Report no. 12/11

Robust Iterative Solution of a Class of Time-Dependent Optimal Control Problems

John W. Pearson, Martin Stoll and Andrew J. Wathen

The fast iterative solution of optimal control problems, and in particular PDE-constrained optimization problems, has become an active area of research in applied mathematics and numerical analysis. In this paper, we consider the solution of a class of time-dependent PDE-constrained optimization problems, specifically the distributed control of the heat equation. We develop a strategy to approximate the $(1, 1)$-block and Schur complement of the saddle point system that results from solving this problem, and therefore derive a block diagonal preconditioner to be used within the MINRES algorithm. We present numerical results to demonstrate that this approach yields a robust solver with respect to step-size and regularization parameter.

Oxford University Mathematical Institute
Numerical Analysis Group
24-29 St Giles'
Oxford, England OX1 3LB
E-mail: john.pearson@maths.ox.ac.uk

July, 2012

1 Introduction

A recently developing area of research in the area of optimal control is the construction of preconditioners for time-dependent PDE-constrained optimization problems [3, 9, 10, 11]. It is desirable to design these preconditioners such that the computation time of the corresponding iterative method grows close to linearly with the problem size, although constructing methods as such often results in a lack of robustness with respect to the regularization term involved in the formulation of the problem. One important such problem is the (distributed) optimal control of the heat equation

$$
\begin{aligned}
\min_{y,u} \quad & J(y,u), && (1.1)\\
\text{s.t} \quad y_t - \nabla^2 y &= u, \quad \text{for } (\mathbf{x},t) \in \Omega \times [0,T],\\
y &= f, \quad \text{on } \partial\Omega,\\
y &= y_0, \quad \text{at } t=0,
\end{aligned}
$$

where y and u denote the *state* and *control*, on some space domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, and some time interval $[0,T]$. Here $J(y,u)$ is some functional that takes into account a norm of the quantity $y - \bar{y}$ (where $\bar{y}$ is some *desired state*), as well as a norm of u. In [7], the authors considered the iterative solution of the problem (1.1) with

$$J(y,u) = \frac{1}{2}\int_0^T \int_{\Omega_1} (y(\mathbf{x},t) - \bar{y}(\mathbf{x},t))^2 \ \mathrm{d}\Omega_1 \mathrm{d}t + \frac{\beta}{2}\int_0^T \int_{\Omega} (u(\mathbf{x},t))^2 \ \mathrm{d}\Omega \mathrm{d}t,$$

where $\Omega_1 = \Omega$ or $\Omega_1 \subset \Omega$, and $\mathbf{x}$ and t denote space and time variables. The authors constructed a solver that was robust with respect to step-size h and *regularization parameter* (or *Tikhonov parameter*) β.

In this paper, we demonstrate that a robust preconditioner may also be constructed for the problem (1.1) with

$$J(y,u) = \frac{1}{2}\int_{\Omega_1} (y(\mathbf{x},T) - \bar{y}(\mathbf{x}))^2 \ \mathrm{d}\Omega_1 + \frac{\beta}{2}\int_0^T \int_{\Omega} (u(\mathbf{x},t))^2 \ \mathrm{d}\Omega \mathrm{d}t,$$

which is similar to the previous problem, except the quantity $y - \bar{y}$ is now only measured at the final time. Discretizing this problem with equal-order finite element basis functions for y, u, and the *adjoint variable* p, results in the matrix system [10]

$$\begin{bmatrix} \mathcal{M}_1 & 0 & \mathcal{K}^T \\ 0 & \beta\tau\mathcal{M}_{1/2} & -\tau\mathcal{M} \\ \mathcal{K} & -\tau\mathcal{M} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \tau\mathcal{M}_1\bar{\mathbf{y}} \\ \mathbf{0} \\ \mathbf{d} \end{bmatrix}, \tag{1.2}$$

where

$$\mathcal{K} = \begin{bmatrix} M+\tau K & & & & \\ -M & M+\tau K & & & \\ & \ddots & \ddots & & \\ & & -M & M+\tau K & \\ & & & -M & M+\tau K \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} M\mathbf{y}_0 + \mathbf{f} \\ \mathbf{f} \\ \vdots \\ \mathbf{f} \\ \mathbf{f} \end{bmatrix},$$

with $\mathcal{M} = \text{blkdiag}\,(M, M, ..., M, M)$, $\mathcal{M}_{1/2} = \text{blkdiag}\left(\frac{1}{2}M, M, ..., M, \frac{1}{2}M\right)$ and $\mathcal{M}_1 = \text{blkdiag}\,(0, 0, ..., 0, M_1)$. Here, $\mathbf{y}$, $\mathbf{u}$ and $\mathbf{p}$ are the vectors containing the values of y, u and p, at each time-step in turn. The vectors $\bar{\mathbf{y}}$, $\mathbf{y}_0$ and $\mathbf{f}$ correspond to the desired state, initial condition and boundary condition respectively, the matrices M and K each denote the standard finite element mass matrix and stiffness matrix on Ω, with M_1 the equivalent mass matrix on Ω_1, and τ denotes the time-step used.

We emphasise that the total size of the matrix system (1.2) is $3nN_t \times 3nN_t$, where n is the size of M and K, and N_t is the number of time-steps used. It is desirable to devise a solution scheme such that one only needs to store the $n \times n$ matrices M and K, rather than any larger blocks.

We will consider the solution of the matrix system (1.2) for the remainder of this paper.

2 Preconditioning the Matrix System

For the majority of this section, we consider the case $\Omega_1 = \Omega$, in which case $M_1 = M$ – we will return to the case $\Omega_1 \subset \Omega$ in Section 2.1. The observation that we use to motivate the structure of our preconditioner for this problem is that if

$$\mathcal{A} = \left[\begin{array}{cc} A & B^T \\ B & 0 \end{array}\right], \quad \mathcal{P} = \left[\begin{array}{cc} A & 0 \\ 0 & BA^{-1}B^T \end{array}\right], \tag{2.1}$$

where $A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{p\times m}$, with $p \leq m$, and $\mathcal{A}$ is non-singular, then $\lambda(\mathcal{P}^{-1}\mathcal{A}) \in \left\{1, \frac{1\pm\sqrt{5}}{2}\right\}$ [5]. Matrices of the form $\mathcal{A}$ are generally referred to as *saddle point systems*, a comprehensive review of which is given in [2]. Therefore an appropriate Krylov subspace method for solving $\mathcal{A}$ should converge in 3 iterations when it is preconditioned with $\mathcal{P}$. The quantity $S := BA^{-1}B^T$ is referred to as the (negative) *Schur complement* of $\mathcal{A}$.

We note that the matrix in (1.2) has the saddle point structure of $\mathcal{A}$, with $A = \text{blkdiag}\left(\tau\mathcal{M}_1, \beta\tau\mathcal{M}_{1/2}\right)$, $B = \left[\begin{array}{cc} \mathcal{K} & -\tau\mathcal{M} \end{array}\right]$. We wish to create a preconditioner for $\mathcal{A}$ of the form $\mathcal{P}$ in (2.1). The first issue that we have is that the $(1,1)$-block A is not invertible. We therefore consider the heuristic (also discussed in [7, 10]) of approximating $\mathcal{A}$ by $\left[\begin{array}{cc} \widetilde{A} & B^T \\ B & 0 \end{array}\right]$, and create an effective preconditioner for this matrix instead. We will demonstrate in the numerical results that this also results in a good preconditioner for $\mathcal{A}$ for the problem we consider. Here $\widetilde{A}$ is a perturbation of A given by $\widetilde{A} = \text{blkdiag}\left(\tau\mathcal{M}_1^{\gamma}, \beta\tau\mathcal{M}_{1/2}\right)$, where $\mathcal{M}_1^{\gamma} = \text{blkdiag}(\gamma M, \gamma M, ..., \gamma M, M)$ for some $0 < \gamma \ll 1$.

We therefore search for a preconditioner of the form $\widehat{\mathcal{P}} = \text{blkdiag}\left(\widehat{A}, \widehat{S}\right)$, where $\widehat{A}$ and $\widehat{S}$ are approximations to $\widetilde{A}$ and $\widetilde{S} := B\widetilde{A}^{-1}B^T$ respectively. We recommend taking $\widehat{A} = \text{blkdiag}\left(\widehat{\mathcal{M}}_1, \beta\tau\widehat{\mathcal{M}}_{1/2}\right)$, where $\widehat{\mathcal{M}}_1$ and $\widehat{\mathcal{M}}_{1/2}$ are approximations to $\mathcal{M}_1^{\gamma}$ and $\mathcal{M}_{1/2}$ respectively. To approximate mass matrices, one may apply the Chebyshev semi-iteration method for consistent mass matrices (as discussed in [13]), or diagonal solves for lumped mass matrices.

We now turn our attention to approximating the Schur complement $\widetilde{S}$ of the perturbed matrix system – this is given by

$$\tilde{S} := \mathcal{K}\widehat{\mathcal{M}}_1^{-1}\mathcal{K}^T + \frac{\tau}{\beta}\mathcal{M}\mathcal{M}_{1/2}^{-1}\mathcal{M} = \mathcal{K}\widehat{\mathcal{M}}_1^{-1}\mathcal{K}^T + \frac{\tau}{\beta}\Theta\widehat{\mathcal{M}}_1^{-1}\Theta,$$

where $\Theta = \mathrm{blkdiag}\left(\sqrt{2\gamma}M, \sqrt{\gamma}M, ..., \sqrt{\gamma}M, \sqrt{2}M\right)$. Writing $\widetilde{S}$ in this way places it into the framework of Theorem 1 below, which was proved in [7].

Theorem 1 *If $\widetilde{S}$ and $\widehat{S}$ are of the form*

$$\widetilde{S} = \mathcal{K}\Phi_1^{-1}\mathcal{K}^T + \frac{1}{\beta}\Phi_2\Phi_1^{-1}\Phi_2, \quad \widehat{S} = \left(\mathcal{K} + \frac{1}{\sqrt{\beta}}\Phi_2\right)\Phi_1^{-1}\left(\mathcal{K} + \frac{1}{\sqrt{\beta}}\Phi_2\right)^T,$$

with Φ_1, Φ_2 symmetric positive definite and $\Phi_1^{-1}\Phi_2 = blkdiag(\alpha_1 I, \alpha_2 I, ..., \alpha_{N_t-1}I, \alpha_{N_t}I)$ where $I \in \mathbb{R}^{n\times n}$ and $\alpha_i > 0$, $i = 1, ..., N_t$, then $\lambda(\widehat{S}^{-1}\widetilde{S}) \in \left[\frac{1}{2}, 1\right)$. □

We may apply Theorem 1 with $\Phi_1 = \widehat{\mathcal{M}}_1$ and $\Phi_2 = \sqrt{\tau}\Theta$, as these choices of Φ_1 and Φ_2 satisfy all the necessary conditions of the theorem, to obtain that

$$\widehat{S} = \left(\mathcal{K} + \sqrt{\frac{\tau}{\beta}}\Theta\right)\widehat{\mathcal{M}}_1^{-1}\left(\mathcal{K} + \sqrt{\frac{\tau}{\beta}}\Theta\right)^T$$

satisfies $\lambda(\widehat{S}^{-1}\tilde{S}) \in \left[\frac{1}{2}, 1\right)$. We therefore conclude that $\widehat{S}$ is a good approximation to $\tilde{S}$.

To summarise, we have motivated the preconditioner

$$\widehat{\mathcal{P}} = \begin{bmatrix} \widehat{\mathcal{M}}_1 & 0 & 0 \\ 0 & \beta\tau\widehat{M}_{1/2} & 0 \\ 0 & 0 & \left(\mathcal{K} + \sqrt{\frac{\tau}{\beta}}\Theta\right)\widehat{\mathcal{M}}_1^{-1}\left(\mathcal{K} + \sqrt{\frac{\tau}{\beta}}\Theta\right)^T \end{bmatrix}$$

for the matrix system (1.2). As our preconditioner is symmetric positive definite, it may be used with the MINRES algorithm [6]. In practice we use a multigrid process to approximate the matrices $\mathcal{K} + \sqrt{\frac{\tau}{\beta}}\Theta$ and its transpose rather than taking exact solves, in order to obtain an optimal preconditioner. We present numerical results to validate the effectivess of this preconditioner in the next section, using algebraic multigrid (AMG) with these matrices.

2.1 The Subdomain Case – $\Omega_1 \subset \Omega$

We also find that our approach can generate an effective method for the case $\Omega_1 \subset \Omega$, where we use lumped mass matrices. In this case, we may again approximate the matrix $\mathcal{A}$ by the perturbation $\begin{bmatrix} \widetilde{A} & B^T \\ B & 0 \end{bmatrix}$, where B is given as before and $\widetilde{A} =$

blkdiag$(\mathcal{M}_1^{(1,\gamma)}, \beta\tau\mathcal{M}_{1/2})$, where $\mathcal{M}_1^{(1,\gamma)} = \text{blkdiag}(\gamma M, ..., \gamma M, M_1^{(1,\gamma)})$, with $M_1^{(1,\gamma)}$ a diagonal matrix equal to M_1, except with the zero entries replaced by the corresponding entries of M multiplied by γ.

We then try to precondition this perturbed matrix as before. It is easy to invert the diagonal matrix $\widetilde{A}$, so our main challenge is to approximate the Schur complement of the perturbed system:

$$\widetilde{S} = \mathcal{K}(\mathcal{M}_1^{(1,\gamma)})^{-1}\mathcal{K}^T + \frac{\tau}{\beta}\mathcal{M}\mathcal{M}_{1/2}^{-1}\mathcal{M}.$$

Using similar reasoning as in [7], we approximate $\widetilde{S}$ by

$$\widetilde{S} = (\mathcal{K} + \widehat{\mathcal{M}})(\mathcal{M}_1^{(1,\gamma)})^{-1}(\mathcal{K} + \widehat{\mathcal{M}})^T,$$
$$\widehat{\mathcal{M}} = \sqrt{\frac{\tau}{\beta}}\text{blkdiag}(\sqrt{2\gamma}M, \sqrt{\gamma}M, ..., \sqrt{\gamma}M, \sqrt{2}\bar{M}_1^{(1,\gamma)}),$$

where $\bar{M}_1^{(1,\gamma)} = (MM_1^{(1,\gamma)})^{1/2}$. We again use AMG to approximate $\widehat{S}^{-1}$ in practice.

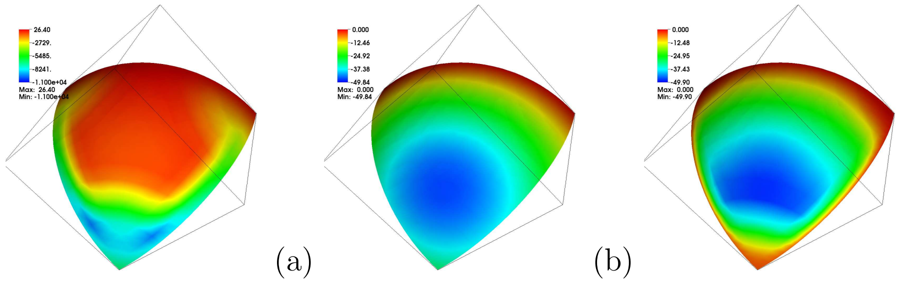
It is not as simple to prove eigenvalue bounds for $\widehat{S}^{-1}\widetilde{S}$ as for the case $\Omega_1 = \Omega$, however we find that our approximation is still an effective one. We demonstrate this with numerical results in the following section.

3 Numerical Results

In order to demonstrate the validity of our approach in practice, we present numerical results for a particular test problem. This problem is of the form (1.1) with Ω taken as the unit cube, homogeneous Dirichlet boundary conditions (i.e. $f = 0$) imposed, initial condition given by $y_0 = 0$, and $\bar{\mathbf{y}} = -64x_0 \exp\left(-\left((x_0 - 0.5)^2 + (x_1 - 0.5)^2 + (x_2 - 0.5)^2\right)\right)$ being the desired state, where $\mathbf{x} = \begin{bmatrix} x_0 & x_1 & x_2 \end{bmatrix}^T$. We used values $T = 1$ and $\tau = 0.05$, therefore taking 20 time-steps.

To obtain our results, we used the MINRES algorithm as written in Chapter 2 of [4], with the stopping criterion being the reduction of the relative pseudo-residual by a factor of 10^{-4}. The experiments were conducted within the deal.II [1] framework, with AMG applied using the Trilinos ML package [12]. We used 2 AMG V-cycles where appropriate, with 10 Chebyshev smoothing steps. For these results we considered **Q1** finite elements, with all mass matrices lumped. The results were generated using a Centos Linux machine with Intel(R) Xeon(R) CPU X5650 @ 2.67GHz CPUs and 48GB of RAM.

In Figure 1 and Table 1, we show results for $\Omega_1 = \Omega$. We used the value $\gamma = \tau\beta$, a choice motivated in [7] – we found that this selection of γ led to better performance than simply choosing a small fixed value. In this case, the MINRES algorithm converged in a small number of iterations for all cases considered. Our preconditioner thus demonstrated robustness with respect to h, β and τ, verifying our theoretical analysis. In Figure 2 and Table 2, we present results for the same problem but with $\Omega_1 = [0,1] \times [\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$ (using $\gamma = \tau\beta h^2$ as motivated in [7]) – we observed rapid convergence in this case as well, though without total parameter-independence. In each case we considered the values $\beta \in \{10^{-2}, 10^{-4}, 10^{-6}\}$, which embrace most values of interest within major applications of PDE-constrained optimization.

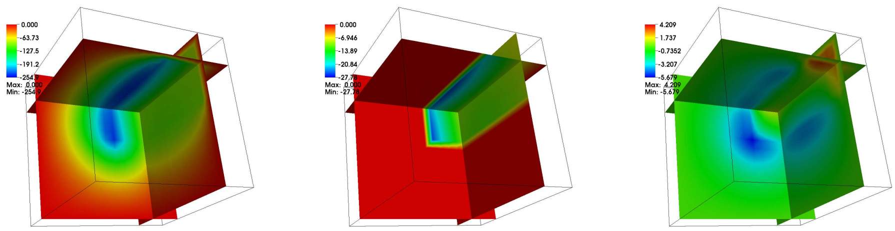


(a) (b) (c)

Figure 1: Plots of (a) control, (b) desired state, and (c) state, at the final time-step with $\beta = 10^{-4}$.

DoF	$\beta = 10^{-2}$	$\beta = 10^{-4}$	$\beta = 10^{-6}$
4913	14(9)	13(9)	9(6)
35937	16(72)	13(59)	12(55)
274625	16(624)	15(677)	13(531)

Table 1: Number of MINRES iterations and computation times to solve the linear system for various node numbers and β values. Also included are the number of degrees of freedom, in space only.



(a) (b) (c)

Figure 2: Plots of (a) control, (b) desired state, and (c) state, at the penultimate time-step with $\beta = 10^{-4}$.

4 Conclusion

In this paper, we have presented an approach for preconditioning the matrix system arising from a class of time-dependent PDE-constrained optimization problems. The method was devised in such a way as to achieve convergence of the MINRES algorithm in a

DoF	$\beta = 10^{-2}$	$\beta = 10^{-4}$	$\beta = 10^{-6}$
4913	17(5)	22(7)	14(5)
35937	18(40)	26(56)	22(50)
274625	18(482)	26(675)	28(706)

Table 2: Number of MINRES iterations and computation times to solve the linear system for various node numbers and β values.

small number of steps for all values of step-size h and regularization parameter β, as well as time-step τ. The approach utilized saddle point theory, efficient representation of the inverse of a mass matrix, an effective Schur complement approximation, and a multigrid process to enact this approximation. We have demonstrated that our method does indeed yield a solver robust to changes of h and β. We emphasise that its implementation requires the storage of only two (relatively small) matrices, a mass matrix and a stiffness matrix, in contrast to a direct method. We believe our approach could also be used to tackle a number of possible extensions of this work, for instance boundary control problems or problems with additional inequality constraints on the state or control.

Acknowledgements. The first author was supported for this work by the Engineering and Physical Sciences Research Council (UK), Grant EP/P505216/1.

References

[1] W. Bangerth, R. Hartmann, and G. Kanschat, ACM Transactions on Mathematical Software, 2007.

[2] M. Benzi, G. H. Golub, and J. Liesen, Acta Numerica, 2010.

[3] M. Benzi, E. Haber, and L. Taralli, to appear in Advances in Computational Mathematics, 2010.

[4] H. C. Elman, D. J. Silvester, and A. J. Wathen, Oxford University Press, New York, 2005.

[5] M. F. Murphy, G. H. Golub, and A. J. Wathen, SIAM Journal on Scientific Computing, 2000.

[6] C. C. Paige, and M. A. Saunders, SIAM Journal on Numerical Analysis, 1975.

[7] J. W. Pearson, M. Stoll, and A. J. Wathen, submitted to SIAM Journal on Matrix Analysis and Applications, 2011.

[8] J. W. Pearson, and A. J. Wathen, to appear in Numerical Linear Algebra with Applications, 2010.

[9] A. Potschka, M. S. Mommer, J. P. Schlöder, and H. G. Bock, Interdisciplinary Center for Scientific Computing, Heidelberg University, 2010.

[10] M. Stoll, and A. J. Wathen, submitted to Journal of Computational Physics, 2010.

[11] M. Stoll, and A. J. Wathen, submitted to Journal of Computational Physics, 2011.

[12] M. W. Gee, C. M. Siefert, J. J. Hu, R. S. Tuminaro, and M. G. Sala, Technical Report SAND2006-2649, Sandia National Laboratories, 2006.

[13] A. J. Wathen, and T. Rees, Electronic Transactions on Numerical Analysis, 2009.