{max}\,\theta)^2\,c_4\,r}$.

Proof. We first prove expression (3.2). Using the upper and lower bounds for $\mathbf{K}$, $\mathbf{E}\mathbf{E}^T$, $\mathbf{N}\mathbf{N}^T$ and $\mathbf{G}$ we obtain:
Upper bound:

$$\begin{aligned}
\mathbf{u}^T\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}\mathbf{u} &\leq (c_2\,\tau + c_7\,s)\,\mathbf{u}^T\mathbf{N}^T\mathbf{E}^{-T}\mathbf{E}^{-1}\mathbf{N}\mathbf{u} && (3.4)\\
&\leq \frac{(c_2\,\tau + c_7\,s)}{(\tau\,\rho_{\min})^2}\,\mathbf{u}^T\mathbf{N}^T\mathbf{N}\mathbf{u} && (3.5)\\
&\leq \frac{(c_2\,\tau + c_7\,s)\,c_6\,\tau^2\,h^d}{(\tau\,\rho_{\min})^2}\,\mathbf{u}^T\mathbf{u} && (3.6)\\
&= \frac{(c_2\,\tau + c_7\,s)c_6\,h^d}{(\rho_{\min})^2}\,\mathbf{u}^T\mathbf{u} && (3.7)\\
&\leq \frac{(c_2\,\tau + c_7\,s)\,c_6}{(\rho_{\min})^2\,c_3\,r\,\tau}\,\mathbf{u}^T\mathbf{G}\mathbf{u} && (3.8)\\
&= \mu_{\max}\,\mathbf{u}^T\mathbf{G}\mathbf{u}. && (3.9)
\end{aligned}$$

Lower bound:

$$\begin{aligned}
\mathbf{u}^T\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}\mathbf{u} &\geq (c_1\,\tau)\,\mathbf{u}^T\mathbf{N}^T\mathbf{E}^{-T}\mathbf{E}^{-1}\mathbf{N}\mathbf{u} && (3.10)\\
&\geq \frac{c_1\,\tau}{4\,(1+\tau\,\rho_{\max}\,\theta)^2}\,\mathbf{u}^T\mathbf{N}^T\mathbf{N}\mathbf{u} && (3.11)\\
&\geq \frac{c_1\,c_5\,\tau^3\,h^d}{4\,(1+\tau\,\rho_{\max}\,\theta)^2}\,\mathbf{u}^T\mathbf{u} && (3.12)\\
&\geq \frac{c_1\,c_5\,\tau^2}{4\,(1+\tau\,\rho_{\max}\,\theta)^2\,c_4\,r}\,\mathbf{u}^T\mathbf{G}\mathbf{u} && (3.13)\\
&= \mu_{\min}\,\mathbf{u}^T\mathbf{G}\mathbf{u}. && (3.14)
\end{aligned}$$

Since $\mu_{\min} > 0$ and $\mu_{\max} > 0$ then $\mathbf{u}^T(\mathbf{G}+\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} \leq (1+\mu_{\max})\,\mathbf{u}^T\mathbf{G}\mathbf{u}$ and $\mathbf{u}^T\mathbf{G}\mathbf{u} \leq (1+\mu_{\min})\,\mathbf{u}^T\mathbf{G}\mathbf{u} \leq \mathbf{u}^T(\mathbf{G}+\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u}$. Therefore:

$$\mathbf{u}^T\mathbf{G}\mathbf{u} \leq \mathbf{u}^T(\mathbf{G}+\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} \leq (1+\mu_{\max})\,\mathbf{u}^T\mathbf{G}\mathbf{u}. \qquad (3.15)$$

This completes the proof. $\square$

First Algorithm. The Schur complement system (3.1) can be solved using a CG (conjugate gradient) algorithm using $\mathbf{G}$ as a preconditioner. Since $\rho_{\min}$ is $O(1)$ and $\rho_{\max}$ is $O(h^{-2})$, it is easy to see that, in Theorem 3.1:

$$\mu_{\min} = O\left(\frac{h^4}{r}\right) \text{ and } \mu_{\max} = O\left(\frac{1+s/\tau}{r}\right). \qquad (3.16)$$

Hence, the rate of convergence of this algorithm will be independent of h, with a condition number bound:

$$\operatorname{cond}\left(\mathbf{G}^{-1}\left(\mathbf{G}+\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}\right)\right) \leq O\left(1+\frac{1+\frac{s}{\tau}}{r}\right). \qquad (3.17)$$

This algorithm is simple to implement however has two drawbacks. It has *inner* and *outer* iterations, and requires applications of $\mathbf{E}^{-1}$ (and $\mathbf{E}^{-T}$) which are not directly parallelizable.

Second Algorithm. Our second algorithm avoids double iterations. Define:

$$\mathbf{w} := -\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}\mathbf{u}.$$

Then, the solution to system (3.1) can be obtained by solving the system:

$$\mathbf{H}\begin{bmatrix}\mathbf{w}\\ \mathbf{u}\end{bmatrix} := \begin{bmatrix}\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T & \mathbf{N}\\ \mathbf{N}^T & -\mathbf{G}\end{bmatrix}\begin{bmatrix}\mathbf{w}\\ \mathbf{u}\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{b}\end{bmatrix}, \tag{3.18}$$

which is symmetric and *indefinite*. The action of $\mathbf{E}^{-1}$ is required only in a pre-computed step to assemble the right hand side input vector $\mathbf{b}$.

Since system (3.18) is symmetric indefinite, it can be solved iteratively using the MINRES algorithm with a positive definite *preconditioner* $\mathbf{P} := \text{blockdiag}(\widetilde{\mathbf{E}}\widetilde{\mathbf{K}}^{-1}\widetilde{\mathbf{E}}^T, \widetilde{\mathbf{G}})$, where the matrix $\widetilde{\mathbf{E}}^{-T}\widetilde{\mathbf{K}}\widetilde{\mathbf{E}}^{-1}$ is required to be spectrally equivalent to $\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}$ and $\widetilde{\mathbf{G}}$ is a preconditioner for the matrix $\mathbf{G}$. The following theorem estimates the condition number of $\mathbf{P}^{-1}\mathbf{H}$ when $\widetilde{\mathbf{E}} = \mathbf{E}$, $\widetilde{\mathbf{K}} = \mathbf{K}$ and $\widetilde{\mathbf{G}} = \mathbf{G}$. In the next section we consider $\widetilde{\mathbf{E}} \neq \mathbf{E}$ and use the parareal algorithm to establish an approximation $\mathbf{E}_n$ to $\mathbf{E}$ (see Theorem 4.1). The case $\widetilde{\mathbf{K}} \neq \mathbf{K}$ follows from Remark 6.

THEOREM 3.2. *Let the matrices* $\mathbf{Z}$, $\mathbf{G}$, $\mathbf{N}$, $\mathbf{\Gamma}$ *and* $\mathbf{E}$ *with bounds described in the Lemma* 2.2. *Let* $\mathbf{P} := \textit{blockdiag}(\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T, \mathbf{G})$ *denote a block diagonal preconditioner for the coefficient matrix* $\mathbf{H}$ *of system* (3.18). *Then, the condition number of the preconditioned system satisfies the bound:*

$$\text{cond}(\mathbf{P}^{-1}\mathbf{H}) \leq O\left(\left(1 + \frac{1 + s/\tau}{r}\right)^{1/2}\right). \tag{3.19}$$

Proof. Since the preconditioner $\mathbf{P}$ is positive definite, we consider the generalized eigenvalue problem given by:

$$\begin{bmatrix}\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T & \mathbf{N}\\ \mathbf{N}^T & -\mathbf{G}\end{bmatrix}\begin{bmatrix}\mathbf{w}\\ \mathbf{u}\end{bmatrix} = \eta\begin{bmatrix}\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T & \\ & \mathbf{G}\end{bmatrix}\begin{bmatrix}\mathbf{w}\\ \mathbf{u}\end{bmatrix}, \tag{3.20}$$

We obtain the equations:

$$(\eta - 1)\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T\mathbf{w} = \mathbf{N}\mathbf{u} \text{ and } (\eta + 1)\mathbf{G}\mathbf{u} = \mathbf{N}^T\mathbf{w}. \tag{3.21}$$

These equations yield $\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}\mathbf{u} = (\eta^2 - 1)\mathbf{G}\mathbf{u}$, where $(\eta^2 - 1)$ is the generalized eigenvalue of $\mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}$ with respect to $\mathbf{G}$. Using Theorem 3.1, we obtain bounds for η as follows:

$$\max|\eta| \leq (1 + \eta_{\max})^{1/2} = O\left(\left(1 + \frac{1 + s/\tau}{r}\right)^{1/2}\right) \tag{3.22}$$

$$\min|\eta| \geq (1 + \eta_{\min})^{1/2} = O(1). \tag{3.23}$$

The desired result now follows, since:

$$\text{cond}(\mathbf{P}^{-1}\mathbf{H}) \leq \frac{\max|\eta|}{\min|\eta|}. \tag{3.24}$$

This completes the proof. □

Remark 3. Applying the matrix $\mathbf{E}$ to a vector is highly unstable, but applying $\mathbf{E}^{-1}$ is stable. The algorithms presented here do not require application of $\mathbf{E}$ or $\mathbf{E}^T$ since:

$$\mathbf{P}^{-1}\mathbf{H} = \begin{bmatrix}\mathbf{I} & \mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}\\ \mathbf{G}^{-1}\mathbf{N}^T & -\mathbf{I}\end{bmatrix}. \tag{3.25}$$

4. Parareal approximation $\mathbf{E}_n^{-T}\widehat{\mathbf{K}}\mathbf{E}_n^{-1}$. The parareal method is an iterative method for solving a parabolic equation based on a *decomposition* of its *temporal* domain $[t_o, t_f]$ into $\hat{k}$ *coarse* sub-intervals of length $\Delta T = (t_f - t_o)/\hat{k}$, setting $T_0 = t_o$ and $T_k = t_o + k\Delta T$ for $1 \le k \le \hat{k}$, see [17]. It determines the solution at the times T_k for $1 \le k \le \hat{k}$ by using a *multiple-shooting* technique which requires solving the parabolic equation on each interval (T_{k-1}, T_k) in *parallel*. To speed up the multiple shooting iteration, the residual equations are "preconditioned" by solving a "coarse" time-grid discretization of the parabolic equation using the time step ΔT.

In this section we formulate a preconditioner $\mathbf{E}_n$ for $\mathbf{E}$ based on n Richardson iterations of the parareal algorithm. Using $\mathbf{E}_n$, an application of $\mathbf{E}_n^{-T}\widehat{\mathbf{K}}\mathbf{E}_n^{-1}$ to a vector $\mathbf{s} \in \mathbb{R}^{(\hat{l}\hat{q})\times(\hat{l}\hat{q})}$ can be computed in three steps. Step 1, apply $\mathbf{E}_n^{-1}\mathbf{s} :\rightarrow \hat{\mathbf{z}}^n$ using n applications of the parareal method (described below). Step 2, multiply $\widehat{\mathbf{K}}\hat{\mathbf{z}}^n :\rightarrow \hat{\mathbf{t}}^n$ (see below). And Step 3, apply $\mathbf{E}_n^{-T}\hat{\mathbf{t}}^n :\rightarrow \mathbf{x}^n$, i.e., the transpose of Step 1.

To describe $\mathbf{E}_n$, we define fine and coarse propagators F and G as follows. The local solution at T_k is defined by marching from T_{k-1} to T_k the θ-scheme on the fine triangulation τ with an initial data Z_{k-1} at T_{k-1}. Let $\hat{m} = (T_k - T_{k-1})/\tau$ and $j_{k-1} = \frac{T_{k-1}-T_0}{\tau}$. It is easy to see that:

$$F_1\, Z_k = F\, Z_{k-1} + S_k, \tag{4.1}$$

where $F := (F_0F_1^{-1})^{\hat{m}-1}F_0 \in \mathbb{R}^{\hat{q}\times\hat{q}}$, $S_k := \sum_{m=1}^{\hat{m}} \left(F_0F_1^{-1}\right)^{\hat{m}-m} s_{j_{k-1}+m}$, $Z_0 = 0$. Imposing continuity $F_1\, Z_k - F\, Z_{k-1} - S_k = 0$ at time T_k, for $1 \le k \le \hat{k}$, yields:

$$\mathbf{C}\,\mathbf{Z} := \begin{bmatrix} F_1 & & & \\ -F & F_1 & & \\ & \ddots & \ddots & \\ & & -F & F_1 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{\hat{k}} \end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_{\hat{k}} \end{bmatrix} := \mathbf{S}. \tag{4.2}$$

The coarse solution at T_k with initial data $Z_{k-1} \in \mathbb{R}^{\hat{q}}$ at T_{k-1} is obtained by applying one coarse time step of the backward Euler method $F_1^{\Delta} Z_k = G\, Z_{k-1}$ where matrix $F_1^{\Delta} := (M_h + A_h\Delta T)$ and $G := M_h \in \mathbb{R}^{\hat{q}\times\hat{q}}$. In the parareal algorithm, the following coarse propagator based on G is employed to precondition system (4.2) via:

$$\begin{bmatrix} Z_1^{i+1} \\ Z_2^{i+1} \\ \vdots \\ Z_{\hat{k}}^{n+1} \end{bmatrix} = \begin{bmatrix} Z_1^{i} \\ Z_2^{i} \\ \vdots \\ Z_{\hat{k}}^{i} \end{bmatrix} + \left(\begin{bmatrix} F_1^{\Delta} & & & \\ -G & F_1^{\Delta} & & \\ & \ddots & \ddots & \\ & & -G & F_1^{\Delta} \end{bmatrix} \right)^{-1} \begin{bmatrix} R_1^{i} \\ R_2^{i} \\ \vdots \\ R_{\hat{k}}^{i} \end{bmatrix}, \tag{4.3}$$

for $0 \le i \le n-1$, where the residual $\mathbf{R}^i := [{R_1^i}^T, ..., {R_{\hat{k}}^i}^T]^T \in \mathbb{R}^{\hat{k}\hat{q}}$ in (4.2) is defined as $\mathbf{R}^i := \mathbf{S} - \mathbf{C}\,\mathbf{Z}^i$, where $\mathbf{Z}^i := [{Z_1^i}^T, ..., {Z_{\hat{k}}^i}^T]^T \in \mathbb{R}^{\hat{k}\hat{q}}$, and $\mathbf{Z}^0 := [0^T, ..., 0^T]^T$.

We can now define $\hat{\mathbf{z}}^n := \mathbf{E}_n^{-1}\mathbf{s}$. Let $\hat{\mathbf{z}}^n$ be the nodal representation of a piecewise linear function $\underline{\hat{z}}^n$ in time with respect to the fine triangulation τ on $[t_o, t_f]$, and continuous inside of each coarse sub-interval $[T_{k-1}, T_k]$, i.e., the function $\underline{\hat{z}}^n$ can be discontinuous across the coarse points T_k, $1 \le k \le \hat{k}-1$, therefore, $\hat{\mathbf{z}}^n \in \mathbb{R}^{(\hat{l}+\hat{k}-1)\hat{q}}$. On each sub-interval $[T_{k-1}, T_k]$, $\underline{\hat{z}}^n$ is defined marching from T_{k-1} to T_k the θ-scheme using fine time steps τ and initial data Z_{k-1}^n at T_{k-1}.

The matrix matrix $\widehat{\mathbf{K}} := \widehat{\mathbf{Z}} + \widehat{\mathbf{\Gamma}}$ with $\widehat{\mathbf{Z}}, \widehat{\mathbf{\Gamma}} \in \mathbb{R}^{((\hat{l}+\hat{k}-1)\hat{q})\times((\hat{l}+\hat{k}-1)\hat{q})}$. The matrix $\widehat{\mathbf{\Gamma}} = s\, diag(0, 0, ..., M_h)$ and $\widehat{\mathbf{Z}} = \hat{D}_\tau \otimes M_h$, $\hat{D}_\tau :=$ blockdiag$(\hat{D}_\tau^1, \ldots, \hat{D}_\tau^{\hat{k}})$, and the

$\hat{D}^1_\tau \in \mathbb{R}^{(\hat{m}\hat{q})\times(\hat{m}\hat{q})}$ and $\hat{D}^k_\tau \in \mathbb{R}^{((\hat{m}+1)\hat{q})\times((\hat{m}+1)\hat{q})}$ for $2 \le k \le \hat{k}$, are the time mass matrices associated to the sub-intervals $[T_{k-1}, T_k]$, Note that $\mathbf{K}$ can be obtained by subassembling $\widehat{\mathbf{K}}$ and $(\hat{l} + \hat{k} - 1)\hat{q} = \hat{m}\hat{q} + (\hat{k}-1)(\hat{m}+1)\hat{q}$.

Remark 4. In the following, we shall express using matrix notation, the parareal algorithm described in the preceding. For convenience, we consider only *unique values* for the solution at the coarse times T_k, although the jumps can be computed using the evolution matrix. In matrix form, the parareal method to solve $\mathbf{E}\,\mathbf{y} = \mathbf{s}$ corresponds to a Schur complement based Richardson iteration. We partition $\mathbf{y} = \left(\mathbf{y}_I^T, \mathbf{y}_B^T\right)^T$ where $\mathbf{y}_I = (\mathbf{y}_I^{(1)^T}, \ldots, \mathbf{y}_I^{(\hat{k})^T})^T$ and $\mathbf{y}_I^{(k)} = (y^T_{(k-1)\hat{m}+1}, \ldots, y^T_{k\hat{m}-1})^T$ are sub-vectors of $\mathbf{y}$ at the times $t_j = j\,\tau$ in (T_{k-1}, T_k), while $\mathbf{y}_B = (y^T_{\hat{m}}, \ldots, y^T_{\hat{k}\hat{m}})^T$ denotes sub-vectors of $\mathbf{y}$ at the times $T_1, \ldots, T_{\hat{k}}$. This block partitions system $\mathbf{E}\,\mathbf{y} = \mathbf{s}$:

$$\begin{bmatrix} \mathbf{E}_{II} & \mathbf{E}_{IB} \\ \mathbf{E}_{BI} & \mathbf{E}_{BB} \end{bmatrix} \begin{bmatrix} \mathbf{y}_I \\ \mathbf{y}_B \end{bmatrix} = \begin{bmatrix} \mathbf{s}_I \\ \mathbf{s}_B \end{bmatrix}, \tag{4.4}$$

where $\mathbf{E}_{II} = \text{blockdiag}(\mathbf{E}_{II}^{(1)}, \ldots, \mathbf{E}_{II}^{(\hat{k})})$ is a block diagonal matrix, with $\mathbf{E}_{II}^{(k)}$ denoting an evolution submatrix of $\mathbf{E}$ on (T_{k-1}, T_k). The matrices $\mathbf{E}_{IB}$, $\mathbf{E}_{BI}$ and $\mathbf{E}_{BB}$ are also submatrices of $\mathbf{E}$, corresponding to the indices B and I. Elimination of $\mathbf{y}_I$ yields a Schur complement system $\mathbf{C}\mathbf{y}_B = (\mathbf{s}_B - \mathbf{E}_{BI}\mathbf{E}_{II}^{-1}\mathbf{s}_I)$, where the Schur complement matrix $\mathbf{C} = (\mathbf{E}_{BB} - \mathbf{E}_{BI}\mathbf{E}_{II}^{-1}\mathbf{E}_{IB})$ can be shown to be block lower bidiagonal, as expressed below, where $F := (F_0 F_1^{-1})^{\hat{m}-1} F_0$. In the parareal method, the following preconditioner $\mathbf{C}_0$ is also employed for $\mathbf{C}$, where $F_1^\Delta = (M_h + \Delta T\, A_h)$:

$$\mathbf{C} \equiv \begin{bmatrix} F_1 & & & \\ -F & F_1 & & \\ & \ddots & \ddots & \\ & & -F & F_1 \end{bmatrix} \quad \text{and} \quad \mathbf{C}_0 \equiv \begin{bmatrix} F_1^\Delta & & & \\ -G & F_1^\Delta & & \\ & \ddots & \ddots & \\ & & -G & F_1^\Delta \end{bmatrix}. \tag{4.5}$$

It can be shown that $\rho(I - C_0^{-1}C) < 1$ yielding a convergent iteration for the Schur complement system. The parareal preconditioner $\mathbf{E}_n$ for $\mathbf{E}$ corresponds to applying n Richardson iterations based on the Schur complement system. To compute the action of $\mathbf{x} = \mathbf{E}_n^{-1}\mathbf{s}$, in step 1 solve $\mathbf{E}_{II}\mathbf{w}_I = \mathbf{s}_I$. In step 2, apply n iterations of the iteration $\mathbf{w}_B^{i+1} = \mathbf{w}_B^i + \mathbf{C}_0^{-1}(\mathbf{s}_B - \mathbf{E}_{BI}\mathbf{w}_I - \mathbf{C}\mathbf{w}_B^i)$ starting with $\mathbf{w}_B^0 = \mathbf{0}$. Define $\mathbf{x}_B = \mathbf{w}_B^n$. In step 3, compute the update $\mathbf{x}_I = \mathbf{E}_{II}^{-1}(\mathbf{s}_I - \mathbf{E}_{IB}\mathbf{x}_B)$.

In the following result, we assume that parameter $\theta = 1$.

THEOREM 4.1. *For any* $\mathbf{s} \in \mathbb{R}^{(\hat{l}\hat{q})\times(\hat{l}\hat{q})}$ *and* $\epsilon \in (0, 1/2)$*, we have:*

$$\gamma_{\min}\left(\mathbf{E}^{-1}\mathbf{s}, \mathbf{K}\mathbf{E}^{-1}\mathbf{s}\right) \le \left(\mathbf{E}_n^{-1}\mathbf{s}, \widehat{\mathbf{K}}\mathbf{E}_n^{-1}\mathbf{s}\right) \le \gamma_{\max}\left(\mathbf{E}^{-1}\mathbf{s}, \mathbf{K}\mathbf{E}^{-1}\mathbf{s}\right),$$

where

$$\begin{cases} \gamma_{\max}(\epsilon) := \min\left((1 + \frac{\rho_n^2(t_f - t_o)}{\tau\epsilon} + 2\epsilon)/(1-2\epsilon), (1 + \frac{\tau\rho_n^2 e^{2(t_f - t_o)max_q(\lambda_q)}}{4\epsilon} + 2\epsilon)/(1-2\epsilon)\right) \\ \gamma_{\min}(\epsilon) := \max\left((1 - \frac{\rho_n^2(t_f - t_o)}{\tau\epsilon} - 2\epsilon)/(1+2\epsilon), (1 - \frac{\tau\rho_n^2 e^{2(t_f - t_o)max_q(\lambda_q)}}{4\epsilon} - 2\epsilon)/(1+2\epsilon)\right). \end{cases} \tag{4.6}$$

Proof. Let $V_h := [v_1, ..., v_{\hat{q}}]$ and $\Lambda_h := \text{diag}\{\lambda_1, ..., \lambda_{\hat{q}}\}$ be the generalized eigenvectors and eigenvalues of A_h with respect to M_h, i.e., $A_h = M_h V_h \Lambda_h V_h^{-1}$. Let $\mathbf{z} := \mathbf{E}^{-1}\mathbf{s}$

with $\underline{z}(t) = \sum_{q=1}^{\hat{q}} \alpha_q(t) v_q$, and $\hat{\mathbf{z}}^n := \mathbf{E}_n^{-1}\mathbf{s}$ with $\underline{\hat{z}}^n(t) = \sum_{q=1}^{\hat{q}} \alpha_q^n(t) v_q$. We note that α_q^n might be discontinuous across the T_k. Then:

$$(\mathbf{E}^{-1}\mathbf{s}, \mathbf{K}\mathbf{E}^{-1}\mathbf{s}) = \|\underline{z}\|^2_{L^2(t_o,t_f;L^2(\Omega))} + s\|\underline{z}(t_f)\|^2_{L^2(\Omega)} = \sum_{q=1}^{\hat{q}} \|\alpha_q\|^2_{L^2(t_o,t_f)} + s|\alpha_q(t_f)|^2,$$

$$(\mathbf{E}_n^{-1}\mathbf{s}, \widehat{\mathbf{K}}\mathbf{E}_n^{-1}\mathbf{s}) = \|\underline{\hat{z}}^n\|^2_{L^2(t_o,t_f;L^2(\Omega))} + s\|\underline{\hat{z}}(t_f)^n\|^2_{L^2(\Omega)} = \sum_{q=1}^{\hat{q}} \|\alpha_q^n\|^2_{L^2(t_o,t_f)} + s|\alpha_q^n(t_f)|^2.$$

First part (Estimation of $\|\alpha_q^n\|^2$). For estimating $\|\alpha_q^n\|^2$ we have

$$\begin{aligned}
\|\alpha_q^n\|^2_{L^2(t_o,t_f)} &= \left(\alpha_q^n - \alpha_q, \alpha_q^n + \alpha_q\right)_{L^2(t_o,t_f)} + \|\alpha_q\|^2_{L^2(t_o,t_f)} \\
&\le \frac{1}{4\epsilon}\|\alpha_q^n - \alpha_q\|^2_{L^2(t_o,t_f)} + \epsilon\|\alpha_q^n + \alpha_q\|^2_{L^2(t_o,t_f)} + \|\alpha_q\|^2_{L^2(t_o,t_f)} \\
&\le \frac{1}{4\epsilon}\|\alpha_q^n - \alpha_q\|^2_{L^2(t_o,t_f)} + 2\epsilon\|\alpha_q^n\|^2_{L^2(t_o,t_f)} + (1+2\epsilon)\|\alpha_q\|^2_{L^2(t_o,t_f)},
\end{aligned}$$

which reduces to:

$$(1-2\epsilon)\|\alpha_q^n\|^2_{L^2(t_o,t_f)} \le (1+2\epsilon)\|\alpha_q\|^2_{L^2(t_o,t_f)} + \frac{1}{4\epsilon}\|\alpha_q^n - \alpha_q\|^2_{L^2(t_o,t_f)}. \tag{4.7}$$

For each $t_l \in [T_{k-1}, T_k]$ we have:

$$|\alpha_q^n(t_l) - \alpha_q(t_l)| = \left((1+\tau\lambda_q)^{-1}\right)^{(t_l - T_{k-1})/\tau} |\alpha_q^n(T_{k-1}) - \alpha_q(T_{k-1})|,$$

and since $\lambda_q > 0$ implies $\left((1+\tau\lambda_q)^{-1}\right)^{(t_l-T_{k-1})/\tau} \le 1$, we obtain:

$$\|\alpha_q^n - \alpha_q\|^2_{L^2(T_{k-1},T_k)} \le \Delta T|\alpha_q^n(T_{k-1}) - \alpha_q(T_{k-1})|^2.$$

Hence:

$$(1-2\epsilon)\|\alpha_q^n\|^2_{L^2(t_o,t_f)} \le (1+2\epsilon)\|\alpha_q\|^2_{L^2(t_o,t_f)} + \frac{t_f - t_o}{4\epsilon}\max_{0\le k\le\hat{k}} |\alpha_q^n(T_k) - \alpha_q(T_k)|^2.$$

Using the Lemma 4.3 (see below) with $\alpha_q(T_0) = 0$ and initial guess $\alpha_q^0(T_k) = 0$, and using

$$\max_{0\le k\le\hat{k}} |\alpha_q(T_k)|^2 = |\alpha_q(T_{k'})|^2 \le \frac{4}{\tau}\min_{\beta}\|\alpha_q(T_{k'}) + \beta t\|^2_{L^2(T_{k'},T_{k'}+\tau)}$$

we obtain:

$$\max_{0\le k\le\hat{k}} |\alpha_q^n(T_k) - \alpha_q(T_k)|^2 \le \rho_n^2 \max_{0\le k\le\hat{k}} |\alpha_q(T_k)|^2 \le \frac{4\rho_n^2}{\tau}\|\alpha_q\|^2_{L^2(t_o,t_f)},$$

and the upper bound follows as:

$$\|\alpha_q^n\|^2_{L^2(t_o,t_f)} \le (1 + \frac{\rho_n^2(t_f - t_o)}{\tau\epsilon} + 2\epsilon)/(1-2\epsilon)\|\alpha_q\|^2_{L^2(t_o,t_f)}. \tag{4.8}$$

The lower bound for $\|\alpha_q^n\|^2$ follows similarly and given by

$$\|\alpha_q^n\|^2_{L^2(t_o,t_f)} \le (1 - \frac{\rho_n^2(t_f - t_o)}{\tau\epsilon} - 2\epsilon)/(1+2\epsilon)\|\alpha_q\|^2_{L^2(t_o,t_f)}. \tag{4.9}$$

Second part (Estimation of $|\alpha_q^n(t_f)|^2$). A similar expression to (4.7) holds, i.e.:

$$(1-2\epsilon)|\alpha_q^n(t_f)|^2 \leq (1+2\epsilon)|\alpha_q(t_f)|^2 + \frac{1}{4\epsilon}|\alpha_q(t_f)-\alpha_q^n(t_f)|^2. \tag{4.10}$$

Notice that, from the Lemma 4.3, we have

$$|\alpha_q(t_f)-\alpha_q^n(t_f)| \leq \max_{1\leq k\leq \hat{k}} |\alpha_q(T_k)-\alpha_q^n(T_k)| \leq \rho_n \max_{1\leq k\leq \hat{k}} |\alpha_q(T_k)|,$$

where $\alpha_q^0(T_k)=0$ for all k. Defining $\max_{0\leq k\leq \hat{k}} |\alpha_q(T_k)| = |\alpha_q(T_{k'})|$ and noting that:

$$\alpha_q(t_f) = e^{-(t_f-T_{k'})\lambda_q}\alpha_q(T_{k'}) \geq \frac{e^{-(t_f-t_o)\lambda_q}}{\tau}\alpha_q(T_{k'}) = \frac{e^{-(t_f-t_o)\lambda_q}}{\tau}\max_{0\leq k\leq \hat{k}} |\alpha_q(T_k)|,$$

we have

$$\max_{1\leq k\leq \hat{k}} |\alpha_q(T_k)|^2 \leq \tau e^{2(t_f-t_o)\max_q(\lambda_q)}|\alpha_q(t_f)|^2,$$

obtaining:

$$(1-2\epsilon)|\alpha_q^n(t_f)|^2 \leq \left(1+2\epsilon+\frac{\tau\rho_n^2 e^{2(t_f-t_o)\lambda_q}}{4\epsilon}\right)|\alpha_q(t_f)|^2. \tag{4.11}$$

The lower bound for $|\alpha_q^n(t_f)|$ follows similarly and it is given by

$$(1+2\epsilon)|\alpha_q^n(t_f)|^2 \leq \left(1-2\epsilon-\frac{\tau\rho_n^2 e^{2(t_f-t_o)\lambda_q}}{4\epsilon}\right)|\alpha_q(t_f)|^2. \tag{4.12}$$

Combining expression (4.8) and (4.11), and (4.9) and (4.12), yields the upper and lower bounds (4.6).

□

Remark 5. Performing straightforward computations we obtain:

$$\gamma_{\max} = \inf_{\epsilon\in(0,1/2)}(\gamma_{\max}(\epsilon)) = \min\left(1+\frac{4}{\sqrt{1+\frac{\tau}{\rho_n^2(t_f-t_o)}}-1}, 1+\frac{4}{\sqrt{1+\frac{4}{\tau\rho_n^2 e^{2(t_f-t_o)max_q(\lambda_q)}}}-1}\right), \tag{4.13}$$

and

$$\gamma_{\min} = \sup_{\epsilon\in(0,1/2)}(\gamma_{\min}(\epsilon)) = \max\left(1+\frac{-4}{\sqrt{1+\frac{\tau}{\rho_n^2(t_f-t_o)}}+1}, 1+\frac{-4}{\sqrt{1+\frac{4}{\tau\rho_n^2 e^{2(t_f-t_o)max_q(\lambda_q)}}}+1}\right). \tag{4.14}$$

Considering $(t_f-t_o)=O(1)$ and defining $\varpi(h) := \sqrt{e^{2(t_f-t_o)max_q(\lambda_q)}}$ then for small values of ρ_n we have

$$\gamma_{\max} \approx \min\left(1+4\rho_n\frac{1}{\sqrt{\tau}}, 1+2\sqrt{\varpi}\rho_n\sqrt{\tau}\right). \tag{4.15}$$

since $max_q(\lambda_q)=O(h^{-2})$.

LEMMA 4.2. *Let the matrices* $\mathbf{Z}$, $\mathbf{G}$, $\mathbf{N}$, $\mathbf{\Gamma}$ *and* $\mathbf{E}$ *satisfy the bounds described in Lemma* 2.2. *Let* $\mathbf{K} = \mathbf{Z} + \mathbf{\Gamma}$ *and* $\widehat{\mathbf{K}} = \widehat{\mathbf{Z}} + \widehat{\mathbf{\Gamma}}$. *Then,*

$$\mathbf{u}^T \left(\mathbf{N}^T \mathbf{E}_n^{-T} \widehat{\mathbf{K}} \mathbf{E}_n^{-1} \mathbf{N} \right) \mathbf{u} \leq (\mu_{\max} \gamma_{\max}) \mathbf{u}^T \mathbf{G} \mathbf{u}, \tag{4.16}$$

where $\mu_{\max}$ *and* $\gamma_{\max}$ *are given in* (3.2) *and* (4.13), *respectively. In terms of the parameters* r, s, h *and* τ, *it will hold that:*

$$\mu_{\max} \gamma_{\max} = O \left(\frac{1 + s/\tau}{r} \gamma_{\max} \right). \tag{4.17}$$

Proof. Using Theorems 4.1 and 3.1, we obtain:

$$\begin{aligned} \mathbf{u}^T \left(\mathbf{N}^T \mathbf{E}_n^{-T} \widehat{\mathbf{K}} \mathbf{E}_n^{-1} \mathbf{N} \right) \mathbf{u} &\leq \gamma_{\max} \mathbf{u}^T \left(\mathbf{N}^T \mathbf{E}^{-T} \mathbf{K} \mathbf{E}^{-1} \mathbf{N} \right) \mathbf{u} \\ &\leq \gamma_{\max} \mu_{\max} \mathbf{u}^T \mathbf{G} \mathbf{u}. \end{aligned} \tag{4.18}$$

Combining (4.18) with (3.16) yields the expression (4.17). ∎

Remark 6. Considering $s = 0$ and taking h constant, then ϖ is a constant. Consequently, for r constant, we obtain $\mu_{\max}\gamma_{\max} = O(\gamma_{\max})$ and $4\rho_n/\sqrt{\tau} \leq 2\rho_n\sqrt{\tau}$, hence $\gamma_{\max} = (1 + 4\rho_n \frac{1}{\sqrt{\tau}})$ in (4.15).

Next, decompose $Z_k = \sum_{q=1}^{\hat{q}} \alpha_q(T_k) v_q$ and $Z_k^n = \sum_{q=1}^{\hat{q}} \alpha_q^n(T_k) v_q$, and denote $\zeta_q^n(T_k) := \alpha_q(T_k) - \alpha_q^n(T_k)$. The convergence of the parareal algorithm for systems follows from the next lemma which is an extension of the results presented in [9].

LEMMA 4.3. *Let* $\Delta T = (t_f - t_o)/\hat{k}$ *and* $T_k = t_o + k\Delta T$ *for* $0 \leq k \leq \hat{k}$. *Then,*

$$\max_{1 \leq k \leq \hat{k}} |\alpha_q(T_k) - \alpha_q^n(T_k)| \leq \rho_n \max_{1 \leq k \leq \hat{k}} |\alpha_q(T_k) - \alpha_q^0(T_k)|,$$

where $\rho_n := \sup_{0<\beta<1} \left(e^{1-1/\beta} - \beta \right)^n \frac{1}{n!} \left| \frac{d^{n-1}}{d\beta^{n-1}} \left(\frac{1-\beta^{\hat{k}-1}}{1-\beta} \right) \right| \leq 0.2984256075^n$.

Proof. Using Theorem 2 from [9] we obtain:

$$\zeta_q^n = \left((1 + \tau\lambda_q)^{-\Delta T/\tau} - \beta_q \right) \mathcal{T}(\beta_q) \zeta_q^{n-1}, \tag{4.19}$$

where $\beta_q := (1 + \lambda_q \Delta T)^{-1}$ and $\mathcal{T}(\beta) := \left\{ \begin{array}{ll} \beta^{j-i-1} & \text{if } j > i, \quad 0 \quad \text{otherwise} \end{array} \right\}$ is a Toeplitz matrix of size $\hat{k}$. Applying (4.19) recursively we obtain:

$$\max_{1 \leq k \leq \hat{k}} |\zeta_q^n| \leq \rho_n^q \max_{1 \leq k \leq \hat{k}} |\zeta_q^0|,$$

where:

$$\rho_n^q := \left\| \left((1 + \tau\lambda_q)^{-\Delta T/\tau} - \beta_q \right)^n \mathcal{T}^n(\beta_q) \right\|_{L^\infty}. \tag{4.20}$$

Since $\lambda_q > 0$ and $\beta_q \leq (1 + \tau\lambda_q)^{-\Delta T/\tau} \leq e^{-\lambda_q \Delta T}$, we obtain

$$|\,(1 + \tau\lambda_q)^{-\Delta T/\tau} - \beta_q| \leq |e^{-\lambda_q \Delta T} - \beta_q| = |e^{1-1/\beta_q} - \beta_q|,$$

which yields:

$$\rho_n^q \leq |e^{1-1/\beta_q} - \beta_q|^n \|\mathcal{T}^n(\beta_q)\|_{L^\infty} \leq \sup_{0<\beta<1} |e^{1-1/\beta} - \beta|^n \|\mathcal{T}^n(\beta)\|_{L^\infty}.$$

By considering $\|\mathcal{T}^n(\beta)\|_\infty \leq \|\mathcal{T}(\beta)\|_\infty^n = \left| \frac{1-\beta^{\hat{k}-1}}{1-\beta} \right|^n$, a simpler upper bound for ρ_n can be obtained:

$$\sup_{0<\beta<1} \left|e^{1-1/\beta} - \beta\right|^n \left|\frac{1-\beta^{\hat{k}-1}}{1-\beta}\right|^n \leq \left(\sup_{0<\beta<1} \frac{e^{1-1/\beta}-\beta}{1-\beta}\right)^n \approx 0.2984256075^n,$$

and the maximum β_* is attained around 0.3528865, independently of n and $\hat{k}$ (β_* presents slight variation for $1 \leq n$ and $6 \leq \hat{k}$, cases of practical interest). □

Third Algorithm. Our third algorithm employs the block diagonal matrix:

$$\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{E}_n \widehat{\mathbf{K}}^{-1} \mathbf{E}_n^T & \\ & \mathbf{G} \end{bmatrix} \tag{4.21}$$

as a preconditioner for (3.18) and solves it using MINRES. The next theorem estimates the condition number when the matrix (4.21) is used as a preconditioner.

THEOREM 4.4. *Let the matrices* $\mathbf{Z}$, $\mathbf{G}$, $\mathbf{N}$, $\mathbf{\Gamma}$ *and* $\mathbf{E}$ *satisfy the bounds described in Lemma* 2.2. *Let* $\tilde{\mathbf{P}} :=$ *blockdiag*$(\mathbf{E}_n \widehat{\mathbf{K}}^{-1} \mathbf{E}_n^T, \mathbf{G})$ *denote a block diagonal preconditioner for the coefficient matrix* $\mathbf{H}$ *of system* (3.18). *Let the bounds for matrix* $\mathbf{E}_n \widehat{\mathbf{K}}^{-1} \mathbf{E}_n^T$ *presented in Lemma* 4.1. *Then, the condition number of the preconditioned system will satisfy the bound:*

$$\text{cond}(\tilde{\mathbf{P}}^{-1}\mathbf{H}) \leq O\left(\left(1 + \frac{1+s/\tau}{r}\right)^{1/2}\right) \frac{\max(\gamma_{\max}, 1)}{\min(\gamma_{\min}, 1)}, \tag{4.22}$$

where $\gamma_{\max}$ *and* $\gamma_{\min}$ *are defined in* (4.13) *and* (4.14), *respectively.*

Proof. To obtain an upper bound consider that $\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}$ is symmetric, then

$$\max|\eta\left(\tilde{\mathbf{P}}^{-1}\mathbf{H}\right)| = \max|\eta\left(\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}\right)| = |||\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}|||_2 \tag{4.23}$$

$$= \sup_{\mathbf{v}\neq 0}\left(\frac{\mathbf{v}^T\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}\mathbf{v}}{\mathbf{v}^T\mathbf{v}}\right) \tag{4.24}$$

$$\leq \sup_{\mathbf{v}\neq 0}\left(\frac{\mathbf{v}^T\mathbf{H}\mathbf{v}}{\mathbf{v}^T\mathbf{P}\mathbf{v}}\right) \sup_{\mathbf{v}\neq 0}\left(\frac{\mathbf{v}^T\mathbf{P}\mathbf{v}}{\mathbf{v}^T\tilde{\mathbf{P}}\mathbf{v}}\right) \tag{4.25}$$

$$\leq \max|\eta\left(\mathbf{P}^{-1}\mathbf{H}\right)| \max(\gamma_{\max}, 1) \tag{4.26}$$

$$\tag{4.27}$$

Analogously, using $\min|\eta\left(\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}\right)| = |||\left(\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}\right)|||_2$, a lower bound is obtained:

$$\min|\eta\left(\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}\right)| \geq \min|\eta\left(\tilde{\mathbf{P}}^{-1/2}\mathbf{H}\tilde{\mathbf{P}}^{-1/2}\right) \min(\gamma_{\min}, 1) \tag{4.28}$$

and consequently,

$$\text{cond}(\tilde{\mathbf{P}}^{-1}\mathbf{H}) \leq O\left(\left(1 + \frac{1+s/\tau}{r}\right)^{1/2}\right) \frac{\max(\gamma_{\max}, 1)}{\min(\gamma_{\min}, 1)}. \tag{4.29}$$

□

Remark 7. A generalization of Theorem 4.4 for matrices $\mathbf{G}_o$ and $\mathbf{E}_n \widehat{\mathbf{K}}_n^{-1} \mathbf{E}_n^T$ (spectrally equivalent to $\mathbf{G}$ and $\mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T$ respectively) follows directly from [13, 14].

5. Numerical Experiments. In this section, we consider the numerical solution of an optimal control problem involving the 1D-heat equation. In this case, the constraints are given by:

$$z_t - z_{xx} = v, \quad 0 < x < 1, \quad t > 0$$

TABLE 5.1
Number of CG iterations for Algorithm 1. The parameters are $s = 0$ ($s = 1$), $q = 1$, $r = 0.0001$, $t_f = 1$, $h = 1/\hat{m}$ and $\tau = 1/\hat{l}$.

$\hat{m} \setminus \hat{l}$	32	64	128	256	512
32	36 (41)	38 (44)	40 (46)	40 (47)	40 (48)
64	36 (41)	38 (44)	40 (46)	40 (47)	40 (48)
128	36 (41)	38 (44)	40 (46)	40 (47)	40 (48)
256	36 (41)	38 (44)	40 (46)	40 (47)	40 (48)
512	36 (41)	38 (44)	40 (46)	40 (47)	40 (48)

with boundary conditions $z(t, 0) = 0$ and $z(t, 1) = 0$ for $t \geq 0$, and with initial data $z(0, x) = 0$ for $x \in [0, 1]$, and with the performance function $\tilde{y} = x(1 - x)e^{-x}$ for all $t \in [0, 1]$. Following [18], we take $q = 1$ and $r = 0.0001$. The backward Euler discretization ($\theta = 1$) is considered in the numerical experiments. As a stopping criteria for the iterative solvers, we take $\|\mathbf{r_k}\|/\|\mathbf{r_0}\| \leq 10^{-9}$ where $\mathbf{r_k}$ is the residual at each iteration k.

TABLE 5.2
Number of CG iterations for Algorithm 1 for different values of τ and s. The parameters are: space discretization $h = 1/32$, $q = 1$, $r = 0.0001$, $t_f = 1$.

$\hat{l} \setminus s$	0	1	10	100
32	36	41	50	68
64	38	44	52	73
128	40	46	57	83
256	40	47	62	89
512	40	48	64	96

TABLE 5.3
Condition number of the preconditioned matrix of Algorithm 1. The parameters are $q = 1$, $r = 0.0001$, $t_f = 1$, $h = 1/32$ and $\tau = 1/64$

$r \setminus s$	10^4	10^2	1	10^{-2}	0
10^{-2}	$4.9\ 10^4$	$5.0\ 10^2$	6.2	1.9	1.9
10^{-4}	$4.7\ 10^6$	$4.8\ 10^4$	$5.2\ 10^2$	93	93
10^{-6}	$4.2\ 10^8$	$4.8\ 10^6$	$5.1\ 10^4$	$9.2\ 10^3$	$9.1\ 10^3$

Algorithm 1: Reduction to u. We consider matrix $\mathbf{G}$ as a preconditioner for system (3.1) and use the PCG method to solve the resulting preconditioned system. For the case where $s = 0$ and in parenthesis $s = 1$, Table 5.1 presents the number of iterations for different time and space meshes. As predicted by the theory in Section 3, see (3.16), the number of iterations remains constant as h is refined. Table 5.1 also shows that the number of iterations deteriorates very weakly when the time discretization τ gets finer. As expected from the analysis, this deterioration is more noticeable for larger s, (see Tables 5.1 and 5.2). Table 5.3 shows that the condition number estimates in Section 3 (see the expression (3.17)) are sharp for different values of parameters r and s.

Algorithm 2. Table 5.6 presents the number of iterations required to solve system (3.18) using MINRES acceleration when both time and space grid sizes are refined. As predicted from the analysis, see (3.19), as the space grid is refined, the

TABLE 5.4
Values of $\mu_{\max}\gamma_{\max}$ when τ is refined.

Parameters $h = 1/10$, $\Delta T = 1/20$ and $s = 0$.

$n \setminus \hat{l}$	200	400	800	1600
$n = 1$	0.864415	1.449299	2.473734	4.371709
$n = 2$	0.070835	0.097852	0.136802	0.193845
$n = 3$	0.007760	0.010765	0.015141	0.021165
$n = 4$	0.000865	0.001224	0.001715	0.002397

Parameters $h = 1/10$, $\Delta T = 1/20$ and $s = 0.01$.

$n \setminus \hat{l}$	200	400	800	1600
$n = 1$	0.864419	1.449305	2.473744	4.371725
$n = 2$	0.070852	0.097885	0.136851	0.193918
$n = 3$	0.007784	0.010800	0.015180	0.021208
$n = 4$	0.000869	0.001229	0.001721	0.002436

Parameters $h = 1/10$, $\Delta T = 1/20$ and $s = 1$.

$n \setminus \hat{l}$	200	400	800	1600
$n = 1$	3.779956	6.866334	12.933472	24.992363
$n = 2$	0.226120	0.333472	0.500244	0.768025
$n = 3$	0.024360	0.034622	0.049315	0.070448
$n = 4$	0.002846	0.004026	0.005699	0.008069

TABLE 5.5
Values of $\mu_{\max}\gamma_{\max}$ for $\tau = 1/800$ and $s = 0/1/10/100$. Parameters $h = 1/10$, $\Delta T = 1/20$.

$n \setminus s$	0	1	10	100
$n = 2$	0.136802	0.500244	2.239858	16.307768
$n = 3$	0.015141	0.049315	0.157836	0.581009
$n = 4$	0.001715	0.005699	0.017459	0.056029

number of iterations remains bounded. As before, for larger s, a deterioration in the number of iterations is observed.

Algorithm 3: parareal preconditioner. Table 5.7 lists the number of MINRES iterations as $\Delta T/\tau$ is varied while τ remains as a constant. We choose $n = 2$. The number of iterations for the MINRES basically remains constant when h is refined and $\hat{k}$ is increased, and so the results indicate scalability. Table 5.8 lists the number of MINRES iterations for $n = 2$ and $\tau = 1/512$ for different values of $\Delta T/\tau$. It indicates optimal order of convergence.

Sharpness of the bound in Lemma 4.2. Table 5.4 lists the value of $\mu_{\max}\gamma_{\max}$, see (4.17 for different values of τ, n (where $n = 7$ is in practice equivalent to an exact solver) and considering $s = 0$. The results in Table 5.4 for $\mu_{\max}\gamma_{\max}$ confirm the sharp dependence with respect to τ since it increases by a $\sqrt{2}$ factor when τ is refined by half and indicate that the method is *scalable* if "n" is kept constant, see Remark 6. Similar behavior if $s \neq 0$ but constant. If s is increased, the values of $\mu_{\max}\gamma_{\max}$ deteriorates, as expected in the expression (4.17) (see Table 5.5).

6. Conclusion. In this paper we have described three approaches for iteratively solving a linear quadratic parabolic optimal control problem. The first method is based on the CG solution of a Schur complement system. This is obtained by reducing the saddle point system to the system associated with the control variable. This method is simple to implement but requires double iteration. The second method avoids double iteration by introducing an auxiliary variable. The resulting system is symmetric and indefinite, so that MINRES can be used. The structure of this method

TABLE 5.6
Number of MINRES iterations for algorithm 2. Parameter are $s = 0$ ($s = 1$), $q = 1$, $r = 0.0001$, $t_f = 1$, $h = 1/\hat{m}$ and $\tau = 1/\hat{l}$.

$\hat{m} \setminus \hat{l}$	32	64	128	256	512
32	56 (60)	58 (62)	60 (64)	60 (66)	60 (68)
64	56 (60)	58 (62)	60 (64)	60 (66)	60 (68)
128	56 (60)	58 (62)	60 (64)	60 (66)	60 (68)
256	56 (60)	58 (62)	60 (64)	60 (66)	60 (68)
512	56 (60)	58 (62)	60 (64)	60 (66)	60 (68)

TABLE 5.7
MINRES iterations using the Parareal algorithm with $n = 2$ as preconditioner. Parameters $r = 0.001/0.0001/0.00001$ and $\tau = 1/512$.

$\hat{k}$	8	16	32	64
$\Delta T/\tau$	64	32	16	8
$h = 1/16$	32 / 62 / 136	32 / 62 / 136	32 / 60 / 132	32 / 60 / 132
$h = 1/32$	32 / 62 / 136	32 / 62 / 136	32 / 62 / 132	32 / 60 / 132
$h = 1/64$	32 / 62 / 136	32 / 62 / 136	32 / 62 / 132	32 / 60 / 132

also allows parallel block preconditioners. The preconditioners described yield a rate of convergence independent of the time and space parameters (under the specific choice of r and s). In the third method, a preconditioner based on the parareal algorithm is also presented which yields a rate of convergence for the MINRES constant when the spatial grid is refined and the number of subdomains is increased.

REFERENCES

[1] A. Battermann and M. Heinkenschloss. Preconditioners for karush-kuhn-tucker matrices arising in the optimal control of distributed systems. In *In W. Desch, F. Kappel, K. Kunisch, editors, Optimal Control of Partial Differential Equations, Vorau 1997*, pages 15 – 32. Birkhäuser Verlag, Basel, Boston, Berlin, 1998.

[2] A. Battermann and E. W. Sachs. Block preconditioner for kkt systems in pde-governed optimal control problems. In *In R. H. W. Hoppe, K.-H. Hoffmann, and V. Schulz, editors, Workshop on Fast Solutions of Discretized Optimization Problems*, pages 1 – 18. Birkhäuser, 2001.

[3] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numer.*, pages 1 – 37, 2005.

[4] M. Berggren and M. Heinkenschloss. Parallel solution of optimal-control problems by time-domain decomposition. In *In M.-O. Bristeau, G. Etgen, W. Fitzgibbon, J. L. Lions, J. Periaux, and M. F. Wheeler, editors, Computational Science for the 21st Century*, pages 102 – 112. J. Wiley, Chichester, 1997.

[5] G. Biros and O. Gattas. Parallel Lagrange-Newton-Krylov-Schur methods for pde-constrained optimization. Part I: The Krylov-Schur solver. *SIAM J. Sci. Comput.*, 27(2):687 – 713, 2005.

[6] G. Biros and O. Ghattas. Lagrange-Newton-Krylov-Schur methods for PDE constrained optimization. Part II: The Lagrange Newton solver and its application to optimal control of steady viscous flows. *SIAM J. Sci. Comput.*, 27(2):714 – 739, 2005.

[7] J. W. Demmel. *Applied Numerical Linear ALgebra.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[8] C. Farhat and M. Chandesris. Time-decomposed parallel time-integrators: theory and feasibility studies for fluid, structure, and fluid-structure applications. *Int. J. Numer. Methods Eng.*, 58(9):1397 – 1434, 2003.

[9] M. J. Gander and S. Vandewalle. On the super linear and linear convergence of the parareal algorithm. In D. E. Keyes and O. B. Widlund, editors, *Domain decomposition methods in science and engineering XVI*, pages 291–298. Springer, Lect. Notes Comput. Sci. Eng.,

TABLE 5.8

MINRES iterations using a parareal with $n = 2/4/7$ as preconditioners. Parameters $r = 0.0001$ and $\Delta T/\tau = 16$.

$\hat{k}$	4	8	16	32
$\hat{l}$	64	128	256	512
$h = 1/16$	62 / 40 / 42	58 / 44 / 44	60 / 50 / 44	60 / 50 / 44
$h = 1/32$	60 / 42 / 42	58 / 44 / 44	60 / 50 / 44	62 / 50 / 44
$h = 1/64$	60 / 42 / 42	58 / 44 / 44	60 / 50 / 44	62 / 50 / 44

2005.

[10] M. J. Gander and S. Vandewalle. Analysis of the parareal time-parallel time-integration method. *SIAM J. Sci. Comput.*, 29(2):556–578, 2007.

[11] E. Haber and U. M. Ascher. Preconditioned all-at-once methods for large, sparse parameter estimation problems. *Inverse Problems*, 17(6):1847–1864, 2001.

[12] M. Heinkenschloss. A time-domain decomposition iterative method for the solution of distributed linear quadratic optimal control problems. *J. Comput. Appl. Math.*, 173(1):169 – 198, 2005.

[13] A. Klawonn. Preconditioners for indefinite problems. Technical report, 1996. Reports CS-NYU, TR1996-716, March.

[14] A. Klawonn. Block triangular preconditioners for saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, 19(1):172 – 184, 1998.

[15] J. L. Lions. *Optimal control of systems governed by partial differential equations.* Springer-Verlag, Berlin-Heidelberg-New York, 1971.

[16] J. L. Lions, Y. Maday, and G. Turinici. Résolution d'edp par un schéma en temps pararéel. *C. R. Acad. Sci. Paris Sr. I Math.*, 332(7):661 – 668, 2001.

[17] J. L. Lions, Y. Maday, and G. Turinici. Résolution d'edp par un schéma en temps pararéel. *C. R. Acad. Sci. Paris Sr. I Math.*, 332(7):661–668, 2001.

[18] A. Locatelli. *Otimal Control: An Introduction.* Birkhäuser,Berlin, 2001.

[19] Y. Maday and G. Turinici. A parareal in time procedure for the control of partial differential equations. *C. R. Math. Acad. Sci. Paris*, 335(4):387 – 392, 2002.

[20] T. Mathew, P. L. Polyakov, G. Russo, and J. Wang. Domain decomposition operator splittings for the solution of parabolic equations. *SIAM J. Sci. Comput.*, 19(3):912 – 932, 1998.

[21] T. Mathew, M. Sarkis, and C. E. Schaerer. Analysis of block matrix preconditioners for elliptic optimal control problems. *Numer. Linear Algebra Appl.*, 14:257 – 279, 2007.

[22] T. P. Mathew, M. Sarkis, and C. E. Schaerer. Block iterative algorithms for the solution of parabolic optimal control problems. In M. Daydé et al., editor, *7th International Meeting High Performance Computing for Computational Science - VECPAR 2006*, pages 452–465. Springer, Lect. Notes Comput. Sci., 4395, 2007.

[23] P. Neittaanäki and D. Tiba. *Optimal control of linear parabolic systems: theory, algorithms and applications.* Marcel Dekker, New York, 1994.

[24] J. Nievergelt. Parallel methods for integrating ordinary differential equations. *Comm. ACM*, 7(12):731 – 733, 1964.

[25] T. Rusten and R. Winther. A preconditioned iterative method for saddle point problems. *SIAM J. Matrix Anal.*, 13(3):887 – 904, 1992.

[26] M. Sarkis, C. E. Schaerer, and T. Mathew. Block diagonal parareal preconditioner for parabolic optimal control problems. In *Domain Decomposition Methods in Science and Engineering XVII*, pages 409–416. Springer, Lect. Notes Comput. Sci. Eng., 60, 2008.

[27] C. E. Schaerer and E. Kaszkurewicz. The shooting method for the numerical solution of ordinary differential equations: a control theoretical perspective. *Internat. J. Systems Sci.*, 32(8):1047 – 1053, 2001.

[28] R. S. Varga. *Matrix Iterative Analysis.* Prentice Hall, Englewood Cliffs, NJ, 1962.