



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

On Symmetric Positive Definite Preconditioners for Multiple Saddle-Point Systems

Citation for published version:

Pearson, JW & Potschka, A 2023, 'On Symmetric Positive Definite Preconditioners for Multiple Saddle-Point Systems', *IMA Journal of Numerical Analysis*. <<https://arxiv.org/abs/2106.12433>>

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

IMA Journal of Numerical Analysis

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



ON SYMMETRIC POSITIVE DEFINITE PRECONDITIONERS FOR MULTIPLE SADDLE-POINT SYSTEMS

JOHN W. PEARSON* AND ANDREAS POTSCHKA†

Abstract. We consider symmetric positive definite preconditioners for multiple saddle-point systems of block tridiagonal form, which can be applied within the MINRES algorithm. We describe such a preconditioner for which the preconditioned matrix has only two distinct eigenvalues, 1 and -1 , when the preconditioner is applied exactly. We discuss the relative merits of such an approach compared to a more widely studied block diagonal preconditioner, specify the computational work associated with applying the new preconditioner inexactly, and survey a number of theoretical results for the block diagonal case. Numerical results validate our theoretical findings.

Key words. Preconditioning; Multiple saddle-point systems; Krylov subspace methods; MINRES

AMS subject classifications. 65F08, 65F10, 65N99, 49M41

1. Introduction. Multiple saddle-point systems and preconditioning strategies for their efficient numerical solution have attracted wide interest of late (see [1, 3, 16, 27], for instance). In this paper we consider multiple saddle-point systems of the following block tridiagonal form:

$$\mathcal{A}_k = \begin{pmatrix} A_0 & B_1^T & O & \dots & O & O \\ B_1 & -A_1 & B_2^T & \ddots & & O \\ O & B_2 & A_2 & B_3^T & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & O \\ O & O & \ddots & B_{k-1} & (-1)^{k-1}A_{k-1} & B_k^T \\ O & O & \dots & O & B_k & (-1)^k A_k \end{pmatrix}, \quad (1.1)$$

with A_0 symmetric positive definite, and A_1, A_2, \dots, A_k symmetric positive semi-definite. These conditions are not necessary to guarantee invertibility of \mathcal{A}_k , but are borne out of the structures of the preconditioners we wish to consider. In particular, one popular choice of preconditioner for \mathcal{A}_k is the block diagonal matrix [16, 23, 27]

$$\mathcal{P}_{D,k} = \begin{pmatrix} A_0 & O & \dots & O & O \\ O & S_1 & \ddots & & O \\ \vdots & \ddots & S_2 & \ddots & \vdots \\ O & & \ddots & \ddots & O \\ O & O & \dots & O & S_k \end{pmatrix}, \quad (1.2)$$

where $S_1 = A_1 + B_1 A_0^{-1} B_1^T$, $S_j = A_j + B_j S_{j-1}^{-1} B_j^T$ for $j = 2, \dots, k$. Note the structure of the Schur complements S_1, S_2, \dots, S_k indicates they are likely to be progressively more difficult to invert or approximate as j increases.

*School of Mathematics, The University of Edinburgh, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, United Kingdom (j.pearson@ed.ac.uk)

†Institute of Mathematics, Clausthal University of Technology, Erzstr. 1, 38678 Clausthal-Zellerfeld, Germany (andreas.potschka@tu-clausthal.de)

MAIN ASSUMPTION. We assume for the forthcoming analysis that $\ker(A_j) \cap \ker(B_j^T) = \{0\}$ for $j = 1, \dots, k$, where $\ker(\cdot)$ denotes the kernel of a matrix. An important consequence of this assumption is that the Schur complements S_1, S_2, \dots, S_k are symmetric positive definite.

The preconditioner $\mathcal{P}_{D,k}$ has the attractive property that it may be applied within the MINRES algorithm [18] when solving linear systems involving \mathcal{A}_k , with a guaranteed convergence rate based on the eigenvalues of the preconditioned matrix $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$. This can be a valuable descriptor: for instance in [6, Theorem 4] and [23, Theorem 5.3] it is shown that for $k = 2$, the eigenvalues of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$ are contained within $[-\frac{1}{2}(1 + \sqrt{5}), \frac{1}{2}(1 - \sqrt{5})] \cup [2 \cos(\frac{3\pi}{7}), 2 \cos(\frac{\pi}{7})]$, approximately $[-1.618, -0.618] \cup [0.445, 1.802]$. However, the distribution of the eigenvalues is not robust with respect to k , as shown by the comprehensive analysis of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$ for $k \geq 1$ in the paper [27], where the matrices are analysed as continuous operators. Indeed under certain assumptions, specifically that A_1, A_2, \dots, A_k are zero matrices, it is shown [27, Theorem 2.6] that the eigenvalues of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$ are given by $2 \cos(\frac{2i+1}{2j+3} \pi)$, for $j = 0, 1, \dots, k$, $i = 0, 1, \dots, j$. The eigenvalues of smallest magnitude therefore tend towards zero as k increases. Hence, robust convergence of MINRES with respect to k cannot be expected when $\mathcal{P}_{D,k}$ is applied.

In this paper we propose a symmetric positive definite preconditioner for which the preconditioned matrix has only two distinct eigenvalues when it is applied exactly, independently of the number of blocks k , and which in general involves approximating the same matrices $A_0, S_1, S_2, \dots, S_k$ as does $\mathcal{P}_{D,k}$. We showcase the benefit of the new preconditioner with practical computations, where the number of approximate inversions of the final (and typically computationally most expensive) Schur complement can frequently be cut in half. We rigorously analyse the eigenvalue clustering for the case where the final Schur complement S_k is approximated. We do not present a result for the much more difficult case in which all Schur complements are approximated, which is a highly challenging task even for the simpler preconditioner $\mathcal{P}_{D,k}$ in the case $k = 2$ (see [6]). While we focus on the block tridiagonal structure (1.1), it is likely that this approach could be generalized to multiple saddle-point systems with larger bandwidth.

After introducing the new preconditioner, providing theoretical results, and discussing implementation details in Section 2, we compare its numerical properties with those of the preconditioner $\mathcal{P}_{D,k}$ in Section 3, and present results of tests on two applications from PDE-constrained optimization in Section 4.

NOTATION. Our definition of \mathcal{A}_k involves $k + 1$ matrix blocks: $k = 1$ therefore corresponds to classical (generalized) saddle-point systems, $k = 2$ to double saddle-point systems, $k = 3$ to triple saddle-point systems, and so on. Let the j -th diagonal block of \mathcal{A}_k contain n_{j-1} rows and columns, so \mathcal{A}_k has dimension $\sum_{j=0}^k n_j$. We denote by I_n an identity matrix of dimension n , and by O a zero matrix of appropriate size. As above, $\mathcal{P}_{D,k}$ is a block diagonal preconditioner of the form (1.2) defined for any integer $k \geq 1$, and $\mathcal{P}_{L,k}$ and $\mathcal{P}_{U,k}$ are analogous block lower triangular and block upper triangular matrices to be defined.

2. An Ideal Symmetric Positive Definite Preconditioner. We now suggest a modified preconditioner, defined for any integer $k \geq 1$:

$$\mathcal{P}_k = \mathcal{P}_{L,k} \mathcal{P}_{D,k}^{-1} \mathcal{P}_{U,k},$$

where

$$\mathcal{P}_{L,k} = \begin{pmatrix} A_0 & O & \dots & O & O \\ B_1 & -S_1 & \ddots & & O \\ O & B_2 & S_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & O \\ O & \dots & O & B_k & (-1)^k S_k \end{pmatrix}$$

and $\mathcal{P}_{U,k} = \mathcal{P}_{L,k}^T$. Note \mathcal{P}_k is clearly symmetric positive definite by the Main Assumption of Section 1. Furthermore the eigenvalues of $\mathcal{P}_k^{-1}\mathcal{A}_k$ have a very convenient structure, as we prove below.

2.1. Theoretical results. The following is a general result about the eigenvalues of the preconditioned matrix $\mathcal{P}_k^{-1}\mathcal{A}_k$, for our new preconditioner:

THEOREM 1. *For any multiple saddle-point system ($k \geq 1$), the preconditioned matrix $\mathcal{P}_k^{-1}\mathcal{A}_k$ has two distinct eigenvalues. Specifically, it has $\sum_{j=0}^{\lfloor k/2 \rfloor} n_{2j}$ eigenvalues equal to 1, and $\sum_{j=1}^{\lfloor (k+1)/2 \rfloor} n_{2j-1}$ eigenvalues equal to -1 .*

Proof. Straightforward linear algebra tells us that

$$\mathcal{P}_{L,k}^{-1}\mathcal{A}_k = \begin{pmatrix} I_{n_0} & S_0^{-1}B_1^T & O & \dots & O & O \\ O & I_{n_1} & -S_1^{-1}B_2^T & \ddots & & O \\ \vdots & \ddots & I_{n_2} & S_2^{-1}B_3^T & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & O \\ O & O & \dots & O & I_{n_{k-1}} & (-1)^{k-1}S_{k-1}^{-1}B_k^T \\ O & O & \dots & \dots & O & I_{n_k} \end{pmatrix}, \quad (2.1)$$

$$\mathcal{P}_{D,k}^{-1}\mathcal{P}_{U,k} = \begin{pmatrix} I_{n_0} & S_0^{-1}B_1^T & O & \dots & O & O \\ O & -I_{n_1} & S_1^{-1}B_2^T & \ddots & & O \\ \vdots & \ddots & I_{n_2} & S_2^{-1}B_3^T & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & O \\ O & O & \dots & O & (-1)^{k-1}I_{n_{k-1}} & S_{k-1}^{-1}B_k^T \\ O & O & \dots & \dots & O & (-1)^k I_{n_k} \end{pmatrix}, \quad (2.2)$$

where $S_0 := A_0$. Using (2.2), it holds that $\mathcal{P}_{U,k}^{-1}\mathcal{P}_{D,k} = (\mathcal{P}_{D,k}^{-1}\mathcal{P}_{U,k})^{-1}$ is an upper triangular matrix with block diagonal entries alternating between identity matrices (of dimensions n_0, n_2, \dots) and negative identity matrices (of dimensions n_1, n_3, \dots). Using this information, along with (2.1), tells us that $\mathcal{P}_k^{-1}\mathcal{A}_k = (\mathcal{P}_{U,k}^{-1}\mathcal{P}_{D,k})(\mathcal{P}_{L,k}^{-1}\mathcal{A}_k)$ is also block upper triangular, with the same block diagonal entries as $\mathcal{P}_{U,k}^{-1}\mathcal{P}_{D,k}$. For example, for $k = 2$,

$$\mathcal{P}_2^{-1}\mathcal{A}_2 = (\mathcal{P}_{U,2}^{-1}\mathcal{P}_{D,2})(\mathcal{P}_{L,2}^{-1}\mathcal{A}_2) = \begin{pmatrix} I_{n_0} & 2S_0^{-1}B_1^T & -2S_0^{-1}B_1^T S_1^{-1}B_2^T \\ O & -I_{n_1} & 2S_1^{-1}B_2^T \\ O & O & I_{n_2} \end{pmatrix}.$$

Reading off block diagonal entries of the triangular matrix $\mathcal{P}_k^{-1}\mathcal{A}_k$ gives the result. \square

Theorem 1 tells us that applying MINRES to solve a linear system involving the matrix \mathcal{A}_k (for any $k \geq 1$), with preconditioner \mathcal{P}_k , leads to guaranteed convergence in at most two iterations in exact arithmetic.

As highlighted above, the final Schur complement S_k may be expected to frequently be the Schur complement of most complex structure, and hence the most difficult to approximate. It is therefore illustrative to highlight a result when a suitable approximation to S_k is applied within the preconditioner.

THEOREM 2. *Let $\widehat{\mathcal{P}}_k$ be the preconditioner \mathcal{P}_k , but with S_k replaced by a symmetric positive definite approximation \widehat{S}_k . Then for any multiple saddle-point system ($k \geq 1$), the eigenvalues of the preconditioned matrix $\widehat{\mathcal{P}}_k^{-1}\mathcal{A}_k$ are equal to 1, -1 , or are within the range of eigenvalues of $(-1)^k\widehat{S}_k^{-1}S_k$.*

Proof. This result may be obtained by applying identical analysis to Theorem 1: $\mathcal{P}_{U,k}^{-1}\mathcal{P}_{D,k}$ and $\mathcal{P}_{L,k}^{-1}\mathcal{A}_k$ have the same structure, except the bottom-right entry of $\mathcal{P}_{L,k}^{-1}\mathcal{A}_k$ is given by $\widehat{S}_k^{-1}S_k$. Reading off block diagonal entries of $\widehat{\mathcal{P}}_k^{-1}\mathcal{A}_k$ gives the result, specifically that it has $\sum_{j=0}^{\lfloor(k-1)/2\rfloor} n_{2j}$ eigenvalues equal to 1, $\sum_{j=1}^{\lfloor k/2\rfloor} n_{2j-1}$ eigenvalues equal to -1 , with the remaining n_k eigenvalues given by those of $(-1)^k\widehat{S}_k^{-1}S_k$. \square

Theorem 2 highlights that the eigenvalues of the preconditioned matrix remain tightly clustered when the final Schur complement S_k is well approximated. Of course, the effective approximation of S_k is an important mathematical question, in particular to ensure that the eigenvalues of $\widehat{S}_k^{-1}S_k$ (and hence those of the preconditioned system) do not become close to zero or very large in magnitude. The design of such an approximation is a problem-specific question, and should be answered to ensure either $\widehat{\mathcal{P}}_{D,k}$ or $\widehat{\mathcal{P}}_k$ is effective in practice. See [6, Theorem 6], which indicates how the approximation of S_k , as well as that of the previous blocks, impacts the eigenvalues of the preconditioned system with the block diagonal preconditioner in the $k = 2$ case. This analysis is complicated even for smaller values of k , suggesting that the corresponding analysis of the preconditioner $\widehat{\mathcal{P}}_k$ with inexact approximations of the blocks would be an interesting and challenging research direction. We highlight that, for many problems, it is possible to design mesh-robust approximations of all Schur complements, in particular S_k , mitigating the issue of outlier eigenvalues and leading to rapid convergence when using $\widehat{\mathcal{P}}_{D,k}$ or $\widehat{\mathcal{P}}_k$: see [6, Section 5.2], [16, Section 8], [27, Section 3], and Section 4.1 of this paper, for examples of such problems.

We now make some further observations about the above results.

REMARK 1. *Applying Theorem 1 with $k = 1$ tells us that when preconditioning a classical (generalized) saddle-point system with*

$$\mathcal{P}_1 = \begin{pmatrix} A_0 & O \\ B_1 & -S_1 \end{pmatrix} \begin{pmatrix} I_{n_0} & A_0^{-1}B_1^T \\ O & -I_{n_1} \end{pmatrix} = \begin{pmatrix} A_0 & B_1^T \\ B_1 & S_1 + B_1A_0^{-1}B_1^T \end{pmatrix},$$

the preconditioned system $\mathcal{P}_1^{-1}\mathcal{A}_1$ has n_0 eigenvalues equal to 1, and n_1 eigenvalues equal to -1 . This is highlighted in [10, 17] for $A_1 = O$. Note that in this case $\mathcal{P}_{D,k}$ itself ensures that the preconditioned matrix has only three distinct eigenvalues, namely 1 and $\frac{1}{2}(1 \pm \sqrt{5})$ [15, 17]. As k increases, however, the clustering properties with the block diagonal preconditioner become less favorable, while \mathcal{P}_k is robust with respect to k when applied exactly.

REMARK 2. The matrix \mathcal{A}_k can be decomposed as $\mathcal{A}_k = \mathcal{P}_{L,k} \bar{\mathcal{P}}_{D,k}^{-1} \mathcal{P}_{U,k}$, where

$$\bar{\mathcal{P}}_{D,k} = \begin{pmatrix} A_0 & O & \dots & O & O \\ O & -S_1 & \ddots & & O \\ \vdots & \ddots & S_2 & \ddots & \vdots \\ O & & \ddots & \ddots & O \\ O & O & \dots & O & (-1)^k S_k \end{pmatrix}.$$

The structures of \mathcal{A}_k and $\mathcal{P}_k = \mathcal{P}_{L,k} \mathcal{P}_{D,k}^{-1} \mathcal{P}_{U,k}$, as well as the eigenvalue result in Theorem 1, inform us that although \mathcal{P}_k should be an excellent preconditioner for \mathcal{A}_k , it is not close to an actual block decomposition of the original matrix, even in the ideal case. Indeed, \mathcal{A}_k is an indefinite matrix, whereas \mathcal{P}_k is positive definite by design, so that it may be accommodated within MINRES. Any preconditioner which results in a very close representation of \mathcal{A}_k^{-1} would in general need to be applied within a non-symmetric iterative method such as GMRES [25].

REMARK 3. Equation (2.1) informs us that using $\mathcal{P}_{L,k}$ (or $\mathcal{P}_{U,k}$) as a preconditioner leads to a preconditioned matrix with only one distinct eigenvalue, specifically 1. A suitable non-symmetric iterative method such as GMRES could therefore be applied. However, the preconditioned matrix in this case would be derogatory: even in the most convenient setting of equally-sized blocks, the minimum polynomial would in general be $(x - 1)^{k+1}$, so convergence would not be independent of k if the preconditioner were exactly applied (see [25, Proposition 2 (4)]). Furthermore, convergence of non-symmetric iterative methods cannot in general be described using eigenvalues, so using $\mathcal{P}_{L,k}$ (or $\mathcal{P}_{U,k}$) as a preconditioner with approximations of the diagonal blocks would not give us a guaranteed convergence rate, unlike MINRES with \mathcal{P}_k or $\mathcal{P}_{D,k}$. In [7, Section 3], as well as considering $\mathcal{P}_{L,k}$ the authors examine preconditioning \mathcal{A}_k with $\bar{\mathcal{P}}_{D,k}$. In our setting, with the additional assumption that each A_j , $j = 0, 1, \dots, k$, is symmetric positive definite, the authors show that the eigenvalues of the preconditioned system are positive real; the authors also analyse situations where \mathcal{A}_k is non-symmetric, if $A_j = O$, $j = 1, \dots, k$. In contrast to the approach of this paper, however, the fact that a preconditioner of the form $\bar{\mathcal{P}}_{D,k}$ is not positive definite means that it may also not be applied within MINRES.

2.2. Implementation. To ensure a practical solver, one would not apply the inverses of the matrices S_0, S_1, \dots, S_k exactly to relevant vectors. One would instead replace these with suitable symmetric positive definite approximations $\hat{S}_0, \hat{S}_1, \dots, \hat{S}_{k-1}$, as well as \hat{S}_k as in Theorem 2, within $\mathcal{P}_{L,k}$, $\mathcal{P}_{D,k}$, $\mathcal{P}_{U,k}$, where $\hat{S}_0 := \hat{A}_0$. Let us again denote the corresponding preconditioner $\hat{\mathcal{P}}_k = \hat{\mathcal{P}}_{L,k} \hat{\mathcal{P}}_{D,k}^{-1} \hat{\mathcal{P}}_{U,k}$, which remains symmetric positive definite. The most efficient way of applying this preconditioner is to apply the inverse of

$$\underbrace{\begin{pmatrix} \hat{S}_0 & O & \dots & O & O \\ B_1 & -\hat{S}_1 & \ddots & & O \\ O & B_2 & \hat{S}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & O \\ O & \dots & O & B_k & (-1)^k \hat{S}_k \end{pmatrix}}_{\hat{\mathcal{P}}_{L,k}} \underbrace{\begin{pmatrix} I_{n_0} & \hat{S}_0^{-1} B_1^T & O & \dots & O \\ O & -I_{n_1} & \hat{S}_1^{-1} B_2^T & \ddots & \vdots \\ \vdots & \ddots & I_{n_2} & \ddots & O \\ O & & \ddots & \ddots & \hat{S}_{k-1}^{-1} B_k^T \\ O & O & \dots & O & (-1)^k I_{n_k} \end{pmatrix}}_{\hat{\mathcal{P}}_{D,k}^{-1} \hat{\mathcal{P}}_{U,k}}.$$

Given the above formulation for applying (the inverse of) $\widehat{\mathcal{P}}_k$ in practice, we make some observations about the computational cost of the preconditioner, and discuss when it is likely to be most effective compared to $\widehat{\mathcal{P}}_{D,k}$.

Computational cost of applying $\widehat{\mathcal{P}}_k^{-1}$. Examining the above decomposition of $\widehat{\mathcal{P}}_k$, we observe that the computational cost of applying the inverse preconditioner involves approximating the inverse operations of $\widehat{S}_0, \widehat{S}_1, \dots, \widehat{S}_{k-1}$ twice per MINRES iteration, and the inverse of the final Schur complement approximation \widehat{S}_k only once. Note that the block diagonal preconditioner $\widehat{\mathcal{P}}_{D,k}$ requires only one application of each block per iteration. However, as shown in Theorems 1 and 2, a key advantage of this approach is that the symmetric positive definite preconditioner $\widehat{\mathcal{P}}_k$ can lead to improved clustering of eigenvalues compared to $\widehat{\mathcal{P}}_{D,k}$, while requiring no additional applications of the most complicated Schur complement.

Parallelism. One advantage of the preconditioner $\mathcal{P}_{D,k}$ as opposed to \mathcal{P}_k studied here is that the inverse operators of the individual blocks of the preconditioner S_0, S_1, \dots, S_k can be approximated in parallel over the number of blocks k . There is no obvious way of doing this for \mathcal{P}_k , due to the block triangular matrices involved, although of course any available parallelism may of course be exploited *within* the blocks. If parallelization is not a focus of a solver, however, the preconditioner \mathcal{P}_k is likely to perform better on a multiple saddle-point system than $\mathcal{P}_{D,k}$, due to the improved clustering of the eigenvalues of the preconditioned matrix.

Potential effectiveness of \mathcal{P}_k . The preconditioner \mathcal{P}_k has two valuable properties when applied to multiple saddle-point systems: it is symmetric positive definite, which allows provable convergence bounds for MINRES based on the preconditioned matrix, and has a small number (only two) of distinct eigenvalues of this preconditioned matrix when the ‘ideal’ preconditioner is applied. It uses the same approximation steps as $\mathcal{P}_{D,k}$, specifically approximating the inverse operations of S_0, S_1, \dots, S_k , with two applications of all but the final Schur complement required per MINRES iteration. The approach is particularly effective if S_0, S_1, \dots, S_{k-1} may be well approximated in a relatively low CPU time, in particular compared to S_k . This is partly due to the fact that S_0, S_1, \dots, S_{k-1} must be approximated twice per iteration within \mathcal{P}_k (as opposed to once within $\mathcal{P}_{D,k}$), and partly because earlier Schur complements (while less complex in structure) may have a detrimental impact if poorly approximated, due to their propagation through the system. The final Schur complement S_k is frequently very complex but may have less influence on the overall quality of the preconditioner. Hence, inaccurate approximations of S_0, S_1, \dots, S_{k-1} could be amplified to a greater extent within \mathcal{P}_k than $\mathcal{P}_{D,k}$, however given their frequently simpler structure it is worthwhile to invest computational effort in approximating these terms. The gain in performance when applying \mathcal{P}_k instead of $\mathcal{P}_{D,k}$ is hence likely to be maximized for larger k , when S_k has by far the most complex structure, in which case a significant reduction in iteration numbers may be expected without a large additional computational workload per iteration.

3. Numerical Comparison with Block Diagonal Preconditioner. In order to offer a comparison between the preconditioners $\mathcal{P}_{D,k}$ and \mathcal{P}_k , we first recapitulate some known results for $\mathcal{P}_{D,k}$. For $k = 1$, that is classical (generalized) saddle-point systems, under the Main Assumption of Section 1, the eigenvalues of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$ are contained within $[-1, \frac{1}{2}(1 - \sqrt{5})] \cup [1, \frac{1}{2}(1 + \sqrt{5})]$ (see [2, Corollary 1], [22, Theorem 4], and [26, Lemma 2.2]). Further, in [6, Theorem 4] and [23, Theorem 5.3] it is shown that for $k = 2$, the bounds are $[-\frac{1}{2}(1 + \sqrt{5}), \frac{1}{2}(1 - \sqrt{5})] \cup [2 \cos(\frac{3\pi}{7}), 2 \cos(\frac{\pi}{7})]$. Below,

Table 3.1: Known theoretical results for minimum and maximum (negative and positive) eigenvalues of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$, for different values of k .

k	Min. -ve	Max. -ve	Min. +ve	Max. +ve	Eigenvalue range (to 3 d.p.)
1	$-2 \cos(\frac{\pi}{3})$	$2 \cos(\frac{3\pi}{5})$	$2 \cos(\frac{\pi}{3})$	$2 \cos(\frac{\pi}{5})$	$[-1, -0.618] \cup [1, 1.618]$
2	$-2 \cos(\frac{\pi}{5})$	$2 \cos(\frac{3\pi}{5})$	$2 \cos(\frac{3\pi}{7})$	$2 \cos(\frac{\pi}{7})$	$[-1.618, -0.618] \cup [0.445, 1.802]$
3	$-2 \cos(\frac{\pi}{7})$	$2 \cos(\frac{3\pi}{9})$	$2 \cos(\frac{3\pi}{7})$	$2 \cos(\frac{\pi}{9})$	$[-1.802, -0.347] \cup [0.445, 1.879]$

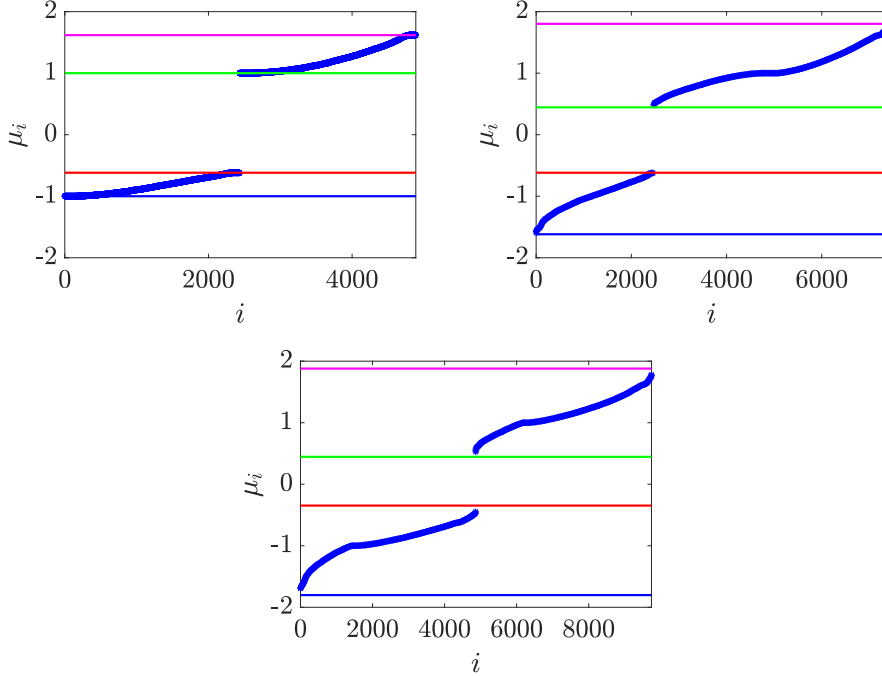


Fig. 3.1: Eigenvalues of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$ for randomly-generated examples, for $k = 1$ (top-left), $k = 2$ (top-right), and $k = 3$ (bottom). Horizontal lines correspond to theoretical bounds.

we extend this analysis to the case $k = 3$, with a proof given in Appendix A:

THEOREM 3. *Let $k = 3$, and assume that the Main Assumption of Section 1 holds. Then all eigenvalues μ of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$ satisfy:*

$$\mu \left(\mathcal{P}_{D,k}^{-1}\mathcal{A}_k \right) \in \left[-2 \cos \left(\frac{\pi}{7} \right), 2 \cos \left(\frac{5\pi}{9} \right) \right] \cup \left[2 \cos \left(\frac{3\pi}{7} \right), 2 \cos \left(\frac{\pi}{9} \right) \right],$$

which to 3 decimal places are $[-1.801, -0.347] \cup [0.445, 1.879]$.

A summary of the above results are contained in Table 3.1, noting that $1 = 2 \cos(\frac{\pi}{3})$, $\frac{1}{2}(1 + \sqrt{5}) = 2 \cos(\frac{\pi}{5})$, $\frac{1}{2}(1 - \sqrt{5}) = 2 \cos(\frac{3\pi}{5})$. Some noticeable patterns emerge in the known bounds for varying k , and it is likely these could be generalized to higher values of k with even more technical analysis. An important observation is that the eigenvalue bounds appear to spread out, in particular with the eigenvalues

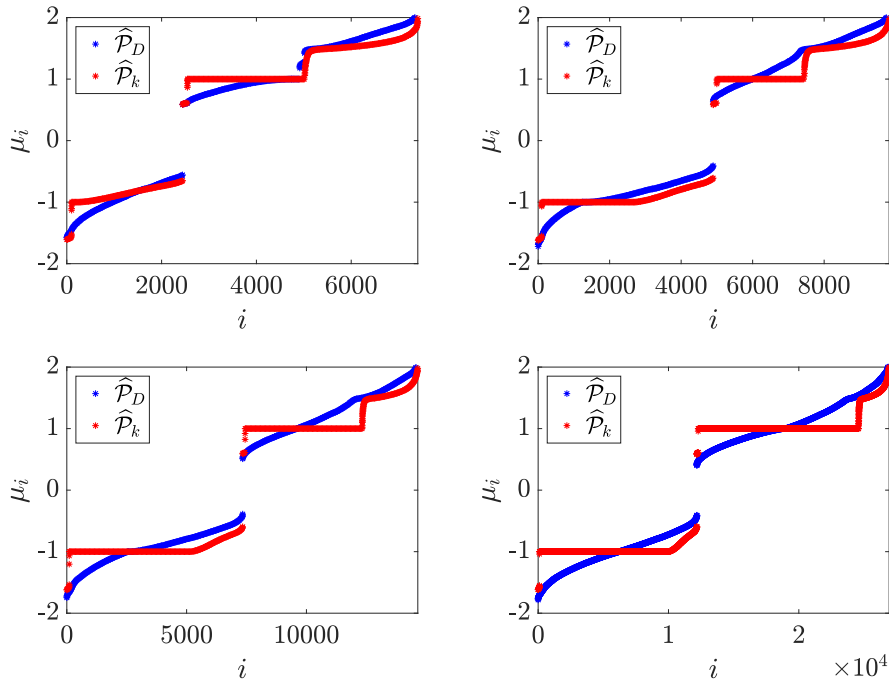


Fig. 3.2: Eigenvalues of $\widehat{\mathcal{P}}_{D,k}^{-1}\mathcal{A}_k$ (blue) and $\widehat{\mathcal{P}}_k^{-1}\mathcal{A}_k$ (red) for randomly-generated examples, for $k = 2$ (top-left), $k = 3$ (top-right), $k = 5$ (bottom-left), and $k = 10$ (bottom-right).

Table 3.2: Average MINRES iterations required for convergence, with randomly-generated examples, for a range of k .

k	1	2	3	4	5	10	15	20
Average DoF	493	754	1,007	1,247	1,498	2,738	3,977	5,240
$\widehat{\mathcal{P}}_{D,k}$ iterations	33.1	59.9	65.6	74.1	74.1	80.4	80.0	80.8
$\widehat{\mathcal{P}}_k$ iterations	30.4	34.0	35.0	34.6	34.8	34.3	33.6	33.6

of smallest magnitude moving closer to the origin, as k increases. A result of the form of Theorem 1 is therefore highly desirable when one considers the utility of the new preconditioner \mathcal{P}_k for a range of k . It is worth pointing out that the majority of the applications considered to date are of the form $k = 2$ (see [1, 3, 6, 7, 16, 23]), with interesting examples with $k = 3$ and $k = 4$ (in the latter case not block tridiagonal) recently considered in [3]. As this is a rapidly-developing field, we believe the theoretical and practical consideration of multiple saddle-point systems with higher numbers of blocks, and in particular the development of robust preconditioners, will enable and drive forward the consideration of more complex such problems.

In Figure 3.1, we show how the eigenvalues of $\mathcal{P}_{D,k}^{-1}\mathcal{A}_k$ behave in practice for a sequence of randomly-constructed examples, for $k = 1, 2, 3$. In more detail, for each value of k we compute 100 random matrices in MATLAB R2018a, and plot all

eigenvalues obtained. We select random dimensions n_j using `fix(20+10*rand)`. To construct matrices A_j , we take the symmetric part of random matrices (formed using `randn`), then add the absolute value of the smallest negative eigenvalue multiplied by an identity matrix, except for A_0 in which case we add 1.01 times this matrix to ensure positive definiteness. We also construct B_j using the `randn` function. The horizontal lines in Figure 3.1 correspond to the theoretical bounds in Table 3.1.

We also examine the behaviour of the eigenvalues of the preconditioned systems when A_0 and the Schur complements are applied inexactly within preconditioners $\widehat{\mathcal{P}}_{D,k}$ and $\widehat{\mathcal{P}}_k$. To present Figure 3.2, for a number of values of k we construct 100 random test problems as above. Letting μ_{\max} and μ_{\min} denote the maximum and minimum eigenvalues of A_0 , we set $\widehat{S}_0 := \widehat{A}_0 = \frac{1}{\mu_{\max} - \mu_{\min}} [(\frac{2}{3}\mu_{\max} - 2\mu_{\min})A_0 + \frac{4}{3}\mu_{\max}\mu_{\min}I_{n_0}]$, guaranteeing that the eigenvalues of $\widehat{A}_0^{-1}A_0$ are contained in $[\frac{1}{2}, \frac{3}{2}]$; we then set $\widehat{S}_j = A_j + B_j\widehat{S}_{j-1}^{-1}B_j^T$ for $j = 1, \dots, k$. This tests the effect of inexactness in the approximation of A_0 ‘propagating’ through the preconditioners $\mathcal{P}_{D,k}$ and \mathcal{P}_k . Figure 3.2 shows that one may frequently expect more of the eigenvalues of $\widehat{\mathcal{P}}_k^{-1}\mathcal{A}_k$ to be clustered around 1 and -1 than those of $\widehat{\mathcal{P}}_{D,k}^{-1}\mathcal{A}_k$, in particular with more small-magnitude eigenvalues further from the origin. Table 3.2 presents the average number of MINRES iterations, using the routine at [19], required to solve 100 (different) randomly-generated test problems with $\widehat{\mathcal{P}}_{D,k}$ and $\widehat{\mathcal{P}}_k$, for a range of k and with n_j now defined using `fix(200+100*rand)`. We solve to a tolerance of 10^{-10} , and refer to [9] for a justification of the stopping condition used – this criterion is motivated using backward error analysis, and measures whether the 2-norm of the current residual is as small as the specified tolerance, multiplied by a factor involving an estimation of $\|\mathcal{A}\|_2\|\widehat{x}\|_2$, where \widehat{x} denotes the approximate solution at the current iterate (see also [19]). The same stopping criterion is used for both preconditioners to ensure a fair comparison. We also report the average number of degrees of freedom (DoF’s) for each k . Although this is a synthetic test case, as one would not exactly factor and then apply Schur complement approximations in practice, the significantly smaller numbers of iterations routinely required by $\widehat{\mathcal{P}}_k$ (typically less than one half of those required by $\widehat{\mathcal{P}}_{D,k}$) further indicate the potential utility of the new preconditioner.

4. Applications and Numerical Results. As further proofs of concept of our new preconditioning strategy, we now present two applications of multiple saddle-point systems from PDE-constrained optimization, and solve them numerically using MINRES with both preconditioners $\mathcal{P}_{D,k}$ and \mathcal{P}_k , using the same approximations for A_0 and the Schur complements. All matrices are generated in Python using the package [13], whereupon we solve the resulting systems in MATLAB R2018a using the MINRES routine at [19] to tolerance 10^{-10} . All tests are carried out on an Intel(R) Core(TM) i7-6700T CPU @ 2.80GHz quad-core processor, and we report the DoF’s of the systems solved.

4.1. Double saddle-point system from PDE-constrained optimization problem. To illustrate the potential improvements of the preconditioner \mathcal{P}_k proposed in this paper as opposed to the block diagonal preconditioner $\mathcal{P}_{D,k}$, even in the case $k = 2$, we consider the discretization of the following PDE-constrained optimization

Table 4.1: Double saddle-point problem: MINRES iterations (top of each cell) and CPU time in seconds (bottom) required for convergence with preconditioner $\mathcal{P}_{D,k}$.

DoF	$h \setminus \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
867	2^{-4}	17	21	24	27	20
		0.0109	0.0137	0.0165	0.0228	0.0156
3,267	2^{-5}	17	21	22	26	18
		0.0397	0.0437	0.0446	0.0515	0.0361
12,675	2^{-6}	14	19	22	25	15
		0.118	0.153	0.172	0.193	0.118
49,923	2^{-7}	14	19	21	20	14
		0.463	0.618	0.676	0.619	0.448
198,147	2^{-8}	16	18	21	17	12
		2.04	2.29	2.62	2.16	1.54
789,507	2^{-9}	14	18	19	14	12
		7.43	9.26	9.74	7.38	6.44
3,151,875	2^{-10}	14	17	18	15	11
		31.6	37.7	39.7	33.6	25.8

problem with boundary observations:

$$\begin{aligned} \min_{u,f} \quad & \frac{1}{2} \|u - \widehat{u}\|_{L^2(\partial\Omega)}^2 + \frac{\alpha}{2} \|f\|_{L^2(\Omega)}^2 \\ \text{s.t.} \quad & \begin{cases} -\Delta u + u + f = 0 & \text{in } \Omega, \\ \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega, \end{cases} \end{aligned}$$

posed on a domain Ω with boundary $\partial\Omega$, with u and f the *state* and *control variables*, \widehat{u} a given *desired state*, and $\alpha > 0$ a given regularization parameter. This problem was examined by the authors of [16], and problems of similar structure were also considered in [27]. As in [16], we discretize this problem using H^1 -conforming piecewise linear Lagrange finite elements ϕ_i on a triangularized mesh, leading to a linear system of the form

$$\begin{pmatrix} \alpha M_h & M_h & O \\ M_h & O & L_h \\ O & L_h & Q_h \end{pmatrix} \begin{pmatrix} f_h \\ p_h \\ u_h \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \widehat{u}_h \end{pmatrix}, \quad (4.1)$$

with $M_h (= M_h^T)$ a finite element *mass matrix*, $L_h (= L_h^T)$ the sum of a *stiffness matrix* and a mass matrix, and Q_h a boundary mass matrix. Here, u_h , f_h are the discretized versions of u , f , with p_h the corresponding discretized *adjoint variable*, and \widehat{u}_h contains terms of the form $\int_{\partial\Omega} \widehat{u} \phi_i$. For our numerical tests, we consider the above problem on the domain $\Omega = (0, 1)^2$, with \widehat{u} generated by solving the forward PDE with ‘true’ control $4x(1-x) + y$, x and y denoting the spatial variables, for a range of mesh parameters h and values of α .

In the notation of multiple saddle-point systems outlined in this paper, the matrix in (4.1) leads to

$$\begin{aligned} A_0 &= \alpha M_h, & A_1 &= O, & A_2 &= Q_h, & B_1 &= B_1^T = M_h, & B_2 &= B_2^T = L_h, \\ S_1 &= \frac{1}{\alpha} M_h, & S_2 &= Q_h + \alpha L_h M_h^{-1} L_h. \end{aligned}$$

Table 4.2: Double saddle-point problem: MINRES iterations (top of each cell) and CPU time in seconds (bottom) required for convergence with preconditioner \mathcal{P}_k .

DoF	$h \backslash \alpha$	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}
867	2^{-4}	8	9	11	12	12
		0.0106	0.00900	0.00889	0.0124	0.0106
3,267	2^{-5}	8	9	9	12	9
		0.0208	0.0216	0.0216	0.0285	0.0210
12,675	2^{-6}	7	9	9	12	8
		0.0658	0.0836	0.0830	0.105	0.0728
49,923	2^{-7}	7	9	9	10	7
		0.290	0.357	0.352	0.363	0.267
198,147	2^{-8}	7	7	9	10	7
		1.11	1.10	1.36	1.46	1.08
789,507	2^{-9}	7	7	9	9	7
		4.50	4.50	5.55	5.55	4.50
3,151,875	2^{-10}	7	7	8	7	7
		19.8	19.7	22.1	19.7	19.7

Table 4.3: Double saddle-point problem: MINRES iterations required for convergence with preconditioners $\mathcal{P}_{D,k}$, \mathcal{P}_k , for varying numbers of Chebyshev semi-iterations.

h	$\mathcal{P}_{D,k}$, # Chebyshev iterations =								\mathcal{P}_k , # Chebyshev iterations =							
	1	2	3	4	5	7	10	20	1	2	3	4	5	7	10	20
2^{-4}	42	37	25	23	24	23	23	23	40	19	12	11	11	9	9	8
2^{-5}	37	34	23	23	22	22	22	22	36	17	13	11	9	9	9	8
2^{-6}	34	29	22	22	22	21	21	21	31	16	11	10	9	9	9	8
2^{-7}	30	28	21	21	21	21	21	21	27	15	11	10	9	9	9	8
2^{-8}	27	26	21	21	21	21	21	21	24	15	10	9	9	9	9	8
2^{-9}	25	23	19	19	19	19	19	19	21	13	9	9	9	9	7	7
2^{-10}	21	24	18	18	18	18	18	18	18	13	8	8	8	8	7	7

For our first tests, within \widehat{A}_0 and \widehat{S}_1 we apply 5 Chebyshev semi-iterations (see [11, 12, 29]) with Jacobi splitting to approximate the action of M_h^{-1} on a vector. To approximate S_2 we take $\widehat{S}_2 = \alpha L_h M_h^{-1} L_h$, and use 2 V-cycles of the HSL_MI20 algebraic multigrid solver [4, 5] (with 2 symmetric Gauss–Seidel iterations as pre/post-smoother) to approximate the action of L_h^{-1} on each occasion. We select these numbers of Chebyshev semi-iterations and multigrid V-cycles as we find these compute inner approximations to satisfactory accuracy, while increasing these numbers only mildly affects the outer MINRES iterations (see Table 4.3 for evidence of this for Chebyshev semi-iterations, which we note are cheap to compute compared to multigrid V-cycles). In Table 4.1, we report the numbers of preconditioned MINRES iterations required for convergence with the (approximately applied) block diagonal preconditioner $\mathcal{P}_{D,k}$, as well as CPU times (reported in seconds), for a range of h and α .¹ In Table 4.2, we report the same quantities with the (approximately applied)

¹We note that the 2-norm of the residual, scaled by the square root of the dimension of the

preconditioner \mathcal{P}_k . We observe that both the iteration numbers and CPU times are lower in each case with preconditioner \mathcal{P}_k , often by a factor of roughly two. This demonstrates the efficacy of the preconditioning approach outlined in this paper. Of course, there are many choices of approximations for the individual blocks, and the relative merits of $\mathcal{P}_{D,k}$ and \mathcal{P}_k will of course depend on these precise choices.

As emphasized above, we anticipate the benefit of \mathcal{P}_k to be more obvious when A_0, S_1 may be well approximated in a computationally cheap way, in comparison to S_2 , as any inexactness in these approximations is approximated twice within \mathcal{P}_k but only once within $\mathcal{P}_{D,k}$. Indeed we observe this feature for problem (4.1): in Table 4.3 we present iteration numbers using both preconditioners for a fixed $\alpha = 10^{-2}$, while varying the number of Chebyshev semi-iterations used to approximate M_h , to test the effect of deliberately taking a worse approximation of A_0, S_1 . We find that $\mathcal{P}_{D,k}$ converges in roughly the same number of iterations as \mathcal{P}_k when only one iteration of Chebyshev semi-iteration is applied to M_h ; this is equivalent to a scaled diagonal approximation of M_h , which is mesh-robust [30] but relatively inaccurate. When two or more Chebyshev semi-iterations are applied to M_h , MINRES converges in both fewer iterations and lower CPU time when accelerated by \mathcal{P}_k than by $\mathcal{P}_{D,k}$. We highlight that the case of taking a large number of Chebyshev semi-iterations resembles the setting of Theorem 2, in the sense that the blocks A_0, S_1 may be approximated to arbitrary accuracy, but this cannot be done for the final Schur complement S_2 in a computationally cheap way.

4.2. Quadruple saddle-point system from PDE-constrained optimization with state constraints. The presence of state constraints in PDE-constrained optimization poses severe difficulties on the analytical as well as on the numerical level (see, e.g., [14]), because in general the Lagrange multipliers are measures and thus exhibit only low regularity. We consider the problem

$$\begin{aligned} \min_{u \in H^1(\Omega), f \in L^2(\Omega)} \quad & \frac{1}{2} \|u - \widehat{u}\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|f\|_{L^2(\Omega)}^2 \\ \text{s.t.} \quad & \begin{cases} -\Delta u = f & \text{in } \Omega, \\ \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega, \\ u \leq \bar{u} & \text{a.e. in } \Omega, \end{cases} \end{aligned}$$

with a continuous upper state bound $\bar{u} \in C(\bar{\Omega})$. Instead of using a regularization approach as in [14], we introduce to the original problem an additional ‘‘lifting’’ variable $\tilde{u} \in L^2(\Omega)$ according to

$$\begin{aligned} \min_{u \in H^1(\Omega), \tilde{u} \in L^2(\Omega), f \in L^2(\Omega)} \quad & \frac{1}{2} \|u - \widehat{u}\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|f\|_{L^2(\Omega)}^2 \\ \text{s.t.} \quad & \begin{cases} -\Delta u = f & \text{in } \Omega, & | \cdot p \\ \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega, & | \cdot \tilde{p} \\ u = \tilde{u} & \text{in } \Omega, & | \cdot \tilde{p} \\ \tilde{u} \leq \bar{u} & \text{a.e. in } \Omega, & | \cdot \tilde{p} \end{cases} \end{aligned}$$

where the last two constraints are supposed to hold in the larger space $L^2(\Omega)$, so that a pointwise interpretation of necessary optimality conditions becomes possible.

problem, was smaller than 10^{-4} upon convergence for all choices of h and α , with either choice of preconditioner; this measure decreased to $\mathcal{O}(10^{-5})$ for some problems with the smallest $\alpha = 10^{-4}$, down to $\mathcal{O}(10^{-8})$ with the largest $\alpha = 1$. For 25 out of the 35 choices of h and α , a smaller residual norm was obtained at the point of convergence when using \mathcal{P}_k than with $\mathcal{P}_{D,k}$.

Table 4.4: Quadruple saddle-point problem: The outer sequential homotopy iterations are mesh-dependent. Computations were performed for a termination tolerance of 10^{-10} with $\lambda = 10^{-10}$ and direct solution of the linear subproblems.

h	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}
Seq. Hom. Iter.	4	4	7	13	26

Table 4.5: Quadruple saddle-point problem: MINRES iterations (top of each cell) and CPU time in seconds (bottom) required for convergence with preconditioner $\mathcal{P}_{D,k}$.

DoF	h	$\lambda = 10^{-7}, \alpha =$			$\lambda = 5 \times 10^{-9}, \alpha =$		
		10^{-6}	10^{-8}	10^{-10}	10^{-6}	10^{-8}	10^{-10}
9,852	2^{-5}	70	60	60	70	77 [†]	48 [†]
		0.553	0.506	0.456	0.601	0.150	0.0951
38,832	2^{-6}	72	60	60	72	52	46
		2.18	1.82	1.85	2.13	1.65	1.33
154,224	2^{-7}	85	62	62	85	51	46
		10.1	7.10	7.22	9.63	6.15	5.79
614,664	2^{-8}	110	79	77	109	57	48
		49.9	36.6	35.3	50.6	26.4	23.1

The corresponding adjoint variables to the equality constraints are $p \in H^1(\Omega)$ and $\tilde{p} \in L^2(\Omega)$. We can then apply the sequential homotopy method of [24] with the limitation that due to a norm gap for the pointwise defined projection operator (see, e.g., [28, Section 3.3]), semismoothness of the resulting non-smooth systems can only be expected for each discretized problem but not in function space. Hence, the overall nonlinear, nonsmooth outer iterations are mesh-dependent (see Table 4.4). However, the required number of outer homotopy iterations on the finest considered mesh are moderate, hence fast methods for the arising linear subsystems are of interest.

With the variable ordering $(f, p, u, \tilde{p}, \tilde{u})$, the discretized linear subproblems (after elimination of active variables and nonlinear transformation to avoid the augmented Lagrangian matrix terms) have the form

$$\begin{pmatrix} (\alpha + \lambda)M_h & -M_h & O & O & O \\ -M_h & -\frac{\lambda}{1+\rho\lambda}L_h & K_h & O & O \\ O & K_h & M_h + \lambda L_h & -M_h & O \\ O & O & -M_h & -\frac{\lambda}{1+\rho\lambda}M_h & (M_h^{(i,\cdot)})^T \\ O & O & O & M_h^{(i,\cdot)} & \lambda M_h^{(i,i)} \end{pmatrix} \begin{pmatrix} f_h \\ p_h \\ u_h \\ \tilde{p}_h \\ \tilde{u}_h^{(i)} \end{pmatrix} = b_h, \tag{4.2}$$

which is a multiple saddle-point system with $k = 4$. Here, K_h is a stiffness matrix, $M_h^{(i,i)}$ is a mass matrix defined over the inactive nodes, and $M_h^{(i,\cdot)}$ is the analogous matrix with rows corresponding to inactive nodes. We also denote $K_h^{(i,i)}$ as a stiffness matrix defined over the inactive nodes. The right-hand side is computed by finite element assembly and is nonzero except for the third block component, in which \hat{u} enters. The parameter $\lambda > 0$ arises within the sequential homotopy method as the reciprocal $\lambda = \frac{1}{\Delta t}$ of a projected backward Euler timestep of size $\Delta t > 0$ on a projected gradient/antigradient flow on the augmented Lagrangian, and $\rho > 0$ is the

Table 4.6: Quadruple saddle-point problem: MINRES iterations (top of each cell) and CPU time in seconds (bottom) required for convergence with preconditioner \mathcal{P}_k .

DoF	h	$\lambda = 10^{-7}, \alpha =$			$\lambda = 5 \times 10^{-9}, \alpha =$		
		10^{-6}	10^{-8}	10^{-10}	10^{-6}	10^{-8}	10^{-10}
9,852	2^{-5}	21 0.299	19 0.258	19 0.252	21 0.306	96 [†] 0.275	40 [†] 0.116
38,832	2^{-6}	27 1.36	20 1.04	20 1.06	27 1.33	17 0.921	17 0.867
154,224	2^{-7}	34 6.56	25 4.93	25 4.96	34 6.45	18 3.74	16 3.48
614,664	2^{-8}	54 41.2	31 24.2	31 24.6	46 35.4	24 19.0	20 16.5

standard augmented Lagrangian parameter. It is set here to a relatively low value of $\rho = 10^{-5}$ for these experiments because this ensured rapid convergence of the outer iteration. Other problems might require higher values of ρ , and both preconditioners also perform well for these values.

In the notation of this paper, the matrix (4.2) gives

$$\begin{aligned}
A_0 &= (\alpha + \lambda)M_h, & A_1 &= \frac{\lambda}{1 + \rho\lambda}L_h, & A_2 &= M_h + \lambda L_h, & A_3 &= \frac{\lambda}{1 + \rho\lambda}M_h, \\
A_4 &= \lambda M_h^{(i,i)}, & B_1 &= B_1^T = B_3 = B_3^T = -M_h, & B_2 &= B_2^T = K_h, & B_4 &= M_h^{(i,:)}, \\
S_1 &= \frac{\lambda}{1 + \rho\lambda}L_h + \frac{1}{\alpha + \lambda}M_h, & S_2 &= M_h + \lambda L_h + K_h \left[\frac{\lambda}{1 + \rho\lambda}L_h + \frac{1}{\alpha + \lambda}M_h \right]^{-1} K_h, \\
S_3 &= \frac{\lambda}{1 + \rho\lambda}M_h + M_h S_2^{-1} M_h, & S_4 &= \lambda M_h^{(i,i)} + M_h^{(i,:)} S_3^{-1} (M_h^{(i,:)})^T.
\end{aligned}$$

This is a highly complicated system, for which we build preconditioners based on approximations tailored to problems with small values of α and λ , a regime of significant numerical interest. The key objective here is to compare the preconditioners $\mathcal{P}_{D,k}$ and \mathcal{P}_k . Our approximation \hat{A}_0 involves 5 Chebyshev semi-iterations approximating the action of M_h^{-1} ; S_1 is approximated by $\frac{1}{\alpha + \lambda}M_h$, which is then handled in the same way. We approximate S_2 by neglecting terms multiplied by λ and applying the ‘matching strategy’ of [20, 21] to approximate:

$$S_2 \approx \underbrace{M_h + (\alpha + \lambda)K_h M_h^{-1} K_h}_{\tilde{S}_2} \approx \frac{1}{\sqrt{2}} \left(M_h + \sqrt{\alpha + \lambda} K_h \right) M_h^{-1} \left(M_h + \sqrt{\alpha + \lambda} K_h \right) =: \hat{S}_2.$$

Using the working of [21] we may show that the eigenvalues of $\hat{S}_2^{-1} \tilde{S}_2$ are contained within $[\frac{1}{\sqrt{2}}, \sqrt{2}]$. The factor of $\frac{1}{\sqrt{2}}$ within \hat{S}_2 is present to give symmetry about 1 (in a geometric sense) of these eigenvalue bounds. To apply \hat{S}_2^{-1} we use 2 V-cycles of the HSL_MI20 multigrid solver to approximate $M_h + \sqrt{\alpha + \lambda} K_h$ in both instances. To approximate the action of S_3^{-1} , we neglect the first term in S_3 and write $\hat{S}_3^{-1} = M_h^{-1} [M_h + (\alpha + \lambda)K_h M_h^{-1} K_h] M_h^{-1}$, where M_h^{-1} is then approximated

using 5 Chebyshev semi-iterations on three occasions. Finally, we approximate

$$\begin{aligned} S_4 &\approx M_h^{(i,:)} S_3^{-1} (M_h^{(i,:)})^T \approx M_h^{(i,:)} M_h^{-1} S_2 M_h^{-1} (M_h^{(i,:)})^T \\ &\approx \frac{1}{\sqrt{2}} \left(M_h^{(i,i)} + \sqrt{\alpha + \lambda} K_h^{(i,i)} \right) \left(M_h^{(i,i)} \right)^{-1} \left(M_h^{(i,i)} + \sqrt{\alpha + \lambda} K_h^{(i,i)} \right) =: \widehat{S}_4, \end{aligned}$$

and use 2 V-cycles of the HSL_MI20 solver to twice approximate $M_h^{(i,i)} + \sqrt{\alpha + \lambda} K_h^{(i,i)}$.² We again emphasize that other choices of approximations for the blocks may yield good results.

For the numerical computations, the matrices are obtained using H^1 -conforming piecewise linear Lagrange finite elements on a triangular mesh approximating the unit disc $\Omega = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\|_2 < 1\}$. The problem data are $\widehat{u}(\mathbf{x}) = 1 - x^2 - y^2$ and $\bar{u}(\mathbf{x}) = \frac{1}{2}$. For easier comparison, we artificially assume for the remaining numerical experiments that the current active set consists of all mesh nodes with distance at most $\frac{1}{2}$ from the origin.

In Tables 4.5 and 4.6 we report the required preconditioned MINRES iterations and CPU times with (approximately applied) preconditioners $\mathcal{P}_{D,k}$ and \mathcal{P}_k , respectively, as well as DoF's of the problems solved. We note that, once again, our new structure of preconditioner \mathcal{P}_k noticeably outperforms the block diagonal preconditioner $\mathcal{P}_{D,k}$ in terms of iteration numbers. Further, apart from two problem instances with a very coarse discretization, the CPU time is lower when \mathcal{P}_k is used. Indeed for large problems the CPU time is significantly and consistently lower. This further justifies the potential utility of this new preconditioner.

5. Conclusion. We proposed a symmetric positive definite preconditioner for block tridiagonal multiple saddle-point systems, such that the preconditioned matrix has only two distinct eigenvalues when the preconditioner is applied exactly, independently of the number of matrix blocks. In the setting where the final Schur complement is the most difficult to cheaply and accurately approximate, which often arises in practice, the approach outlined here can offer a significant advantage over the widely-used block diagonal preconditioner. Valuable future work would include the analysis of the preconditioner when individual blocks are approximated, for different values of k , and implementation of the solver for different physical applications.

Acknowledgements. The authors are grateful to two anonymous referees for their careful reading and their valuable comments. We thank Michele Benzi and Andy Wathen for their helpful feedback. JWP gratefully acknowledges support from the Engineering and Physical Sciences Research Council (UK) grant EP/S027785/1.

Appendix A. Proof of Theorem 3. We prove Theorem 3 by contradiction, through assuming that the eigenvalues μ are not contained in the intervals stated in the result. We define the notation $V \succ W$ to mean that $V - W$ is (symmetric) positive definite; similarly we define $V \succeq W$, $V \prec W$, and $V \preceq W$ to mean that $V - W$ is positive semi-definite, negative definite, and negative semi-definite, respectively.

Following the structure of the proof in [23, Theorem 5.2], we examine the associ-

²We make an exception to this strategy, indicated by \dagger in Tables 4.5 and 4.6, when coarsening failed within the HSL_MI20 routine due to the very specific parameter regime, in which case we instead approximate S_2^{-1} and S_4^{-1} with 5 Chebyshev semi-iterations applied to M_h and $M_h^{(i,i)}$, respectively.

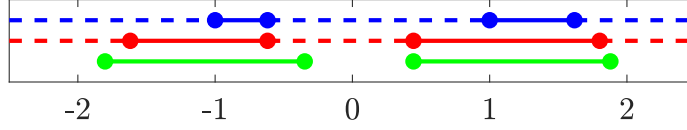


Fig. A.1: Illustration of intervals within which we allow the possibilities of eigenvalues of $\mathcal{P}_{D,3}^{-1}\mathcal{A}_3$ belonging (bold lines), as we progress through the proof, as well as intervals still to be analysed (dotted lines). Intervals are given after analysing the second block row (top, blue), third block row (middle, red), and our conclusions after analysing the fourth block row (bottom, green).

ated eigenproblem:

$$\begin{pmatrix} A_0 & B_1^T & O & O \\ B_1 & -A_1 & B_2^T & O \\ O & B_2 & A_2 & B_3^T \\ O & O & B_3 & -A_3 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \end{pmatrix} = \mu \begin{pmatrix} A_0 & O & O & O \\ O & S_1 & O & O \\ O & O & S_2 & O \\ O & O & O & S_3 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \end{pmatrix}. \quad (\text{A.1})$$

In the subsequent analysis, we progress through the second, third, and fourth block rows of (A.1), establishing conditions upon which invertibility of appropriate matrices holds, and hence upon which we may advance to the analysis of the next block row. When the relevant conditions are not satisfied, we exclude certain intervals from the analysis and allow the possibility that eigenvalues of $\mathcal{P}_{D,3}^{-1}\mathcal{A}_3$ are contained within these intervals. Finally, when analysing the fourth block row, we exclude the possibility of eigenvalues belonging to particular intervals by contradiction. See Figure A.1 for an illustration of the deductions made as we progress through the block rows of (A.1).

(First and) Second block row: The first block row of (A.1) tells us that $p\mathbf{x}_1 = -A_0^{-1}B_1^T\mathbf{x}_2$, where $p = 1 - \mu \neq 0$ by assumption. Substituting this into the second block row of (A.1), simple algebra gives that $\mathbf{x}_2 = p[qB_1A_0^{-1}B_1^T + rA_1]^{-1}B_2^T\mathbf{x}_3$ where $q = 1 + \mu - \mu^2$, $r = 1 - \mu^2$, if $\mu \in (-\infty, -1) \cup (\frac{1}{2}(1 - \sqrt{5}), 1) \cup (\frac{1}{2}(1 + \sqrt{5}), +\infty)$ in which case $qB_1A_0^{-1}B_1^T + rA_1$ is guaranteed to be invertible. We allow the possibility that eigenvalues are contained in $[-1, \frac{1}{2}(1 - \sqrt{5})] \cup [1, \frac{1}{2}(1 + \sqrt{5})]$, and hence exclude these intervals from the remaining analysis (see top line of Figure A.1).

Third block row: When $qB_1A_0^{-1}B_1^T + rA_1$ is invertible, the third block row of (A.1) tells us that

$$[-pA_2 + \Phi]\mathbf{x}_3 = B_3^T\mathbf{x}_4,$$

$$\text{where } \Phi := B_2 \left[\frac{1}{\mu}B_1A_0^{-1}B_1^T + \frac{1}{\mu}A_1 \right]^{-1} B_2^T - B_2 \left[\frac{q}{p}B_1A_0^{-1}B_1^T + \frac{r}{p}A_1 \right]^{-1} B_2^T,$$

excepting the case $\mu = 0$, as the decomposition $\mathcal{A}_k = \mathcal{P}_{L,k}\bar{\mathcal{P}}_{D,k}^{-1}\mathcal{P}_{U,k}$ guarantees the invertibility of \mathcal{A}_k . We now consider values of μ for which $-pA_2 + \Phi$ is invertible:

- If $\mu > 2 \cos(\frac{\pi}{7})$, $-pA_2$ is positive semi-definite. As $p, q, r < 0$, $\frac{1}{\mu} < \frac{q}{p}$, and $\frac{1}{\mu} < \frac{r}{p}$, we see that $\frac{1}{\mu}B_1A_0^{-1}B_1^T + \frac{1}{\mu}A_1 \prec \frac{q}{p}B_1A_0^{-1}B_1^T + \frac{r}{p}A_1$, hence Φ is positive semi-definite. Hence, $-pA_2 + \Phi$ is positive definite: any vector belonging to the nullspaces of $-pA_2$ and Φ , contradicts the positive definiteness of S_2 .

- If $\mu \in (0, 2 \cos(\frac{3\pi}{7}))$, then $p, q, r > 0$, $\frac{1}{\mu} > \frac{q}{p}$, and $\frac{1}{\mu} > \frac{r}{p}$. It follows that Φ is negative semi-definite, as is $-pA_2$. Hence, $-pA_2 + \Phi$ is negative definite.
- If $\mu \in (\frac{1}{2}(1 - \sqrt{5}), 0)$, $-pA_2$ is negative semi-definite. As $\frac{1}{\mu} < 0$, $\frac{q}{p}, \frac{r}{p} > 0$, it follows that Φ is negative definite, hence so is $-pA_2 + \Phi$.
- If $\mu < -\frac{1}{2}(1 + \sqrt{5})$, then $p > 0$, $q, r < 0$, $\frac{1}{\mu} > \frac{q}{p}$, and $\frac{1}{\mu} > \frac{r}{p}$. It follows that Φ is negative semi-definite, as is $-pA_2$. Hence $-pA_2 + \Phi$ is negative definite.

Hence $\mathbf{x}_3 = [-pA_2 + \Phi]^{-1} B_3^T \mathbf{x}_4$ when $\mu \in (-\infty, -\frac{1}{2}(1 + \sqrt{5})) \cup (\frac{1}{2}(1 - \sqrt{5}), 0) \cup (0, 2 \cos(\frac{3\pi}{7})) \cup (\cos(\frac{\pi}{7}), +\infty)$. We now exclude the intervals $[-\frac{1}{2}(1 + \sqrt{5}), \frac{1}{2}(1 - \sqrt{5})] \cup [2 \cos(\frac{3\pi}{7}), 2 \cos(\frac{\pi}{7})]$ from the remaining analysis, accepting that eigenvalues may be contained within the intervals (see middle line of Figure A.1).

Fourth block row: Within the remaining intervals under consideration, we may substitute $\mathbf{x}_3 = [-pA_2 + \Phi]^{-1} B_3^T \mathbf{x}_4$ into the fourth block row of (A.1) to give:

$$\begin{aligned} & (\mu + 1) \mathbf{x}_4^T A_3 \mathbf{x}_4 \tag{A.2} \\ & = \mathbf{x}_4^T \left(B_3 [-pA_2 + \Phi]^{-1} B_3^T - B_3 \left[\frac{1}{\mu} A_2 + B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T \right]^{-1} B_3^T \right) \mathbf{x}_4. \end{aligned}$$

The left-hand side of (A.2) is non-negative if $\mu > -1$ and non-positive if $\mu < -1$; we use this to contradict the presence of eigenvalues μ in certain intervals (note that the possibility $\mu = 0$ is already excluded):

- $\underline{\mu > 2 \cos(\frac{\pi}{9})}$: To exclude this interval we may show that the right-hand side of (A.2) is non-positive: if it is zero we may argue that $\mathbf{x}_4 \in \text{null}(B_3^T)$ and the problem reduces to the case $k = 2$ (see [6, Theorem 4] and [23, Theorem 5.3]), and if it is negative we have a contradiction. In this regime we have that $-pA_2 \succeq \frac{1}{\mu} A_2$; further, as $p < 0$, $r < q < 0$, and $\frac{1}{\mu} + \frac{p}{q} < \mu$, we have that

$$\begin{aligned} & B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T + B_2 \left[\frac{q}{p} B_1 A_0^{-1} B_1^T + \frac{r}{p} A_1 \right]^{-1} B_2^T \\ & \preceq \left(\frac{1}{\mu} + \frac{p}{q} \right) B_2 [B_1 A_0^{-1} B_1^T + A_1]^{-1} B_2^T \preceq B_2 \left[\frac{1}{\mu} B_1 A_0^{-1} B_1^T + \frac{1}{\mu} A_1 \right]^{-1} B_2^T, \end{aligned}$$

and hence $\Phi \succeq B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T$. These facts, and the result that $W \prec V \Leftrightarrow V^{-1} \prec W^{-1}$, for symmetric positive definite V, W [8, Theorem 4.5], tell us that the right-hand side of (A.2) is non-positive, as required.

- $\underline{\mu \in (0, 2 \cos(\frac{3\pi}{7}))}$: Here we have that $-pA_2 \preceq 0$, $\frac{1}{\mu} A_2 \succeq 0$, $\Phi \preceq 0$, and $B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T \succeq 0$. We thus see the right-hand side of (A.2) is non-positive, so exclude this interval for μ by the same logic as above.
- $\underline{\mu \in (2 \cos(\frac{5\pi}{9}), 0)}$: We have that $-pA_2 \succeq \frac{1}{\mu} A_2$. Further, as $p > 0$, $r > q > 0$, and $\mu - \frac{p}{q} > \frac{1}{\mu}$, we also write

$$\begin{aligned} & - \left(B_2 \left[-\frac{1}{\mu} B_1 A_0^{-1} B_1^T - \frac{1}{\mu} A_1 \right]^{-1} B_2^T + B_2 \left[\frac{q}{p} B_1 A_0^{-1} B_1^T + \frac{r}{p} A_1 \right]^{-1} B_2^T \right) \\ & \succeq \left(\mu - \frac{p}{q} \right) B_2 [B_1 A_0^{-1} B_1^T + A_1]^{-1} B_2^T \succeq B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T. \end{aligned}$$

Hence $\Phi \succeq B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T$, and so the right-hand side of (A.2) is non-positive, allowing us to exclude this interval by the logic above.

- $\mu < -2 \cos(\frac{\pi}{7})$: We have that $-pA_2 \preceq \frac{1}{\mu}A_2$. Further, as $p > 0$, $q < r < 0$, and $\frac{1}{\mu} + \frac{p}{r} > \mu$, we also write

$$\begin{aligned} & B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T + B_2 \left[\frac{q}{p} B_1 A_0^{-1} B_1^T + \frac{r}{p} A_1 \right]^{-1} B_2^T \\ & \succeq \left(\frac{1}{\mu} + \frac{p}{r} \right) B_2 [B_1 A_0^{-1} B_1^T + A_1]^{-1} B_2^T \succeq B_2 \left[\frac{1}{\mu} B_1 A_0^{-1} B_1^T + \frac{1}{\mu} A_1 \right]^{-1} B_2^T. \end{aligned}$$

Hence $\Phi \preceq B_2 [\mu B_1 A_0^{-1} B_1^T + \mu A_1]^{-1} B_2^T$, and so the right-hand side of (A.2) is non-negative, allowing us to exclude this interval by the logic above.

The result is thus proved by contradiction, through excluding the remaining intervals.

REFERENCES

- [1] F. P. Ali Beik and M. Benzi, *Iterative methods for double saddle point systems*, SIAM J. Matrix Anal. Appl. **39**, 902–921 (2018)
- [2] O. Axelsson and M. Neytcheva, *Eigenvalue estimates for preconditioned saddle point matrices*, Numer. Linear Alg. Appl. **13**, 339–360 (2006)
- [3] A. Beigl, J. Sogn, and W. Zulehner, *Robust preconditioners for multiple saddle point problems and applications to optimal control problems*, SIAM J. Matrix Anal. Appl. **41**, 1590–1615 (2020)
- [4] J. Boyle, M. Mihajlović, and J. Scott, *HSL_MI20: an efficient AMG preconditioner for finite element problems in 3D*, Int. J. Numer. Meth. Eng. **82**, 64–98 (2010)
- [5] *HSL_MI20 unsymmetric system: algebraic multigrid preconditioner*, code available at https://www.hsl.rl.ac.uk/catalogue/hsl_mi20.html
- [6] S. Bradley and C. Greif, *Eigenvalue bounds for double saddle-point systems*, arXiv preprint [arXiv:2110.13328](https://arxiv.org/abs/2110.13328) (2021)
- [7] M. Cai, G. Ju, and J. Li, *Schur complement based preconditioners for twofold and block tridiagonal saddle point problems*, arXiv preprint [arXiv:2108.08332](https://arxiv.org/abs/2108.08332) (2022)
- [8] M. Fiedler and V. Pták, *A new positive definite geometric mean of two positive definite matrices*, Linear Algebra Appl. **251**, 1–20 (1997)
- [9] D. C.-L. Fong and M. Saunders, *CG versus MINRES: An empirical comparison*, SJU Journal for Science **17**, 44–62 (2012)
- [10] P. E. Gill, W. Murray, D. B. Ponceleón, and M. A. Saunders, *Preconditioners for indefinite systems arising in optimization*, SIAM J. Matrix Anal. Appl. **13**, 292–311 (1992)
- [11] G. H. Golub and R. S. Varga, *Chebyshev semi-iterative methods, successive over-relaxation iterative methods, and second order Richardson iterative methods, Part I*, Numer. Math. **3**, 147–156 (1961)
- [12] G. H. Golub and R. S. Varga, *Chebyshev semi-iterative methods, successive over-relaxation iterative methods, and second order Richardson iterative methods, Part II*, Numer. Math. **3**, 157–168 (1961)
- [13] T. Gustafsson and G. D. McBain, *scikit-fem: a Python package for finite element assembly*, J. Open Source Softw. **5**, 2369 (2020)
- [14] K. Ito and K. Kunisch, *Semi-smooth Newton methods for state-constrained optimal control problems*, Systems Control Lett. **50**, 221–228 (2003)
- [15] Y. A. Kuznetsov, *Efficient iterative solvers for elliptic finite element problems on nonmatching grids*, Russ. J. Numer. Anal. Math. M. **10**, 187–211 (1995)
- [16] K.-A. Mardal, B. F. Nielsen, and M. Nordaas, *Robust preconditioners for PDE-constrained optimization with limited observations*, BIT Numer. Math. **57**, 405–431 (2017)
- [17] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput. **21**, 1969–1972 (2000)
- [18] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal. **12**, 617–629 (1975)
- [19] C. C. Paige, M. A. Saunders, et al., *MINRES: sparse symmetric equations*, code available at <https://web.stanford.edu/group/SOL/software/minres/>
- [20] J. W. Pearson and A. J. Wathen, *A new approximation of the Schur complement in preconditioners for PDE-constrained optimization*, Numer. Linear Alg. Appl. **19**, 816–829 (2012)

- [21] J. W. Pearson and A. J. Wathen, *Fast iterative solvers for convection–diffusion control problems*, Electron. Trans. Numer. Anal. **40**, 294–310 (2013)
- [22] J. Pearson, *Fast Iterative Solvers for PDE-Constrained Optimization Problems*, DPhil thesis, University of Oxford (2013)
- [23] J. W. Pearson and A. Potschka, *A preconditioned inexact active-set method for large-scale nonlinear optimal control problems*, arXiv preprint [arXiv:2112.05020](https://arxiv.org/abs/2112.05020) (2021)
- [24] A. Potschka and H. G. Bock, *A sequential homotopy method for mathematical programming problems*, Math. Program., **187**, 459–486 (2021)
- [25] Y. Saad and M. H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput. **7**, 856–869 (1986)
- [26] D. Silvester and A. Wathen, *Fast iterative solution of stabilised Stokes systems. Part II: using general block preconditioners*, SIAM J. Numer. Anal. **31**, 1352–1367 (1994)
- [27] J. Sogn and W. Zulehner, *Schur complement preconditioners for multiple saddle point problems of block tridiagonal form with application to optimization problems*, IMA J. Numer. Anal. **39**, 1328–1359 (2019)
- [28] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, MOS-SIAM Series on Optimization, Vol. 11., SIAM, Philadelphia, PA (2011)
- [29] A. Wathen and T. Rees, *Chebyshev semi-iteration in preconditioning for problems including the mass matrix*, Electron. Trans. Numer. Anal. **34**, 125–135 (2009)
- [30] A. J. Wathen, *Realistic eigenvalue bounds for the Galerkin mass matrix*, IMA J. Numer. Anal. **7**, 449–457 (1987)