Block Diagonal Parareal Preconditioner for Parabolic Optimal Control Problems

Marcus Sarkis[1,2], Christian E. Schaerer[1], and Tarek Mathew[1]

[1] IMPA, Dona Castorina 110, Rio de Janeiro, RJ 22460-320, Brazil.
{msarkis,cschaer}@impa.br
[2] WPI, 100 Institute Road, Worcester, MA 01609, USA.
{tmathew}@poonithara.org

Summary. We describe a block matrix iterative algorithm for solving a linear-quadratic parabolic optimal control problem (OCP) on a finite time interval. We derive a reduced symmetric indefinite linear system involving the control variables and auxiliary variables, and solve it using a preconditioned MINRES iteration, with a symmetric positive definite block diagonal preconditioner based on the parareal algorithm. Theoretical and numerical results show that the preconditioned algorithm converges at a rate independent of the mesh size h, and has parallel scalability.

1 Introduction

Let (t_0, t_f) denote a time interval, let $\Omega \subset \mathbb{R}^2$ be a 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 $H = L^2(\Omega)$; see [2]. Denote the state variable space as $\mathcal{Y} = \{z \in L^2(t_o, t_f; Y) : z_t \in L^2(t_o, t_f; Y')\}$, where it can be shown that $\mathcal{Y} \subset \mathcal{C}^0([t_o, t_f]; H)$; see [2]. Given $y_o \in H$, we consider the following state equation on (t_0, t_f) with $z \in \mathcal{Y}$:

$$\begin{cases} z_t + \mathcal{A}z = \mathcal{B}v \quad \text{for} \quad t_o < t < t_f, \\ \qquad z(0) = y_o. \end{cases} \tag{1}$$

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) is well posed, see [2], and we indicate the dependence of z on $v \in \mathcal{U}$ using the notation $z(v)$. Given a target function $\hat{y}$ in $L^2(t_o, t_f; H)$ and parameters $q > 0$, $r > 0$, we shall employ the following cost function which we associate with the state equation (1):

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

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

$$J(y,u) \;:=\; \min_{v\in\mathcal{U}} \, J(z(v),v). \tag{3}$$

Since $q, r > 0$, the optimal control problem (3) is well posed, see [2].

Our presentation is organized as follows: In § 2 we discretize (3) using a finite element method and backward Euler discretization, yielding a large scale saddle point system. In § 3, we introduce and analyze a symmetric positive definite block diagonal preconditioner for the saddle point system, based on the *parareal* algorithm [3]. In § 4, we present numerical results which illustrate the scalability of the algorithm.

2 The Discretization and the Saddle Point System

To discretize the state equation (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 $\mathbb{P}_1$ conforming finite element space $Y_h \subset Y$ for $z(t,\cdot)$, and the $\mathbb{P}_0$ 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. To indicate their time dependence we denote $\underline{z}$ and $\underline{v}$.

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

$$J_h(\underline{z},\underline{v}) := \frac{1}{2}\int_{t_o}^{t_f} (\underline{z}-\underline{\hat{y}})^T(t) M_h (\underline{z}-\underline{\hat{y}})(t)\, dt + \frac{r}{2}\int_{t_o}^{t_f} \underline{v}^T(t) R_h \underline{v}(t)\, dt \tag{4}$$

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

$$M_h \dot{\underline{z}} + A_h\, \underline{z} = B_h \underline{v}, \;\; \text{for} \;\; t_o < t < t_f; \;\; \text{and} \;\; \underline{z}(t_o) = y_o^h. \tag{5}$$

Here $(\underline{z} - \underline{\hat{y}}^h)(t)$ denotes the tracking error, where $\underline{\hat{y}}^h(t)$ and y_0^h belong to Y_h and are approximations of $\underline{\hat{y}}(t)$ and y_o (for instance, use $L^2(\Omega)$-projections into Y_h). The matrices $M_h, A_h \in \mathbb{R}_h^{\hat{q}\times\hat{q}}$, $B_h \in \mathbb{R}^{\hat{q}\times\hat{p}}$ 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 (4) and (5), 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 a discontinuous function and constant in each sub-interval (t_{l-1}, t_l) with associated basis functions $\{\chi_l\}_{l=1}^{\hat{l}}$. Denoting $v_l \in \mathbb{R}^{\hat{p}}$ as the nodal representation of $\underline{v}(t_l - (\tau/2))$, we have $\underline{v}(t) = \sum_{l=1}^{\hat{l}} v_l \chi_l(t)$.

The corresponding discretization of the expression (4) results in:

$$J_h^\tau(\mathbf{z},\mathbf{v}) = \frac{1}{2}\,(\mathbf{z}-\hat{\mathbf{y}})^T \mathbf{K}(\mathbf{z}-\hat{\mathbf{y}}) + \frac{1}{2}\,\mathbf{v}^T\mathbf{G}\mathbf{v} + (\mathbf{z}-\hat{\mathbf{y}})^T\mathbf{g}. \tag{6}$$

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

Employing the backward Euler discretization of (5) in time, yields:

$$\mathbf{E}\,\mathbf{z} + \mathbf{N}\,\mathbf{v} = \mathbf{f}, \tag{7}$$

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

$$\mathbf{E} = \begin{bmatrix} F_h & & & \\ -M_h & F_h & & \\ & \ddots & \ddots & \\ & & -M_h & F_h \end{bmatrix}, \tag{8}$$

where $F_h = (M_h + \tau A_h) \in \mathbb{R}^{\hat{q}\times\hat{q}}$. 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 $\mathcal{L}_h(\mathbf{z}, \mathbf{v}, \mathbf{q})$ for minimizing (6) subject to constraint (7) is:

$$\mathcal{L}_h^\tau(\mathbf{z}, \mathbf{v}, \mathbf{q}) = J_h^\tau(\mathbf{z}, \mathbf{v}) + \mathbf{q}^T(\mathbf{E}\mathbf{z} + \mathbf{N}\mathbf{v} - \mathbf{f}). \tag{9}$$

To obtain a discrete saddle point formulation of (9), we apply optimality conditions for $\mathcal{L}_h(\cdot,\cdot,\cdot)$. This yields the symmetric indefinite linear system:

$$\begin{bmatrix} \mathbf{K} & \mathbf{0} & \mathbf{E}^T \\ \mathbf{0} & \mathbf{G} & \mathbf{N}^T \\ \mathbf{E} & \mathbf{N} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{K}\hat{\mathbf{y}} - \mathbf{g} \\ \mathbf{0} \\ \mathbf{f} \end{bmatrix}, \tag{10}$$

where $\hat{\mathbf{y}} := [(\hat{y}_1^h)^T, \dots, (\hat{y}_{\hat{l}}^h)^T]^T \in \mathbb{R}^{\hat{l}\hat{q}}$. Eliminating $\mathbf{y}$ and $\mathbf{p}$ in (10), and defining $\mathbf{b} := \mathbf{N}^T\mathbf{E}^{-T}\left(\mathbf{K}\mathbf{E}^{-1}\mathbf{f} - \mathbf{K}\hat{\mathbf{y}} + \mathbf{g}\right)$ yields the *reduced* Hessian system:

$$(\mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N})\mathbf{u} = \mathbf{b}. \tag{11}$$

The matrix $\mathbf{H} := \mathbf{G} + \mathbf{N}^T\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}$ is symmetric positive definite and $(\mathbf{u}, \mathbf{G}\mathbf{u}) \le (\mathbf{u}, \mathbf{H}\mathbf{u}) \le \mu(\mathbf{u}, \mathbf{G}\mathbf{u})$, where $\mu = O(1 + \frac{1}{r})$; for details see [4]. As a result, the Preconditioned Conjugate Gradient method (PCG) can be used to solve (11), but each matrix-vector product with $\mathbf{H}$ requires the solution of two linear systems, one with $\mathbf{E}$ and one with $\mathbf{E}^T$. To avoid double iterations, we define the auxiliary variable $\mathbf{w} := -\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}\mathbf{N}\mathbf{u}$. Then (11) will be equivalent to the symmetric indefinite system:

$$\begin{bmatrix} \mathbf{E}\mathbf{K}^{-1}\mathbf{E}^T & \mathbf{N} \\ \mathbf{N}^T & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\mathbf{b} \end{bmatrix}. \tag{12}$$

The system (12) is ill-conditioned and will be solved using the MINRES algorithm with a preconditioner of the form $\mathbf{P} := \mathrm{diag}(\mathbf{E}_n^{-T}\hat{\mathbf{K}}\mathbf{E}_n^{-1}, \mathbf{G}^{-1})$; see [5]. For a fixed number of parareal sweeps n, $\mathbf{E}_n^{-1}$ and $\mathbf{E}_n^{-T}$ are linear operators. We next define the operator $\mathbf{E}_n^{-1}$ and then analyze the spectral equivalence between $\mathbf{E}^{-T}\mathbf{K}\mathbf{E}^{-1}$ and $\mathbf{E}_n^{-T}\hat{\mathbf{K}}\mathbf{E}_n^{-1}$.

3 Parareal Approximation $\mathbf{E}_n^{-T}\hat{\mathbf{K}}\mathbf{E}_n^{-1}$

An application of $\mathbf{E}_n^{-T}\hat{\mathbf{K}}\mathbf{E}_n^{-1}$ to a vector $\mathbf{s} \in \mathbb{R}^{(\hat{l}\hat{q})\times(\hat{l}\hat{q})}$ is performed as follows: Step 1, apply $\mathbf{E}_n^{-1}\mathbf{s} :\longrightarrow \hat{\mathbf{z}}^n$ using n applications of the parareal method described below. Step 2, multiply $\hat{\mathbf{K}}\mathbf{z}^n :\longrightarrow \hat{\mathbf{t}}$ where $\hat{\mathbf{K}} := \hat{D}_\tau \otimes M_h$, $\hat{D}_\tau :=$ blockdiag$(\hat{D}_\tau^1, \ldots, \hat{D}_\tau^{\hat{k}})$, and the $\hat{D}_\tau^k$ are the time mass matrices associated to the sub-intervals $[T_{k-1}, T_k]$. And Step 3, apply $\mathbf{E}_n^{-T}\hat{\mathbf{t}}^n :\longrightarrow \mathbf{x}$, i.e., the transpose of Step 1.

To describe $\mathbf{E}_n$, we partition the time interval $[t_o, t_f]$ 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 $1 \le k \le \hat{k}$. We define fine and coarse propagators F and G as follows. The local solution at T_k is defined marching the backward Euler method from T_{k-1} to T_k 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 it is easy to see that:

$$M_h Z_k = F Z_{k-1} + S_k, \tag{13}$$

where $F := (M_h F_h^{-1})^{\hat{m}} M_h \in \mathbb{R}^{\hat{q}\times\hat{q}}$, $S_k := \sum_{m=1}^{\hat{m}} \left(M_h F_h^{-1}\right)^{\hat{m}-m+1} s_{j_{k-1}+m}$ with $Z_0 = 0$. Imposing the continuity condition at time T_k, for $1 \le k \le \hat{k}$, i.e., $M_h Z_k - F Z_{k-1} - S_k = 0$, we obtain the system:

$$\begin{bmatrix} M_h & & & \\ -F & M_h & & \\ & \ddots & \ddots & \\ & & -F & M_h \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{\hat{k}} \end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_{\hat{k}} \end{bmatrix}. \tag{14}$$

The coarse solution at T_k with initial data $Z_{k-1} \in \mathbb{R}^{\hat{q}}$ at T_{k-1} is given by one coarse time step of the backward Euler method $M_h Z_k = G Z_{k-1}$ where $G := M_h(M_h + A_h\Delta T)^{-1} M_h \in \mathbb{R}^{\hat{q}\times\hat{q}}$. In the parareal algorithm, the coarse propagator G is used for preconditioning the system (14) via:

$$\begin{bmatrix} Z_1^{n+1} \\ Z_2^{n+1} \\ \vdots \\ Z_{\hat{k}}^{n+1} \end{bmatrix} = \begin{bmatrix} Z_1^n \\ Z_2^n \\ \vdots \\ Z_{\hat{k}}^n \end{bmatrix} + \left(\begin{bmatrix} M_h & & & \\ -G & M_h & & \\ & \ddots & \ddots & \\ & & -G & M_h \end{bmatrix} \right)^{-1} \begin{bmatrix} R_1^n \\ R_2^n \\ \vdots \\ R_{\hat{k}}^n \end{bmatrix}, \tag{15}$$

where the residual vector $\mathbf{R}^n := [R_1^{nT}, ..., R_{\hat{k}}^{nT}]^T \in \mathbb{R}^{\hat{k}\hat{q}}$ is defined in the usual way from the equation (14).

We are now in position to 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 $[t_o, t_f]$, however continuous only inside each coarse sub-interval $[T_{k-1}, T_k]$, i.e., the function $\underline{\hat{z}}^n$ can be discontinuous across the 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 the backward Euler method from T_{k-1} to T_k on the fine triangulation τ with initial condition Z_{k-1}^n at T_{k-1}.

Theorem 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:*

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

$$\text{where} \begin{cases} \gamma_{\max} := (1 + \frac{\rho_n^2 (t_f - t_o)}{\tau\epsilon} + 2\epsilon)/(1-2\epsilon), \\ \gamma_{\min} := (1 - \frac{\rho_n^2 (t_f - t_o)}{\tau\epsilon} - 2\epsilon)/(1+2\epsilon). \end{cases}$$

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:

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

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

and therefore:

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

which reduces to:

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

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

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

and since $\lambda_q > 0$ implies $(1+\tau\lambda_q)^{-(t_l - T_{k-1})/\tau} \le 1$, 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 1 (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 (16) follows. The lower bound follows similarly.

Remark 1. Performing straightforward computations we obtain:

$$\min_{\epsilon} \gamma_{\max}(\epsilon) = 1 + \frac{4}{\sqrt{1 + \frac{\tau}{\rho_n^2 (t_f - t_o)}} - 1}.$$

Hence, for small values of ρ_n, we have $\gamma_{\max} - 1 \approx 4\sqrt{\frac{\rho_n^2 (t_f - t_o)}{\tau}}$. The dependence of $\gamma_{max} - 1$ with respect to τ is sharp as evidenced in Table 1 (see below) since it increases by a $\sqrt{2}$ factor when τ is refined by half.

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 it is an extension of the results presented in [1].

Lemma 1. *Let $\Delta T = (t_f - t_o)/\hat{k}$ and $T_k = t_o + k\Delta T$ for $0 \le k \le \hat{k}$. Then,*

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

where $\rho_n := \sup_{0<\beta<1} \left(e^{1-1/\beta} - \beta\right)^n \frac{1}{n!} \left| \frac{d^{n-1}}{d\beta^{n-1}} \left(\frac{1-\beta^{\hat{k}-1}}{1-\beta} \right) \right| \le 0.2984^n$.

Proof. Using Theorem 2 from [1] we obtain:

$$\zeta_q^n = \left((1 + \lambda_q \tau)^{-\Delta T/\tau} - \beta_q \right) \mathcal{T}(\beta_q) \zeta_q^{n-1}, \tag{16}$$

where $\beta_q := (1 + \lambda_q \Delta T)^{-1}$ and $\mathcal{T}(\beta) := \left\{ \beta^{j-i-1} \text{ if } j > i,\, 0 \text{ otherwise} \right\}$ is a Toeplitz matrix of size $\hat{k}$. Applying (16) recursively we obtain:

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

where:

$$\rho_n^q := \left\| \left((1 + \lambda_q \tau)^{-\Delta T/\tau} - \beta_q \right)^n \mathcal{T}^n(\beta_q) \right\|_{L^\infty}. \tag{17}$$

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

$$|\,(1 + \lambda_q \tau)^{-\Delta T/\tau} - \beta_q| \le |e^{-\lambda_q \Delta T} - \beta_q| = |e^{1-1/\beta_q} - \beta_q|, \tag{18}$$

which yields:

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

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

$$\sup_{0<\beta<1} \left| e^{1-1/\beta} - \beta \right|^n \left| \frac{1-\beta^{\hat{k}-1}}{1-\beta} \right|^n \le \left(\sup_{0<\beta<1} \frac{e^{1-1/\beta} - \beta}{1-\beta} \right)^n \approx 0.2984^n,$$

and the maximum is attained around $\beta_* = 0.358$, independently of n and $\hat{k}$ (β_* presents slight variation for $1 \le n$ and $6 \le \hat{k}$, cases of practical interest).

4 Numerical Experiments

The optimal control problem we consider involves the 1D-heat equation:

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

with boundary conditions $z(t,0) = z(t,1) = 0$ for $t \in [0,1]$, and initial data $z(0,x) = 0$ for $x \in [0,1]$. The control variable $v(\cdot)$ corresponds to the forcing term, and the target function is the nodewise interpolation of the function $\hat{y}(t,x) = x(1-x)e^{-x}$. We choose a tolerance $tol \leq 10^{-6}$ for the left preconditioned MINRES.

Table 1 lists the value of $(\gamma_{\max} - 1)$ for different values of τ and n. The results confirm Remark 1. Table 2 lists the number of MINRES iterations as ΔT and τ vary while $(\Delta T/\tau)$ remains constant. Choosing $n = 2, 4, 7$ iterations for the Parareal, the number of iterations for the MINRES basically remains constant when h or τ are refined, and so the results indicate scalability. Table 3 lists the number of MINRES iterations for $n = 2$ and $\tau = (1/512)$ for different values of $(\Delta T/\tau)$. It indicates also scalability with respect to ΔT. Like in [4], we observe numerically that the number of MINRES iterations grows logarithmically with respect to $1/r$.

Table 1. Values of $\gamma_{max} - 1$ when τ is refined. Parameters $h = 1/10$ and $\Delta T = 1/20$.

$n \setminus \hat{l}$	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

Table 2. 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

References

[1] M. J. Gander and S. Vandewalle. On the super linear and linear convergence of the parareal algorithm. In *Domain Decomposition Methods in Science and*

Table 3. 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

Engineering XVI, volume 55 of *Lect. Notes Comput. Sci. Engrg.*, pages 291–298. Springer, Berlin, 2007.

[2] J. L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag, Berlin-Heidelberg-New York, 1971.

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

[4] 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, *High Performance Computing for Computational Science - VECPAR 2006*, pages 452–465. Springer, Lect. Notes Comput. Sci., 4395, 2007.

[5] C. C. Paige and M. A. Saunders. Solutions of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12(4):617–629, 1975.