ON THE SPECTRAL CLUSTERING OF A CLASS OF MULTIGRID PRECONDITIONERS

JOSÉ PABLO LUCERO LORCA *, CONOR MCCOID †, AND MICHAL OUTRATA ‡

Abstract. We consider an algebraic multigrid (AMG) scheme for the direct solution of complex-valued square linear systems based on a recursive 2×2 block partitioning of the coefficient matrix and study the optimal choices of its components. In particular, we complement existing results that characterize optimal choices for a nonsymmetric-cycle methods by analyzing the spectral behavior of its symmetrized variant. We analyze the error propagation operator of the two-level symmetric-cycle method by calculating its invariant subspaces and its nonzero eigenvalues that govern the behavior of the error after a single cycle. We show that the error propagation operator can be studied separately for pairs of modes, working as bases of the invariant subspaces. The main result is an explicit choice of smoothing parameters that makes all the pairs of modes respond identically, forcing the nontrivial eigenvalues of the error propagation operator to collapse to a single, a priori known value, i.e., we give an explicit choice for the smoothing parameters so that the nonzero part of the spectrum of the error propagation operator collapses to a single, a priori known point on the real line. As a consequence, we give a closed-form formula for the inverse of a general square matrix that, to the best of our knowledge, is new in the literature. Additionally, the framework provides a clear, self-contained description of an ideal AMG W-cycle, offering a concrete target for the design of related schemes. We illustrate the theory with direct applications to general matrices and with analyses of representative matrices arising in numerical methods.

1. Introduction.

Large linear algebraic systems of the form

$$L\mathbf{u} = \mathbf{b}, \quad L \in \mathbb{C}^{N\times N}, \quad \mathbf{b} \in \mathbb{C}^N, \tag{1.1}$$

are ubiquitous in scientific computing and modeling involving partial differential equations. Hence, the efficiency of solving eq. (1.1) is paramount for these applications.

Many of the most powerful solution techniques exploit some sort of hierarchy to approximate the discretized solution. Examples include multigrid methods [21, 42], domain decomposition methods with coarse correction [13], hierarchical direct solvers [22, 33], and many parallel-in-time methods [19]. Multigrid methods in particular construct such hierarchies by combining smoothing with coarse representations of the problem on each level of the hierarchy so that at different levels the method focuses on different components of the error. Their success then depends on the interaction of the method's components: the smoother, the smoothing parameters (i.e., damping and steps), the operators that transfer information between the levels of the hierarchy (i.e., restrictions and prolongations) and the coarse problems that represent the global behavior of the solution. Despite the wide use of multigrid methods, their processes are, in many cases, not understood in full detail, especially for non-Hermitian problems. In some applications, one must often rely on case-dependent analysis or numerical experimentation to obtain insight into the method's mechanism for their problem [28, 23].

We consider a standard *ansatz* for the algebraic multigrid methods (AMG), i.e., we assume L is regular and is partitioned into a 2×2 block structure with square diagonal blocks. This partition of L may arise from a physical decomposition, a coloring or ordering of the unknowns or simply from algebraic convenience; in all cases, the diagonal blocks represent subproblems that can be inverted or approximately solved efficiently. We begin by recalling that there exists a direct two-level *nonsymmetric*

*affiliation
†McMaster University, Ontario, Canada
‡Charles University, Prague, Czechia

cycle method with a *singular* smoother induced by this blocking, see [30, 26, 29]. The question is then:

> "*Under what conditions can we construct a direct two-level* symmetric *cycle method (with the same number of pre- and post-smoothing steps) with a* regular *smoother?*"

This question is practically relevant – the vast majority of problems of the form of eq. (1.1) are solved only approximately through iterative methods. Many if not most of these iterative methods are formulated as an economical modification or approximation of direct solvers. This shows that a new direct solver opens the doors for new iterative methods, for example see [30, 26]. This remains true even in the case where the direct solver itself may suffer from numerical issues. Multigrid literature shows the condition on the regularity of the smoother is a natural one, frequently requiring the smoother to be a convergent method on its own[42]; we comment on the utility of the symmetry of the cycle at the end of this section.

In this paper we give the following answer to the above question:

> "*If the Schur complement matrix pencil corresponding to the* 2×2 *blocking of* L *is diagonalizable, then we can construct such a direct two-level method.*"

In the language of the multigrid community, we take the block-Jacobi smoother induced by the blocking of L, the ideal restriction and prolongation operators used in the nonsymmetric-cycle direct method and the inherited coarse problem and then show that with these choices we can construct a direct symmetric W-cycle method.

We begin with a spectral analysis of the error propagation operator based on the smoothing parameters, under the assumption of diagonalizability of the Schur complement matrix pencil and show that the diagonalizability of this pencil ensures a decomposition of $\mathbb{C}^N$ into a direct sum of low-dimensional invariant subspaces of the error propagation operator. Analyzing its spectral properties can then be done independently in each of these subspaces, making the calculations tractable. We show that within each of these subspaces the error propagation operator has only a single nonzero eigenvalue, which depends on an eigenvalue of the pencil and on the smoothing parameters. Next, we give explicit formulas for the values of the smoothing parameters (depending on the number of pre- and post-smoothing steps) so that the nonzero eigenvalues become identical across all of the invariant subspaces, i.e., collapsing the spectrum of the error propagation operator into two *a priori* known points, one of which is zero. The derivation of this result is purely algebraic and hence independent of the origins of L, making the result relevant to a wide set of problems of the form of eq. (1.1). We then show several implications for making the two-level method direct and point out that it can be called recursively, obtaining a multilevel direct method.

If a direct multigrid (or other) method is modified or approximated, we usually use Krylov subspace methods (KSM) to ensure or enhance its convergence, i.e., we use it as a preconditioner for KSM. Classical choices include the generalized minimal residual (GMRES) method or, in the Hermitian positive definite (HPD) setting, the conjugate gradient (CG) method (see [25]). A single application of the modified/approximated multigrid method is carried out at each iteration of the KSM, thus building a more suitable Krylov subspace for the approximation of the solution compared to the unpreconditioned KSM. Importantly, if the multilevel method uses a nonsymmetric cycle and we use it as a preconditioner for an HPD problem, then CG

can no longer be used, leading to loss of efficiency. This further motivates formulating a *symmetric cycle, positive definite* direct method (for cases when L is HPD) so that even after its (suitable) approximation we can use it as a preconditioner for CG. We also note that in some cases, e.g., for CG, the convergence can be understood and predicted based on the eigenproperties of the preconditioned system. The spectral analysis of the error propagation operator of the two-level method directly yields analogous results for the preconditioned system when the two-level method is used as a preconditioner.

The rest of this paper is structured as follows: we first refer to the existing results in the field in Section 1.1 and then recall some background material in Section 2. In Section 3 we develop the spectral properties of the error propagation operator of the two-level method and in Section 4 we use these result to derive the optimal choice of the smoothing parameters for the eigenvalue clustering of the error propagation operator and show how to obtain a direct method in Section 5. We provide numerical insights into the use of the method in Section 6.

1.1. Related work. Multigrid methods have a long history, beginning with analyses of relaxation for elliptic difference equations [16, 17, 2, 4, 21, 42, 9], quickly accompanied with their algebraic counterparts – the algebraic multigrid – that evolved from the *coarse* and *fine* splitting framework, also called CF-splitting, introduced and developed in [5, 6, 38]. This viewpoint leads to a direct two-level method approach that is generally impractical but serves as the theoretical foundation for many AMG algorithms, see Algorithm 2.1 in Section 2.

The nonsymmetric AMG literature offers several approaches for adapting this two-level method to challenging problems. A natural approach is to approximate ideal restriction and prolongation operators under fixed sparsity patterns with constrained optimization techniques [36, 43]. Petrov–Galerkin smoothed aggregation introduces transfer operators adapted to nonsymmetric systems [40, 8]. AIR and nAIR algorithms [32] develop further the reduction-based AMG, showing that local approximations of ideal restriction combined with F-relaxation can yield a robust two-level method for general nonsymmetric problems. There are also adaptive strategies [30, 7] or nonsymmetric methods based on approximate ideal restriction [31].

Local Fourier Analysis (LFA) provides a tool for analyzing the multigrid convergence and smoothing behavior on structured grids, complementary to the analysis present in the above cited work. Classical LFA results [10, 4, 9] analyze the Fourier symbol of the operator and the relaxation scheme to understand the smoothing efficiency and the asymptotic two-grid contraction factor. Variants of this approach have been applied to discontinuous Galerkin methods and related discretizations, including settings where exact coarse-grid elimination leads to a direct two-level solver [27].

For the HPD setting, the question of an optimal algorithm has been considered in [15, 44], including the energy-minimizing interpolation.

2. Background. The AMG viewpoint is based on ordering the degrees of freedom (DoFs) into *fine* and *coarse* DoFs, writing L in a block form[1]:

$$L = \begin{bmatrix} L_{ff} & L_{fc} \\ L_{cf} & L_{cc} \end{bmatrix} \equiv \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

[1]To the best of our knowledge, the AMG community labels the blocks with subscripts c (coarse) and f (fine). Here, for the ease of notation as well as accessibility to a wider numerical linear algebra community, we denote these blocks $A, \ldots, D$.

with $L_{ff} \equiv A \in \mathbb{C}^{(N-n)\times(N-n)}$ and $L_{cc} \equiv D \in \mathbb{C}^{n\times n}$ being invertible. With this blocking, one can write the action of L^{-1} as a two-level scheme, simply due to the Schur complement formulation of the LDU factorization of L, see [30, 26, 18, 37, 11]. We recall this scheme in Algorithm 2.1, using the ideal prolongation, restriction and smoothing operators

$$P_{\text{ideal}} = \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}, \quad R_{\text{ideal}} = \begin{bmatrix} \star & I \end{bmatrix} \quad \text{and} \quad S^{-1}_{\text{ideal}} = \begin{bmatrix} A^{-1} & \\ & \end{bmatrix}, \tag{2.1}$$

that in turn induce the ideal coarse space operator $M_{0,\text{ideal}} := R_{\text{ideal}} L P_{\text{ideal}}$, which corresponds to the Schur complement $M_{0,\text{ideal}} \equiv D - CA^{-1}B$. Notice that Algorithm 2.1 is a direct two-level method with any choice of the first rectangular block of $R_{\text{ideal}} \in \mathbb{C}^{n\times N}$; a natural choice in the AMG literature is to replace $\star$ with $-CA^{-1}$.

Algorithm 2.1 Two-level realization of $L^{-1} : \mathbf{v} \mapsto L^{-1}\mathbf{v}$

$\mathbf{x} \leftarrow \mathbf{0}$
$\mathbf{x} \leftarrow \mathbf{x} + S^{-1}_{\text{ideal}}(\mathbf{v} - L\mathbf{x})$ ▷ pre-smoothing
$\mathbf{x} \leftarrow \mathbf{x} + P_{\text{ideal}} M^{-1}_{0,\text{ideal}} R_{\text{ideal}}(\mathbf{v} - L\mathbf{x})$ ▷ coarse correction
return $L^{-1}\mathbf{v} \leftarrow \mathbf{x}$

Algorithm 2.1 can be used recursively, inducing a direct, multilevel AMG method – provided the coarse solve is eventually computed exactly. As noted in [29], this algebraic structure can be interpreted as a nonsymmetric V-cycle with no post-smoothing and a pre-smoothing acting only on the fine DoFs governed by the block A.

We consider a symmetric adaptation of Algorithm 2.1, with m pre- and post-smoothing steps, each relaxed by a smoothing parameter $\alpha_i \in \mathbb{C}$, as summarized in Algorithm 2.2. The symmetric method is characterized by the prolongation, restriction and smoothing operators P, R, S^{-1} and the smoothing parameters $\alpha_1, \ldots, \alpha_m$ – we automatically use the algebraically induced coarse operator $M_0 := RLP$.

Algorithm 2.2 Action of a two-level multigrid preconditioner M^{-1} on $\mathbf{v} \in \mathbb{C}^N$

$\mathbf{x} \leftarrow \mathbf{0}$
for $i = 1, \ldots, m$ **do**
    $\mathbf{x} \leftarrow \mathbf{x} + \alpha_i S^{-1}(\mathbf{v} - L\mathbf{x})$ ▷ pre-smoothing
end for
$\mathbf{x} \leftarrow \mathbf{x} + P M_0^{-1} R(\mathbf{v} - L\mathbf{x})$ ▷ coarse correction with $M_0 = RLP$
for $i = 1, \ldots, m$ **do**
    $\mathbf{x} \leftarrow \mathbf{x} + \alpha_i S^{-1}(\mathbf{v} - L\mathbf{x})$ ▷ post-smoothing
end for
return $M^{-1}\mathbf{v} \leftarrow \mathbf{x}$

Aiming for a direct method with a regular smoother, we take inspiration from Algorithm 2.1 and set

$$P = \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}, \quad R = \begin{bmatrix} -CA^{-1} & I \end{bmatrix}, \quad S^{-1} = \begin{bmatrix} A^{-1} & \\ & D^{-1} \end{bmatrix}, \tag{2.2}$$

which then implies $M_0 = D - CA^{-1}B$, i.e., the coarse space operator corresponds to the Schur complement of A in L, same as in Algorithm 2.1.

We note that eq. (2.2) corresponds to taking S^{-1} as the *block Jacobi smoother*, as opposed to using just its leading principal submatrix in Algorithm 2.1 (Jacobi-type smoothers are a common choice in the wider multigrid community); we adapted the prolongation and restriction operators directly from eq. (2.1) with $\star = -CA^{-1}$, ensuring that if L is HPD then so are M^{-1} and M_0^{-1}.

To have a concrete method, it remains to specify the smoothing parameters $\alpha_1, \dots \alpha_m$, a problem treated in detail in Section 4. We conclude this section by recalling the error propagation operators of Algorithm 2.2. These are commonly used in multigrid analysis (see, e.g., [30, eq. (2) and below]).

We denote by $E^{(s)} \equiv E^{(s)}(\alpha)$ the error propagation of a single pre- or post-smoothing step and, similarly, we denote by $E^{(c)}$ the error propagation operator of the coarse space correction. Then

$$E^{(s)}(\alpha) = I - \alpha S^{-1} L, \quad \text{and} \quad E^{(c)} = I - P M_0^{-1} R L.$$

By eq. (2.2), we evaluate these operators as

$$E^{(s)}(\alpha) = \begin{bmatrix} (1-\alpha)I & -\alpha A^{-1}B \\ -\alpha D^{-1}C & (1-\alpha)I \end{bmatrix} \quad \text{and} \quad E^{(c)} = \begin{bmatrix} I & A^{-1}B \\ & \end{bmatrix}, \tag{2.3}$$

so that the full error propagation operator $E \equiv E(\alpha_1, \dots, \alpha_m)$ corresponding to Algorithm 2.2 becomes

$$E(\alpha_1, \dots, \alpha_m) = \left(\prod_{i=m}^{1} E^{(s)}(\alpha_i) \right) E^{(c)} \left(\prod_{i=1}^{m} E^{(s)}(\alpha_i) \right). \tag{2.4}$$

The error propagation operator E is linked with the preconditioned system $M^{-1}L$ by

$$M^{-1}L = I - E, \tag{2.5}$$

showcasing that any insight on the eigenproperties, the field of values or the pseudospectra of E can be directly translated to the preconditioned system $M^{-1}L$ and then used for the convergence analysis of iterative methods applied to it. In particular, clustering the spectrum of E around the origin translates to clustering of the spectrum of $M^{-1}L$ around 1.

3. Spectral properties of E. This section analyzes the spectral properties of the error propagation operator $E(\alpha_1, \dots, \alpha_m)$ of Algorithm 2.2 for general smoothing parameters. Some of the analysis follows from similarities between the error propagation operators in eq. (2.3) with the block matrices encountered in the literature for block-preconditioners of saddle-point systems, see, in particular [41, 24] but also [12, 3, 34] and references therein. Related ideas can be found in [21, §2.4] for geometric multigrid.

Consider $E^{(s)}(\alpha)$ for some $\alpha \neq 0$, let $\mu := (\lambda + \alpha - 1)/\alpha$ and assume $\mu \neq 0$. The characteristic polynomial of $E^{(s)}(\alpha)$ is given by

$$\begin{aligned} \det(\lambda I - E^{(s)}) &= \det \begin{bmatrix} \mu\alpha I & \alpha A^{-1}B \\ \alpha D^{-1}C & \mu\alpha I \end{bmatrix} = (\mu\alpha)^N \det \begin{bmatrix} I & \mu^{-1}A^{-1}B \\ \mu^{-1}D^{-1}C & I \end{bmatrix} \\ &= (\mu\alpha)^{N-n} \det\left(I_{N-n} - \mu^{-2}A^{-1}BD^{-1}C\right) = \left(\frac{\alpha}{\mu^2}\right)^{N-n} \det\left(\mu^2 I_{N-n} - T\right), \end{aligned}$$

where $T := A^{-1}BD^{-1}C$. Thus, the eigenvalues of the error propagation operator for smoothing are parametrized by those of T and by α – in the literature, an eigenpair $(\lambda, \mathbf{w})$ of T is more commonly thought of as a generalized eigenpair of the Schur complement matrix pencil $\{BD^{-1}C, A\}$, i.e., $(BD^{-1}C)\mathbf{w} = \lambda A\mathbf{w}$, but here we will work with T rather than with the pencil $\{BD^{-1}C, A\}$.

The spectrum of $E^{(c)}$ is equal to $\{0, 1\}$ and so it seems intuitive that the spectrum of E should also be parametrized by that of T. Next, we show that this intuition carries through – if T is diagonalizable, then we can essentially block-diagonalize both $E^{(s)}(\alpha)$ and $E^{(c)}$ in the same basis[2]. First, we showcase a simplified version of our argument for the case $n = N - n = N/2$ where the blocks $A, \ldots, D$ are square *and* assuming T is regular. On one hand, this leads to convenient simplifications that lay out the structure of the operators clearly; on the other hand, these assumptions are *mostly unrealistic* as we almost always have $n > N - n$. We treat the general case in analogous (though more tedious) fashion afterwards. We note that the case $n < N - n$ can be handled "by symmetry" – in the calculation above, we could use the other Schur complement formulation for the determinant, featuring the matrix $D^{-1}CA^{-1}B$ instead of T. This would lead to calculations that are analogous to the ones we carry out below, replacing the eigenpairs of T with those of $D^{-1}CA^{-1}B$ and exchanging the blocks $D^{-1}C$ and $A^{-1}B$ wherever necessary.

3.1. Invariant subspaces – the simplified case. In this section we compute the invariant subspaces of the error propagation operator assuming that $n = N/2$, where n is the size of the coarse space and N the size of the fine space, and $T \in \mathbb{C}^{N/2\times N/2}$ is regular and diagonalizable. Hence, we have a full set of $N/2$ eigenpairs $(\lambda_1, \mathbf{w}_1), \ldots, (\lambda_{N/2}, \mathbf{w}_{N/2})$, where all eigenvalues are nonzero. Let the eigendecomposition of T be $W\Lambda W^{-1}$. Define the matrix $\tilde{V}$ as

$$\tilde{V} := \begin{bmatrix} W\sqrt{\Lambda} & W\sqrt{\Lambda} \\ D^{-1}CW & -D^{-1}CW \end{bmatrix} = \begin{bmatrix} W & \\ & D^{-1}CW \end{bmatrix} \begin{bmatrix} \sqrt{\Lambda} & \sqrt{\Lambda} \\ I & -I \end{bmatrix}. \tag{3.1}$$

Since T is regular so is C and $\tilde{V}$. We can transform $E^{(s)}(\alpha)$ and $E^{(c)}$ into the basis given by the columns of $\tilde{V}$. A direct calculation gives

$$\begin{aligned} E^{(s)}\tilde{V} &= \begin{bmatrix} (1-\alpha)I & -\alpha A^{-1}B \\ -\alpha D^{-1}C & (1-\alpha)I \end{bmatrix} \begin{bmatrix} W\sqrt{\Lambda} & W\sqrt{\Lambda} \\ D^{-1}CW & -D^{-1}CW \end{bmatrix} \\ &= \begin{bmatrix} (1-\alpha)W\sqrt{\Lambda} - \alpha W\Lambda & (1-\alpha)W\sqrt{\Lambda} + \alpha W\Lambda \\ -\alpha D^{-1}CW\sqrt{\Lambda} + (1-\alpha)D^{-1}CW & -\alpha D^{-1}CW\sqrt{\Lambda} - (1-\alpha)D^{-1}CW \end{bmatrix} \\ &= \begin{bmatrix} W\sqrt{\Lambda} & W\sqrt{\Lambda} \\ D^{-1}CW & -D^{-1}CW \end{bmatrix} \begin{bmatrix} (1-\alpha-\alpha\sqrt{\Lambda})I & \\ & (1-\alpha+\alpha\sqrt{\Lambda})I \end{bmatrix}, \\ E^{(c)}\tilde{V} &= \begin{bmatrix} I & A^{-1}B \\ & \end{bmatrix} \begin{bmatrix} W\sqrt{\Lambda} & W\sqrt{\Lambda} \\ D^{-1}CW & -D^{-1}CW \end{bmatrix} = \begin{bmatrix} W(\Lambda+\sqrt{\Lambda}) & W(\Lambda-\sqrt{\Lambda}) \\ & \end{bmatrix} = \\ &= \begin{bmatrix} W\sqrt{\Lambda} & W\sqrt{\Lambda} \\ D^{-1}CW & -D^{-1}CW \end{bmatrix} \begin{bmatrix} \frac{1}{2}(I+\sqrt{\Lambda}) & \frac{1}{2}(I-\sqrt{\Lambda}) \\ \frac{1}{2}(I+\sqrt{\Lambda}) & \frac{1}{2}(I-\sqrt{\Lambda}) \end{bmatrix}. \end{aligned} \tag{3.2}$$

There are several important observations to be made about eq. (3.2). We notice that $E^{(s)}(\alpha)$ *diagonalizes* in this basis. Moreover, the eigenbasis formed by the columns of $\tilde{V}$ is *independent* of α. Therefore, $E^{(s)}(\alpha)$ and $E^{(s)}(\beta)$ commute for any

[2]We note that an analogous calculation and intuition also carries through if we work with the "complementary" matrix $D^{-1}CA^{-1}B$. Algebraically, all of the derivations that follow can be based on $D^{-1}CA^{-1}B$ with some minor adjustments.

$\alpha, \beta \in \mathbb{C}$. The error propagation operator $E^{(c)}$ does not diagonalize, but becomes a block matrix such that each block is a diagonal matrix.

Suppose the columns of $\tilde{V}$ are reorganized such that for all $\lambda \in \sigma(T)$ the vectors $D^{-1}C\mathbf{w}_\lambda$ and $-D^{-1}C\mathbf{w}_\lambda$ are next to each other. This corresponds to the permutation $\pi = [1, n/2+1, 2, n/2+2, \dots]$, the so-called *red-black* reordering. Call the new matrix $V = \tilde{V}\Pi$, where Π denotes the permutation matrix that reorders the columns according to π. Clearly, $E^{(s)}(\alpha)$ still diagonalizes in this new basis, with the entries of the diagonal matrix permuted accordingly. Moreover, $E^{(c)}$ is transformed into a block-diagonal matrix with 2×2 blocks, namely

$$V^{-1}E^{(c)}V = \frac{1}{2}\begin{bmatrix} \begin{bmatrix} 1+\sqrt{\lambda_1} & 1-\sqrt{\lambda_1} \\ 1+\sqrt{\lambda_1} & 1-\sqrt{\lambda_1} \end{bmatrix} & & \\ & \ddots & \\ & & \begin{bmatrix} 1+\sqrt{\lambda_{N/2}} & 1-\sqrt{\lambda_{N/2}} \\ 1+\sqrt{\lambda_{N/2}} & 1-\sqrt{\lambda_{N/2}} \end{bmatrix} \end{bmatrix}.$$

In other words, we have found $N/2$ two-dimensional invariant subspaces $\mathcal{V}_{\lambda_1}, \dots, \mathcal{V}_{\lambda_{N/2}}$ of the error propagation operators $E^{(c)}, E^{(s)}(\alpha)$ and thus also of any of their compositions. Therefore, we can analyze the global error propagation operator $E(\alpha_1, \dots, \alpha_m)$ defined in eq. (2.4) on the subspace $\mathcal{V} := \mathcal{V}_{\lambda_1} \oplus \dots \oplus \mathcal{V}_{\lambda_{N/2}}$ (with the basis V) using 2×2 matrices, allowing for a direct calculation. In the next section we derive analogous results in the general case but first, we briefly comment on the symbol $\sqrt{\Lambda} \equiv \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{N/2}})$ in the following remark.

REMARK 1 (Branch choice and the case $\lambda = 0$). *Throughout, the symbol $\sqrt{\lambda}$ corresponds to the principal branch of the square root, i.e.,*

$$\sqrt{\lambda} := \exp\left(\tfrac{1}{2}\log\lambda\right),$$

where log *denotes the principal complex logarithm, with branch cut on $(-\infty, 0]$ and $\arg\lambda \in (-\pi, \pi]$. None of the conclusions depend on the sign of $\sqrt{\lambda}$ as the resulting expressions are either even in $\sqrt{\lambda}$ or involve only the symmetric combination of $\sqrt{\lambda}$ and $-\sqrt{\lambda}$. If the spectrum of T intersects the branch cut, one may equivalently select any fixed branch that is analytic on a neighborhood of $\sigma(T)$; the analysis proceeds componentwise on each connected subset of $\sigma(T)$ disjoint from the cut.*

3.2. Invariant subspaces – the general case. In this section, we show that the ideas of the simplified case in Section 3.1 extend to more general setting. We assume only that $n \geq N/2$ and T is diagonalizable. Zero may be an eigenvalue of T, and we denote its multiplicity by k. Eigenvectors with eigenvalue zero are denoted $\mathbf{w}_{\lambda=0}$, while all others are denoted $\mathbf{w}_{\lambda\neq 0}$.

The nonzero eigenvalues. For each eigenpair with nonzero eigenvalue $(\lambda, \mathbf{w}_{\lambda\neq 0})$ of T define the two-dimensional subspace $\mathcal{V}_{\lambda\neq 0} \subset \mathbb{C}^N$ through

$$V_{\lambda\neq 0} := \begin{bmatrix} \sqrt{\lambda}\mathbf{w}_{\lambda\neq 0} & \sqrt{\lambda}\mathbf{w}_{\lambda\neq 0} \\ D^{-1}C\mathbf{w}_{\lambda\neq 0} & -D^{-1}C\mathbf{w}_{\lambda\neq 0} \end{bmatrix} \quad \text{and} \quad \mathcal{V}_{\lambda\neq 0} = \operatorname{Im}(V_{\lambda\neq 0}). \tag{3.3}$$

This is analogous to eq. (3.1), and a direct calculation yields the analogue of eq. (3.2)

$$
\begin{aligned}
E^{(s)}V_{\lambda\neq 0} &= \begin{bmatrix} (1-\alpha)I & -\alpha A^{-1}B \\ -\alpha D^{-1}C & (1-\alpha)I \end{bmatrix} \begin{bmatrix} \sqrt{\lambda}\mathbf{w}_{\lambda\neq 0} & \sqrt{\lambda}\mathbf{w}_{\lambda\neq 0} \\ D^{-1}C\mathbf{w}_{\lambda\neq 0} & -D^{-1}C\mathbf{w}_{\lambda\neq 0} \end{bmatrix} \\
&= \begin{bmatrix} \left((1-\alpha)\sqrt{\lambda}-\alpha\lambda\right)\mathbf{w}_{\lambda\neq 0} & \left(-(1-\alpha)\sqrt{\lambda}+\alpha\lambda\right)\mathbf{w}_{\lambda\neq 0} \\ \left(1-\alpha-\alpha\sqrt{\lambda}\right)D^{-1}C\mathbf{w}_{\lambda\neq 0} & -\left(1-\alpha+\alpha\sqrt{\lambda}\right)D^{-1}C\mathbf{w}_{\lambda\neq 0} \end{bmatrix} \\
&= V_{\lambda\neq 0}\begin{bmatrix} (1-\alpha-\alpha\sqrt{\lambda}) & \\ & (1-\alpha+\alpha\sqrt{\lambda}) \end{bmatrix}, \\
E^{(c)}V_{\lambda\neq 0} &= \begin{bmatrix} I & A^{-1}B \\ & \end{bmatrix} \begin{bmatrix} \sqrt{\lambda}\mathbf{w}_{\lambda\neq 0} & \sqrt{\lambda}\mathbf{w}_{\lambda\neq 0} \\ D^{-1}C\mathbf{w}_{\lambda\neq 0} & -D^{-1}C\mathbf{w}_{\lambda\neq 0} \end{bmatrix} = \frac{1}{2}V_{\lambda\neq 0}\begin{bmatrix} 1+\sqrt{\lambda} & 1-\sqrt{\lambda} \\ 1+\sqrt{\lambda} & 1-\sqrt{\lambda} \end{bmatrix}.
\end{aligned}
$$

Hence, we can analyze $E^{(s)}(\alpha)$ and $E^{(c)}$ on $\mathcal{V}_{\lambda\neq 0}$ using only the 2×2 matrices[3]

$$
\begin{aligned}
\mathtt{E}^{(s)}_{\lambda\neq 0} \equiv \mathtt{E}^{(s)}_{\lambda\neq 0}(\alpha) &:= \begin{bmatrix} 1-\alpha-\alpha\sqrt{\lambda} & \\ & 1-\alpha+\alpha\sqrt{\lambda} \end{bmatrix}, \\
\mathtt{E}^{(c)}_{\lambda\neq 0} &:= \frac{1}{2}\begin{bmatrix} 1+\sqrt{\lambda} & 1-\sqrt{\lambda} \\ 1+\sqrt{\lambda} & 1-\sqrt{\lambda} \end{bmatrix}.
\end{aligned} \tag{3.4}
$$

Since $\mathtt{E}^{(s)}_{\lambda\neq 0}(\alpha_i)$ and $\mathtt{E}^{(s)}_{\lambda\neq 0}(\alpha_j)$ commute for any $\alpha_i, \alpha_j \in \mathbb{C}$ we set

$$
\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0} \equiv \mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}(\alpha_1,\ldots,\alpha_m) := \prod_{i=1}^{m}\mathtt{E}^{(s)}_{\lambda\neq 0}(\alpha_i) \equiv \prod_{i=m}^{1}\mathtt{E}^{(s)}_{\lambda\neq 0}(\alpha_i). \tag{3.5}
$$

Therefore, the error propagation operator E in eq. (2.4) restricted to $\mathcal{V}_{\lambda\neq 0}$, denoted by $E\big|_{\mathcal{V}_{\lambda\neq 0}}$, is similar to the following 2×2 matrix $\mathtt{E}_{\lambda\neq 0}$,

$$
\mathtt{E}_{\lambda\neq 0} \equiv \mathtt{E}_{\lambda\neq 0}(\alpha_1,\ldots,\alpha_m) := \mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\mathtt{E}^{(c)}_{\lambda\neq 0}\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}. \tag{3.6}
$$

We realize there is a rank-one factorization of $\mathtt{E}^{(c)}_{\lambda\neq 0}$ as

$$
\mathtt{E}^{(c)}_{\lambda\neq 0} = \mathbf{u}\,\mathbf{v}(\lambda)^*, \quad \text{with} \quad \mathbf{u} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}(\lambda)^* = \begin{bmatrix} 1+\sqrt{\lambda} & 1-\sqrt{\lambda} \end{bmatrix}. \tag{3.7}
$$

This implies that $\mathtt{E}_{\lambda\neq 0}$ is also rank-one and only one of its two eigenvalues is nonzero. Apply $\mathtt{E}_{\lambda\neq 0}$ to the vector $\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\mathbf{u}$:

$$
\begin{aligned}
\mathtt{E}_{\lambda\neq 0}\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\mathbf{u} &= \left(\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\mathbf{u}\mathbf{v}(\lambda)^*\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\right)\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\mathbf{u}, \\
&= \mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\mathbf{u}\left(\mathbf{v}(\lambda)^*\left(\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\right)^2\mathbf{u}\right).
\end{aligned} \tag{3.8}
$$

Thus, the nonzero eigenvalue of $\mathtt{E}_{\lambda\neq 0}$ is $\rho_m(\lambda) := \mathbf{v}(\lambda)^*\left(\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}\right)^2\mathbf{u}$, simplifying to

$$
\rho_m(\lambda) = \tfrac{1}{2}\left(1+\sqrt{\lambda}\right)\prod_{i=1}^{m}\left((1-\alpha_i)-\alpha_i\sqrt{\lambda}\right)^2 + \tfrac{1}{2}\left(1-\sqrt{\lambda}\right)\prod_{i=1}^{m}\left((1-\alpha_i)+\alpha_i\sqrt{\lambda}\right)^2. \tag{3.9}
$$

[3]Notice that $\mathbf{w}_{\lambda\neq 0} \notin \mathrm{Ker}(C)$ by definition, thus $\dim\left(\mathcal{V}_{\lambda\neq 0}\right) = 2$, even if $\mathrm{Ker}(C) \neq \{\mathbf{0}\}$.

Since $(\lambda, \mathbf{w}_{\lambda\neq 0})$ is an arbitrary eigenpair corresponding to a nonzero eigenvalue of T, eq. (3.9) gives the spectrum of E restricted to the invariant subspace $\bigoplus_{\lambda\neq 0}\mathcal{V}_{\lambda\neq 0}$. The eigenvectors corresponding to these eigenvalues are computed in Theorem 3.1 below.

The zero eigenvalues. For each eigenpair with zero eigenvalue $(\lambda, \mathbf{w}_{\lambda\neq 0})$ of T define the two-dimensional subspace $\mathcal{V}_{\lambda\neq 0} \subset \mathbb{C}^N$ through

$$V_{\lambda=0} := \begin{bmatrix} \mathbf{0} & \mathbf{w}_{\lambda=0} \\ D^{-1}C\mathbf{w}_{\lambda=0} & \mathbf{0} \end{bmatrix} \quad \text{and} \quad \mathcal{V}_{\lambda=0} = \operatorname{Im}(V_{\lambda=0}). \tag{3.10}$$

Analogously to the nonzero case, a direct calculation gives

$$\begin{aligned} E^{(s)}V_{\lambda=0} &= \begin{bmatrix} (1-\alpha)I & -\alpha A^{-1}B \\ -\alpha D^{-1}C & (1-\alpha)I \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{w}_{\lambda=0} \\ D^{-1}C\mathbf{w}_{\lambda=0} & \mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0} & (1-\alpha)\mathbf{w}_{\lambda=0} \\ (1-\alpha)D^{-1}C\mathbf{w}_{\lambda=0} & -\alpha D^{-1}C\mathbf{w}_{\lambda=0} \end{bmatrix} = V_{\lambda=0}\left(I_2 - \alpha \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}\right), \\ E^{(c)}V_{\lambda=0} &= \begin{bmatrix} I & A^{-1}B \\ & \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{w}_{\lambda=0} \\ D^{-1}C\mathbf{w}_{\lambda=0} & \mathbf{0} \end{bmatrix} = V_{\lambda=0} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

As in the nonzero case, we can analyze $E^{(s)}(\alpha), E^{(c)}$ on $\mathcal{V}_{\lambda=0}$ using the 2×2 matrices

$$\mathtt{E}^{(c)}_{\lambda=0} := \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathtt{E}^{(s)}_{\lambda=0} \equiv \mathtt{E}^{(s)}_{\lambda=0}(\alpha) := I_2 - \alpha J \quad \text{with} \quad J := \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}. \tag{3.11}$$

We observe an unpleasant feature of $E^{(s)}\big|_{\mathcal{V}_{\lambda=0}} : \mathtt{E}^{(s)}_{\lambda=0}$ is no longer diagonal and its Jordan basis depends on α. Nevertheless, a direct calculation reveals that for any $\alpha, \beta \in \mathbb{C}$ we have

$$(I_2 - \alpha J)(I_2 - \beta J) = (I_2 - \beta J)(I_2 - \alpha J)$$

and so $\mathtt{E}^{(s)}_{\lambda=0}(\alpha), \mathtt{E}^{(s)}_{\lambda=0}(\beta)$ commute for any $\alpha, \beta \in \mathbb{C}$, justifying the definition

$$\mathtt{E}^{(s,\mathrm{full})}_{\lambda=0} \equiv \mathtt{E}^{(s,\mathrm{full})}_{\lambda=0}(\alpha_1, \ldots, \alpha_m) := \prod_{i=1}^{m} \mathtt{E}^{(s)}_{\lambda=0}(\alpha_i) \equiv \prod_{i=m}^{1} \mathtt{E}^{(s)}_{\lambda=0}(\alpha_i).$$

Setting

$$\mathtt{E}_{\lambda=0} \equiv \mathtt{E}_{\lambda=0}(\alpha_1, \ldots, \alpha_m) := \mathtt{E}^{(s,\mathrm{full})}_{\lambda=0}\mathtt{E}^{(c)}_{\lambda=0}\mathtt{E}^{(s,\mathrm{full})}_{\lambda=0},$$

we can write $\mathtt{E}^{(s,\mathrm{full})}_{\lambda=0}$ as a polynomial p in J, obtaining

$$\mathtt{E}_{\lambda=0} = p(J)\mathtt{E}^{(c)}_{\lambda=0}p(J), \quad \text{with} \quad p(t) := \prod_{i=1}^{m}(1-\alpha_i t) \equiv 1 + \sum_{i=1}^{m}\beta_i t^i. \tag{3.12}$$

Calculating further, we obtain

$$J^i = \begin{bmatrix} 1 & i \\ & 1 \end{bmatrix}, \quad \mathtt{E}^{(c)}_{\lambda=0}J^i = \mathtt{E}^{(c)}_{\lambda=0}, \quad \text{and} \quad J^i\mathtt{E}^{(c)}_{\lambda=0} = \begin{bmatrix} 0 & i \\ 0 & 1 \end{bmatrix}$$

and so we can evaluate $\mathtt{E}_{\lambda=0}$ as

$$\begin{aligned}\mathtt{E}_{\lambda=0} &= p(J)\cdot \mathtt{E}^{(c)}_{\lambda=0}\cdot p(J) = p(J)\cdot\left(\mathtt{E}^{(c)}_{\lambda=0}+\sum_{i=1}^{m}\beta_i \mathtt{E}^{(c)}_{\lambda=0}J^i\right) = p(J)\cdot \mathtt{E}^{(c)}_{\lambda=0}p(1)\\ &= \left(\mathtt{E}^{(c)}_{\lambda=0}+\sum_{i=1}^{m}\beta_i J^i \mathtt{E}^{(c)}_{\lambda=0}\right)p(1) = \begin{bmatrix} 0 & \left(\sum\limits_{i=1}^{m} i\beta_i\right)\\ 0 & 1+\left(\sum\limits_{i=1}^{m}\beta_i\right)\end{bmatrix} p(1)\\ &= \begin{bmatrix} 0 & p(1)\left(\sum\limits_{i=1}^{m} i\beta_i\right)\\ 0 & p(1)^2\end{bmatrix}.\end{aligned} \tag{3.13}$$

Thus, the error propagation operator restricted to the invariant subspace $\mathcal{V}_{\lambda=0}$ has one eigenvalue equal to zero and the other to

$$p(1)^2 = \prod_{i=1}^{m}(1-\alpha_i)^2.$$

This is then true for all eigenpairs of T with zero eigenvalue, thereby calculating the spectrum of E restricted to the invariant subspace $\bigoplus_{\lambda=0}\mathcal{V}_{\lambda=0}$. Moreover, the eigenvectors corresponding to these eigenvalues can be explicitly evaluated. Notice also that all of the calculations above simplify notably if $D^{-1}C\mathbf{w}_{\lambda=0}=\mathbf{0}$, i.e., if $\mathbf{w}_{\lambda=0}\in \mathrm{Ker}\,(C)$. We summarize the results of this section below.

THEOREM 3.1. *Let $\alpha_1,\ldots,\alpha_m\in\mathbb{C}$. Suppose that $T = A^{-1}BD^{-1}C$ is diagonalizable with eigenvalues $\lambda_1,\ldots,\lambda_{N-n}$ and associated eigenvectors $\mathbf{w}_\lambda$ for $\lambda\in\{\lambda_1,\ldots,\lambda_{N-n}\}$. Define $k_0 := \dim(\mathrm{Ker}\,(C))$, $k := \dim(\mathrm{Ker}\,(T))$, β_i as in eq. (3.12) and set*

$$\mathcal{V}_\lambda := \mathrm{span}\left(\begin{bmatrix}\mathbf{w}_\lambda\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\mathbf{0}\\ D^{-1}C\mathbf{w}_\lambda\end{bmatrix}\right) \quad\text{and}\quad \rho^{\pm}(\lambda) := \prod_{i=1}^{m}(1-\alpha_i)\pm\alpha_i\sqrt{\lambda}.$$

Then for each λ, $\mathcal{V}_\lambda$ is an invariant subspace of E. Furthermore $\bigoplus_\lambda \mathcal{V}_\lambda$ is an invariant subspace of E and the statements below hold for $E\big|_{\mathcal{V}_\lambda}$, regardless of k, k_0 and α.

1. *If $\mathbf{w}_\lambda\in\mathrm{Ker}(C)$, then $\lambda=0$, $\dim(\mathcal{V}_\lambda)=1$ and the eigenpair of E corresponding to $\mathcal{V}_\lambda$ is $(\rho_m,\mathbf{z}_1)$ with*

$$\rho_m := \prod_{i=1}^{m}(1-\alpha_i)^2 \qquad\text{and}\qquad \mathbf{z}_1 := \begin{bmatrix}\mathbf{w}_\lambda\\ \mathbf{0}\end{bmatrix}.$$

   Across all λ, k_0 eigenpairs are of this form.
2. *If $\mathbf{w}_\lambda\notin\mathrm{Ker}(C)$ and $\lambda=0$, then $\dim(\mathcal{V}_\lambda)=2$ and the two eigenpairs of E corresponding to $\mathcal{V}_\lambda$ are $(\rho_m,\mathbf{z}_1)$ and $(0,\mathbf{z}_2)$ with*

$$\rho_m = \prod_{i=1}^{m}(1-\alpha_i)^2,\quad \mathbf{z}_1 = \begin{bmatrix}\left(\prod_{i=1}^{m}(1-\alpha_i)\right)\mathbf{w}_\lambda\\ \sum\limits_{i=1}^{m} i\beta_i\cdot D^{-1}C\mathbf{w}_\lambda\end{bmatrix},\quad \mathbf{z}_2 = \begin{bmatrix}\mathbf{0}\\ D^{-1}C\mathbf{w}_\lambda\end{bmatrix}$$

   up to the normalization of the eigenvectors. Across all λ, $2(k-k_0)$ eigenpairs are of this form.

3. *If* $\mathbf{w}_\lambda \notin \mathrm{Ker}(C)$ *and* $\lambda \neq 0$*, then* $\dim(\mathcal{V}_\lambda) = 2$ *and the two eigenpairs of* E *corresponding to* $\mathcal{V}_\lambda$ *are* $(\rho_m, \mathbf{z}_1)$ *and* $(0, \mathbf{z}_2)$ *with*

$$\rho_m(\lambda) = \tfrac{1}{2}\left(1+\sqrt{\lambda}\right)\left(\rho^-(\lambda)\right)^2 + \tfrac{1}{2}\left(1-\sqrt{\lambda}\right)\left(\rho^+(\lambda)\right)^2,$$

and

$$\mathbf{z}_1 = \frac{1}{2}\begin{bmatrix} \sqrt{\lambda}(\rho^- + \rho^+)\mathbf{w}_\lambda \\ (\rho^- - \rho^+)D^{-1}C\mathbf{w}_\lambda \end{bmatrix}, \quad \mathbf{z}_2 = \begin{bmatrix} \sqrt{\lambda}\left(\frac{\sqrt{\lambda}-1}{\rho^-} - \frac{\sqrt{\lambda}+1}{\rho^+}\right)\mathbf{w}_\lambda \\ \left(\frac{\sqrt{\lambda}-1}{\rho^-} - \frac{\sqrt{\lambda}+1}{\rho^+}\right)D^{-1}C\mathbf{w}_\lambda \end{bmatrix},$$

if $\mathtt{E}^{(s,\mathrm{full})}$ *is regular. Otherwise, either* ρ^+ *or* ρ^- *is zero and*

$$\begin{cases} \mathbf{z}_1 = \frac{\rho^+}{2}\begin{bmatrix} \sqrt{\lambda}\mathbf{w}_\lambda \\ -D^{-1}C\mathbf{w}_\lambda \end{bmatrix}, \quad \mathbf{z}_2 = \begin{bmatrix} \sqrt{\lambda}\mathbf{w}_\lambda \\ D^{-1}C\mathbf{w}_\lambda \end{bmatrix} & \text{if } \rho^- = 0, \\ \mathbf{z}_1 = \frac{\rho^-}{2}\begin{bmatrix} \sqrt{\lambda}\mathbf{w}_\lambda \\ D^{-1}C\mathbf{w}_\lambda \end{bmatrix}, \quad \mathbf{z}_2 = \begin{bmatrix} \sqrt{\lambda}\mathbf{w}_\lambda \\ -D^{-1}C\mathbf{w}_\lambda \end{bmatrix} & \text{if } \rho^+ = 0 \end{cases}$$

up to the normalization of the eigenvectors. Across all λ*,* $2((N-n)-k)$ *eigenpairs are of this form.*

Proof. Taking any $\lambda \in \sigma(T)$, we notice that the columns of $V_{\lambda\neq 0}$ in eq. (3.3) are linear combinations of the columns of $V_{\lambda=0}$ in eq. (3.10) and hence $\mathcal{V}_\lambda$ coincides both with $\mathcal{V}_{\lambda\neq 0}$ or $\mathcal{V}_{\lambda=0}$, depending on λ. Therefore $\mathcal{V}_\lambda$ is an invariant subspace of E and so is $\bigoplus_\lambda \mathcal{V}_\lambda$.

Statement 1. Taking $\mathbf{w}_\lambda \in \mathrm{Ker}(C)$, we see that $\dim(\mathcal{V}_\lambda) = 1$ and as it is an invariant subspace of E we conclude that $\mathbf{z}_1$ is an eigenvector of E. Moreover, eq. (3.13) shows that the corresponding eigenvalue is $\prod_{i=1}^{m}(1-\alpha_i)^2$, obtaining the k_0 eigenpairs.

Statement 2. Taking $\mathbf{w}_\lambda \notin \mathrm{Ker}(C)$, with $\lambda = 0$, we observe that $\dim(\mathcal{V}_\lambda) = 2$. We recall that eq. (3.11) shows that $E\big|_{\mathcal{V}_\lambda}$ is similar to $\mathtt{E}_{\lambda=0}$, and eq. (3.13) shows that the latter's eigenpairs are

$$\left(\prod_{i=1}^{m}(1-\alpha_i)^2, \begin{bmatrix} \sum_{i=1}^{m} i\beta_i \\ \prod_{i=1}^{m}(1-\alpha_i) \end{bmatrix}\right) \quad \text{and} \quad \left(0, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right),$$

up to normalization. Pre-multiplying the eigenvectors by $V_{\lambda=0}$ gives $2(k-k_0)$ eigenpairs of E and the statement.

Statement 3. Taking $\mathbf{w}_\lambda \notin \mathrm{Ker}(C)$, with $\lambda \neq 0$, then $\dim(\mathcal{V}_\lambda) = 2$. We recall that eq. (3.6) shows that $E\big|_{\mathcal{V}_\lambda}$ is similar to $\mathtt{E}_{\lambda\neq 0}$, the latter's eigenvalues being 0 and ρ_m, seeeq. (3.7) and (3.9). Assuming $\mathtt{E}^{(s,\mathrm{full})}$ is regular, eq. (3.8) shows that its eigenvector with nonzero eigenvalue is $\mathtt{E}^{(s,\mathrm{full})}\mathbf{u}$ and its eigenvector with zero eigenvalue must lie in the kernel of $\mathbf{v}(\lambda)^*\mathtt{E}^{(s,\mathrm{full})}_{\lambda\neq 0}$, i.e.,

$$\mathtt{E}^{(s,\mathrm{full})}\mathbf{u} = \frac{1}{2}\begin{bmatrix} \rho^- \\ \rho^+ \end{bmatrix} \quad \text{and} \quad \left(\mathtt{E}^{(s,\mathrm{full})}\right)^{-1}(\mathbf{v}(\lambda))^\perp = \begin{bmatrix} \frac{\sqrt{\lambda}-1}{\rho^-} \\ \frac{\sqrt{\lambda}+1}{\rho^+} \end{bmatrix}$$

up to the normalization. If $\mathtt{E}^{(s,\mathrm{full})}$ is singular, i.e., either ρ^- or ρ^+ is equal to zero, then either the vector $\mathbf{e}_1 = [1,0]^T$ or $\mathbf{e}_2 = [0,1]^T$ lies in the kernel of $\mathbf{v}(\lambda)^*\mathtt{E}^{(s,\mathrm{full})}$ and

hence becomes its eigenvector with eigenvalue 0. Pre-multiplying the eigenvectors by $V_{\lambda\neq 0}$ gives $2((N-n)-k)$ eigenpairs of E and the statement. ∎

Notice that the two formulas for ρ_m in Theorem 3.1 coincide – inserting $\lambda = 0$ into the latter gives the former. We have calculated $2(N-n)-k_0$ eigenpairs of E, i.e.,

$$\{0\} \cup \bigcup_{\lambda\in\sigma(T)} \rho_m(\lambda) \subset \sigma(E),$$

where each $\rho_m(\lambda)$ is a simple eigenvalue and 0 has the multiplicity $N-n-k_0$. Hence, the spectrum of the preconditioned system satisfies (with the corresponding multiplicities)

$$\{1\} \cup \bigcup_{\lambda\in\sigma(T)} 1-\rho_m(\lambda) \subset \sigma(M^{-1}L).$$

Next, the question needs to be asked – can we compute the entire spectrum of E? To do that, we need to understand the error propagation operator E also on the rest of $\mathbb{C}^N$ and we explore this in the next section.

3.3. Complement of the invariant subspaces. Following Theorem 3.1, we define invariant subspaces whose direct sum is the eigenspace of E parametrized by T[4]

$$\mathcal{V}^{(1)} := \bigoplus_{\lambda\,:\,\mathbf{w}_\lambda\in\mathrm{Ker}(C)} \mathcal{V}_\lambda \qquad \text{and} \qquad \mathcal{V}^{(2)} := \bigoplus_{\lambda\,:\,\mathbf{w}_\lambda\notin\mathrm{Ker}(C)} \mathcal{V}_\lambda. \tag{3.14}$$

We note that by construction all $\mathbf{w}_\lambda$ are linearly independent and thereby $\mathcal{V}_{\lambda_i}\cap\mathcal{V}_{\lambda_j} = \delta_{ij}\mathcal{V}_{\lambda_i}$ and hence all of these direct sums are well defined.

The goal of this section is to show that the complement of $\mathcal{V}^{(1)}\oplus\mathcal{V}^{(2)}$ is contained in $\mathrm{Ker}(E^{(c)})$ and therefore irrelevant from the viewpoint of $E(\alpha_1,\ldots,\alpha_m)$, see eq. (2.4). A direct calculation shows that

$$\mathrm{Ker}(E^{(c)}) = \left\{ \begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix} \in \mathbb{C}^N \mid \mathbf{x}_2\in\mathbb{C}^n \text{ and } \mathbf{x}_1 = -A^{-1}B\mathbf{x}_2 \in \mathbb{C}^{N-n} \right\}, \tag{3.15}$$

is of dimension n. Similarly, we calculate the dimensions of $\mathcal{V}^{(1)}\oplus\mathcal{V}^{(2)}$ as

$$\dim\left(\mathcal{V}^{(1)}\oplus\mathcal{V}^{(2)}\right) = \dim\left(\mathcal{V}^{(1)}\right) + \dim\left(\mathcal{V}^{(2)}\right) = N-n+(N-n-k_0).$$

Thereby

$$\dim\left(\mathcal{V}^{(1)}\oplus\mathcal{V}^{(2)}\right) + \dim\left(\mathrm{Ker}(E^{(c)})\right) = N+(N-n-k_0) \geq N,$$

and so if

$$\dim\left(\left(\mathcal{V}^{(1)}\oplus\mathcal{V}^{(2)}\right)\cap\mathrm{Ker}(E^{(c)})\right) \leq (N-n-k_0), \tag{3.16}$$

then the complement of $\mathcal{V}^{(1)}\oplus\mathcal{V}^{(2)}$ lies in $\mathrm{Ker}(E^{(c)})$.

[4]The superscripts also reflect dimensions of the summed subspaces, i.e., $\mathcal{V}^{(1)}$ is a direct sum of one-dimensional subspaces while $\mathcal{V}^{(2)}$ is a direct sum of two-dimensional subspaces.

PROPOSITION 3.2. *Let*

$$\mathcal{S} := \left(\mathcal{V}^{(1)} \oplus \mathcal{V}^{(2)}\right) \cap \mathrm{Ker}(E^{(c)}),$$

with $\mathcal{V}^{(1,2)}$ *as in eq.* (3.14)*. Then* $\mathcal{S} = \mathcal{V}^{(2)} \cap \mathrm{Ker}(E^{(c)})$ *and*

$$\dim\left(\mathcal{S} \cap \mathrm{Ker}(E^{(c)})\right) = (N - n - k_0),$$

with $k_0 = \dim\left(\mathrm{Ker}\,(C)\right)$ *and*

$$\mathbb{C}^N = \mathcal{V}^{(1)} \oplus \left(\mathcal{V}^{(2)} \backslash \mathcal{S}\right) \oplus \mathcal{S} \oplus \left(\mathrm{Ker}(E^{(c)}) \backslash \mathcal{S}\right).$$

As a result, E is diagonalizable.

Proof. In order to show the decomposition of $\mathbb{C}^N$ we first notice that

$$\left(\mathcal{V}^{(1)} \oplus \mathcal{V}^{(2)}\right) \cap \mathrm{Ker}(E^{(c)}) = \left(\mathcal{V}^{(1)} \cap \mathrm{Ker}(E^{(c)})\right) \oplus \left(\mathcal{V}^{(2)} \cap \mathrm{Ker}(E^{(c)})\right).$$

Recalling eqs. (3.14) and (3.15) we observe that the intersection $\mathcal{V}^{(1)} \cap \mathrm{Ker}(E^{(c)})$ is trivial as $\mathcal{V}_\lambda$ corresponding to $\mathbf{w}_\lambda \in \mathrm{Ker}(C)$ is an eigenspace of E corresponding to a nonzero eigenvalue and hence cannot intersect with $\mathrm{Ker}(E^{(c)})$.

Taking $\mathbf{x} \in \mathcal{V}^{(2)} \cap \mathrm{Ker}(E^{(c)})$, there exist $\zeta_{1,2} \in \mathbb{C}$ and $\lambda \in \sigma(T)$ corresponding to $\mathbf{w}_\lambda \notin \mathrm{Ker}(C)$ such that

$$\mathbf{x} \equiv \begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix} = \begin{bmatrix}\zeta_1 \mathbf{w}_\lambda \\ \zeta_2 D^{-1} C \mathbf{w}_\lambda\end{bmatrix} \quad \text{and} \quad \mathbf{x}_1 = -A^{-1} B \mathbf{x}_2.$$

Plugging the latter into the former gives the condition $\zeta_1 = -\lambda\zeta_2$ and thereby showcases that for any such λ, $\mathcal{V}_\lambda \cap \mathrm{Ker}(E^{(c)})$ is a one-dimensional vector space. Therefore, $\mathcal{V}^{(2)} \cap \mathrm{Ker}(E^{(c)})$ has dimension equal to the number of such $\lambda \in \sigma(T)$, i.e., to $N - n - k_0$, yielding the statement.

Since $\mathbb{C}^N$ can be decomposed as stated, by the above construction we can choose a basis of these four subspaces composed of eigenvectors of E thanks to Theorem 3.1 and eq. (3.15), i.e., E is diagonalizable. ∎

The crucial observation is that the complement of $\mathcal{V}^{(1)} \oplus \mathcal{V}^{(2)}$ *lies fully* within $\mathrm{Ker}(E^{(c)})$. Therefore, for arbitrary smoothing parameters $\alpha_1, \ldots, \alpha_m$ the components of the error lying in the complement of $\mathcal{V}^{(1)} \oplus \mathcal{V}^{(2)}$ are zero-ed out by $E^{(c)}$ and *cannot* be re-introduced by the post-smoothing since $\mathcal{V}^{(1)} \oplus \mathcal{V}^{(2)}$ is an invariant subspace of $E^{(s,\mathrm{full})}$. Altogether, we have justified that we can analyze the error propagation operator E *purely on* $\mathcal{V}^{(1)} \oplus \mathcal{V}^{(2)}$, where Theorem 3.1 gives us the full spectral information of E.

4. Spectral clustering. This section harvests the results of Section 3 and we show how to choose the optimal smoothing parameters $\alpha_1, \ldots, \alpha_m$. Starting with $m = 1$, the nonzero eigenvalues of E become

$$\begin{aligned}\rho_1(\alpha, \lambda) &= \frac{1}{2}(1 + \sqrt{\lambda})\left((1 - \alpha - \alpha\sqrt{\lambda})^2 + \frac{1}{2}(1 - \sqrt{\lambda})(1 - \alpha + \alpha\sqrt{\lambda})^2\right) \\ &= (1 - \alpha)^2 + \alpha^2\lambda - 2\alpha(1 - \alpha)\lambda^2.\end{aligned}$$

From the viewpoint of choosing a suitable α, $\rho_1(\alpha, \lambda)$ is a parametric polynomial whose values give the nonzero eigenvalues of E. Clearly we cannot nullify this polynomial

by a single choice of α for all possible λ to obtain a direct solver. Instead, we seek to perfectly cluster all the nonzero eigenvalues of E to a single value, making it satisfy a degree 2 minimal polynomial. To do so, we make $\rho_1(\alpha, \lambda)$ independent of λ. The root of the derivative of $\rho_1(\alpha, \lambda)$ with respect to λ is equal to $\alpha = 2/3$.

Similarly simple formulas can be found for general m. The setting for general m is analogous to the above: Theorem 3.1 gives $\rho_m \equiv \rho_m(\alpha_1, \ldots, \alpha_m, \lambda)$ as

$$\rho_m = \tfrac{1}{2}\left(1+\sqrt{\lambda}\right)\prod_{i=1}^{m}\left((1-\alpha_i)-\alpha_i\sqrt{\lambda}\right)^2 + \tfrac{1}{2}\left(1-\sqrt{\lambda}\right)\prod_{i=1}^{m}\left((1-\alpha_i)+\alpha_i\sqrt{\lambda}\right)^2. \tag{4.1}$$

To analyze it, we first present a convenient identity, based on [20, §1.39, 1.396(3) and 1.396(2)].

LEMMA 4.1. *Let $m \geq 1$ and define $\theta_i := \frac{2\pi i}{2m+1}$ for $i = 1, \ldots, m$. Then for any $x \in \mathbb{C}$ we have the identity*

$$\prod_{i=1}^{m}(\cos\theta_i \pm t) = \frac{x^{2m+1} \pm 1}{2^m x^m (x \pm 1)},$$

where $t = \frac{1}{2}(x + x^{-1})$.

Proof. From $t = \frac{1}{2}(x + x^{-1})$ we have, for each i,

$$\cos\theta_i \pm t = \cos\theta_i \pm \frac{x + \frac{1}{x}}{2} = \frac{2\cos\theta_i x \pm \left(x^2+1\right)}{2x} = \frac{x^2 \pm 2x\cos\theta_i + 1}{2x}.$$

Taking the product over $i = 1, \ldots, m$ gives

$$\prod_{i=1}^{m}(\cos\theta_i \pm t) = \prod_{i=1}^{m}\frac{x^2 \pm 2x\cos\theta_i + 1}{2x} = \frac{1}{(2x)^m}\prod_{i=1}^{m}\left(x^2 \pm 2x\cos\theta_i + 1\right),$$

and by [20, §1.39, 1.396(3) and 1.396(2)] we get

$$\prod_{i=1}^{m}\left(x^2 + 2x\cos\theta_i + 1\right) = \frac{x^{2m+1}+1}{x+1}, \quad \prod_{i=1}^{m}\left(x^2 - 2x\cos\theta_i + 1\right) = \frac{x^{2m+1}-1}{x-1}.$$

Substituting each case yields the claim. ∎

Equipped with Lemma 4.1, we can formulate the key result of this section.

PROPOSITION 4.2. *Let $m \geq 1$ and take ρ_m as in eq. (4.1), i.e.,*

$$\rho_m(\alpha_1, \ldots, \alpha_m, \tau) = \tfrac{1}{2}(1+\tau)\prod_{i=1}^{m}\left((1-\alpha_i)-\alpha_i\tau\right)^2 + \tfrac{1}{2}(1-\tau)\prod_{i=1}^{m}\left((1-\alpha_i)+\alpha_i\tau\right)^2,$$

where $\tau = \sqrt{\lambda}$ may take complex values. Moreover, as in Lemma 4.1 we set

$$\theta_i = \frac{2\pi i}{2m+1} \quad \text{and} \quad c_i = -\cos\theta_i, \quad i = 1, \ldots, m.$$

Then, for

$$\alpha_i^{opt} := \frac{1}{1+c_i}$$

we have $\rho_m(\alpha_1^{opt}, \ldots, \alpha_m^{opt}, \tau) = \frac{1}{(2m+1)^2}$.

Proof. By definition we have

$$c_i = (1-\alpha_i^{opt})/\alpha_i^{opt}, \quad \text{so that} \quad (1-\alpha_i^{opt}) \mp \alpha_i^{opt}\tau = \alpha_i^{opt}(c_i \mp \tau),$$

and we can write

$$\rho_m = \left(\prod_{i=1}^m \alpha_i^{opt}\right)^2 \left[\tfrac{1}{2}(1+\tau)\prod_{i=1}^m (c_i-\tau)^2 + \tfrac{1}{2}(1-\tau)\prod_{i=1}^m (c_i+\tau)^2\right]. \tag{4.2}$$

Using the change of variables $\tau = \frac{1}{2}(x+x^{-1})$ with $x \in \mathbb{C}\backslash\{0\}$ as in Lemma 4.1 we have

$$\prod_{i=1}^m (c_i-\tau)^2 = \left(\frac{x^{2m+1}+1}{2^m x^m(x+1)}\right)^2 \text{ and } \prod_{i=1}^m (c_i+\tau)^2 = \left(\frac{x^{2m+1}-1}{2^m x^m(x-1)}\right)^2,$$

so that substituting back into eq. (4.2) yields

$$\begin{aligned}
\rho_m &= \left(\prod_{i=1}^m \alpha_i^{opt}\right)^2 \left\{ \tfrac{1}{2}\left(1+\tfrac{x+\frac{1}{x}}{2}\right)\left(\frac{x^{2m+1}+1}{2^m x^m(x+1)}\right)^2 + \tfrac{1}{2}\left(1-\tfrac{x+\frac{1}{x}}{2}\right)\left(\frac{x^{2m+1}-1}{2^m x^m(x-1)}\right)^2 \right\} \\
&= \left(\prod_{i=1}^m \alpha_i^{opt}\right)^2 \left\{ \frac{(x+1)^2}{4x}\left(\frac{x^{2m+1}+1}{2^m x^m(x+1)}\right)^2 - \frac{(x-1)^2}{4x}\left(\frac{x^{2m+1}-1}{2^m x^m(x-1)}\right)^2 \right\} \\
&= \left(\prod_{i=1}^m \alpha_i^{opt}\right)^2 \left\{ \frac{(x^{2m+1}+1)^2 - (x^{2m+1}-1)^2}{2^{2m+2}x^{2m+1}} \right\} \\
&= \left(\prod_{i=1}^m \alpha_i^{opt}\right)^2 \left\{ \frac{\left[(x^{2m+1}+1)-(x^{2m+1}-1)\right]\left[(x^{2m+1}+1)+(x^{2m+1}-1)\right]}{2^{2m+2}x^{2m+1}} \right\} \\
&= \left(\prod_{i=1}^m \alpha_i^{opt}\right)^2 \left\{ \frac{1}{2^{2m+2}x^{2m+1}} \cdot 2 \cdot 2x^{2m+1} \right\} = \frac{\left(\prod_{i=1}^m \alpha_i^{opt}\right)^2}{2^{2m}} = \left(\frac{2^{-m}}{\prod_{i=1}^m \cos\theta_i - 1}\right)^2.
\end{aligned}$$

Using Lemma 4.1 once again for the denominator gives

$$\prod_{i=1}^m (\cos\theta_i - t) = \frac{x^{2m+1}-1}{2^m x^m(x-1)} = \frac{(x-1)\left(\sum_{j=0}^{2m} x^{2m-j}\right)}{2^m x^m(x-1)} = \frac{\sum_{j=0}^{2m} x^{2m-j}}{2^m x^m}$$

taking $t=1$ proves the statement: ∎

$$\rho_m = \left(\frac{2^{-m}}{\prod_{i=1}^m \cos\theta_i - 1}\right)^2 = \frac{1}{(2m+1)^2}.$$

We summarize these results in the context of Algorithm 2.2 below.

COROLLARY 4.3. *Consider Algorithm 2.2 with $n \geq N-n$, the choices of prolongation, restriction and coarse solve given in eq. (2.2) and take the smoothing parameters α_i as*

$$\alpha_i = \frac{1}{1-\cos\left(\frac{2\pi i}{2m+1}\right)}, \qquad i = 1, \ldots, m.$$

Moreover, assume that $T = A^{-1}BD^{-1}C$ is diagonalizable. Then the spectrum of the preconditioned system $M^{-1}L$ is given as

$$\sigma\left(M^{-1}L\right) = \left\{1, 1 - \frac{1}{(2m+1)^2}\right\},$$

regardless of the size of the system or the entries of L.

Proof. Due to eq. (2.5) we have $\sigma\left(M^{-1}L\right) = 1 - \sigma(E)$ and Theorem 3.1 combined with Proposition 3.2 gives

$$\sigma(E) = \{0\} \cup \bigcup_{\lambda \in \sigma(T)} \rho_m\left(\alpha_1, \ldots, \alpha_m, \lambda\right).$$

The result then follows from Proposition 4.2. ∎

5. The direct method. Denoting by M^{-1} the action of Algorithm 2.2 as in Corollary 4.3, we have that $M^{-1}L$ is diagonalizable and has a two-point spectrum $\{1, \rho_m^{opt}, \}$ with $\rho_m^{opt} \equiv 1 - 1/(2m+1)^2$. Therefore, the minimal polynomial of $M^{-1}L$ is

$$\left(M^{-1}L - I\right)\left(M^{-1}L - \rho_m^{opt} I\right) = 0, \tag{5.1}$$

and a rearrangement of the terms gives

$$L^{-1} = \left(\left(1 + \frac{1}{\rho_m^{opt}}\right) I - \frac{1}{\rho_m^{opt}} M^{-1}L\right) M^{-1}. \tag{5.2}$$

We have successfully constructed a direct two-level method for eq. (1.1); it requires two applications of Algorithm 2.2 and a single application of L in the matrix-vector product sense. In this sense, our choices of S^{-1}, P, R, M_0 and $\alpha_1, \ldots, \alpha_m$ are optimal, analogous to the optimality of the choices in eq. (2.1) for Algorithm 2.1.

Krylov subspace methods. Instead of calculating the minimal polynomial of $M^{-1}L$ explicitly, an intuitive alternative is to use a KSM to solve the system with $M^{-1}L$ – since $M^{-1}L$ is diagonalizable, both CG and GMRES will converge in at most two iterations for any right-hand side vector in exact arithmetic[5], requiring at most two application of M^{-1} and L each. A multilevel adaptation is possible – in exact arithmetic, we can identify the two iterations of CG/GMRES for the system with $M^{-1}L$ with the inverse $(M^{-1}L)^{-1}$ at each level and we obtain a W-cycle structure, replacing the explicit formula (5.2) with two CG/GMRES iterations.

Since we need two iterations of a KSM using Algorithm 2.2 as a two-level preconditioner (inverting M_0 explicitly), a recursive, direct multilevel method takes the form of a W-cycle. Both Algorithm 2.1 and Algorithm 2.2 produce the exact solution after one iteration of their respective cycles when used as a preconditioner of a KSM, but the former is a simpler V-cycle – the W-cycle is the price to pay for the "symmetrization" of the method and regularity of the smoother. Next, we highlight two other takeaways from eq. (5.1), which open doors to further numerical approximations of the direct method.

[5]While the CG convergence behavior is *fully* described by the spectrum of the system matrix, this isn't the case for GMRES (even in exact arithmetic; see [25, Sections 5.7.2-5.7.4]). However, if the system matrix is diagonalizable with precisely ℓ distinct eigenvalues, then, in exact arithmetic, GMRES converges in at most ℓ steps.

In fact, "CG/GMRES to precondition CG/GMRES" is an idea that has been considered before and led to the so-called *flexible GMRES* method [39] (fGMRES) or *flexible* or *generalized CG* method [1, 35] (fCG, gCG). The term *flexible* has been used to highlight that the KSM of choice does not require the same preconditioner at each iteration; in other words the preconditioner can be *non-linear*.

This becomes useful for numerical modifications or approximations of (parts of) Algorithm 2.2, i.e., of M^{-1}. As a result of these approximations, two iterations of CG/GMRES no longer correspond to the exact solver and since KSM are, generally speaking, non-linear solvers we naturally obtain a non-linear preconditioner within the recursion. Therefore we need to use the flexible versions of CG/GMRES if we want to use an approximation of the multilevel version of Algorithm 2.2 coupled with CG/GMRES. As a result, we can still run two iterations (or more) of an *optimal polynomial method*, such as CG or GMRES, at each level of the recursion rather than applying a fixed polynomial as suggested by eq. (5.2).

Newton-Schulz method. An interesting way to interpret eq. (5.2) is to realize that the damped Newton's method for finding $\left(M^{-1}L\right)^{-1}$ as the root of the function $F(X) := X^{-1} - M^{-1}L$ reads

$$X_{i+1} = X_i + \omega \left(X_i - X_i M^{-1} L X_i\right),$$

with a damping parameter ω. Taking $X_0 = I$ as the initial guess and the damping parameter as $\omega = 1/\rho_m^{opt}$, the first iteration corresponds to eq. (5.2). In other words, eq. (5.2) corresponds to one iteration of the Newton-Schultz with line-search, by choosing an optimal damping factor ω. We obtain an iterative formulation that allows for generalizations of the multilevel method in the case of approximations or modifications. Moreover, the Newton link might allow us to tackle some nonlinear operators either by finding nonlinear surrogates of the ingredients of Algorithm 2.2, or by using it to approximate the inverse of a Jacobian. Most approaches hinge on our capacity to reconstruct the ingredients for a new iterate of the nonlinear method – we plan to expand on this in future work.

6. Adaptation of the direct solver to practice. This section highlights some practical considerations and examples, showcasing the potential of the above results beyond the abstract setting. A number of these examples come from active research efforts and will be presented in full detail in forthcoming papers.

We want to emphasize that, in our opinion, the results of the above sections are important on their own in raising the hypothetical flagpole in the form of Algorithm 2.2 combined with Corollary 4.3. This gives practitioners a clear message of *what are we trying to approximate* for the Algorithm 2.2 setup and what can we expect in return.

6.1. Conditioning of the preconditioned system. We start by numerically testing the direct solver version of Algorithm 2.2 with Corollary 4.3. For this purpose we take a somewhat involved test problem that will allow us to demonstrate our numerical experience with the direct solver – the goal is to take a complex-valued, non-normal matrix $L \in \mathbb{C}^{N\times N}$, for which we can directly control the non-normality as well as its definitness.

We do that by combining an HPD part and a skew-Hermitian part, scaling each to our liking. We assemble $W, X \in \mathbb{C}^{N\times N}$ by drawing the real and imaginary parts of the matrices from the distribution $\mathcal{N}(0,1)$, independently for each entry, and set $H = W^*W + \eta I$ and $K = \frac{1}{2}(X - X^*)$. Hence, with high probability, H is HPD for

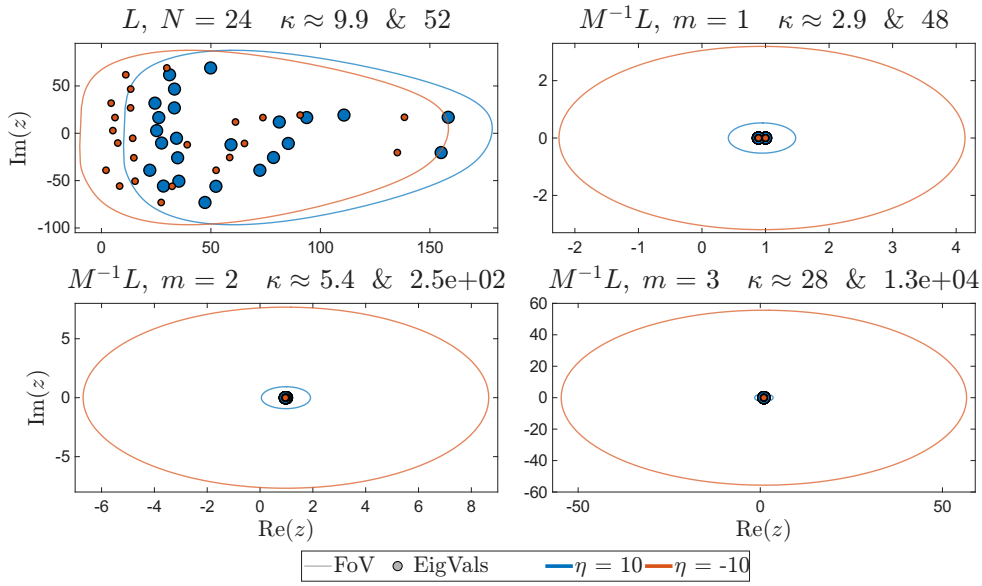

Fig. 1: Field of values and spectra for $L, M^{-1}L \in \mathbb{C}^{24\times 24}$ with $\gamma = 10$. We also give the condition number κ of L or $M^{-1}L$ in the title of each of the subplots, first for $\eta = 10$ and then for $\eta = -10$.

$\eta > 0$ and hermitian, indefinite for $\eta < 0$; on the other hand K is skew-Hermitian. Taking their linear combination as

$$L = H + \gamma K,$$

with some $\gamma > 0$, we obtain (with high-probability) a non-normal matrix L. Analogously, we also generate a random, complex-valued right-hand side vector and solve the resulting system as outlined in Section 5, with two GMRES iterations used at each level, continuing the recursion till the coarse space is a scalar problem.

In Figure 1, we illustrate some properties of L and $M^{-1}L$ when L is definite ($\eta = \gamma = 10$) or indefinite ($\eta = -10, \gamma = 10$); in both cases L is well-conditioned (condition number $\leq 10^2$). First, we notice the numerically perfect spectral clustering[6] when using Algorithm 2.2 combined with Corollary 4.3, precisely as predicted, regardless of the definitness of L. Next, if L is definite, then the preconditioned system is also well-conditioned even for larger number of pre- and post-smoothings. However, when L becomes indefinite, the preconditioned system becomes increasingly *ill-conditioned* as the number of pre- and post-smoothings grows. For $m = 6$ we obtain $\kappa \approx 10^8$ for L indefinite, compared to $\kappa \approx 10^4$ for L definite. This sounds dangerous and it is – left preconditioned GMRES measures the convergence by the relative *preconditioned* residual, i.e., the direct solver may deliver the correct solution to a wrong problem. For $m = 3$, i.e., the bottom-right-most subplot, the *relative error* of the resulting approximate solution becomes roughly of order 10^{-10} (while for L definite, it stays at the level of 10^{-14}).

[6]The clustering is on the level $n^{2m+1}\varepsilon_{mach}$, which aligns with the rounding error level due to the assembly of $M^{-1}L = I - E^{(s,\text{full})}E^{(c)}E^{(s,\text{full})}$.

Consider L first with η and then $-\eta$, we observe this gap in conditioning of the preconditioned system to shrink (or widen) as $\eta \to 0$ (or $\eta \to +\infty$) but grows steadily as we increase m. We note that from practical viewpoint, the latter is not too pressing as we are usually not interested in running a larger number of pre- and post-smoothing. On the other hand, indefinite problems are of practical importance and there we advice caution as we believe that further understanding is necessary, especially for practical adaptations of Algorithm 2.2.

6.2. Sparsity of prolongation and restriction operators. With most multigrid methods we expect that the application of the matrices S^{-1}, P, R and M_0^{-1} is *cheap*. For P, R this classically follows from their structural sparsity while for S^{-1} this is often achieved by construction, e.g., taking point-Jacobi or block-Jacobi (with small blocksize) smoothers; M_0^{-1} then should inherit these properties by recursion. Using Algorithm 2.2 with the choices of operators as in Corollary 4.3 directly, these practical requirements will often be violated. In this section we consider adaptations that still outperform the widely used choices for a simple model problem.

For the sake of brevity, we consider the model problem of two–dimensional Poisson equation with homogeneous Dirichlet boundary conditions on the unit square, discretizing $-\Delta$ on a uniform finite–difference mesh. Obtaining a system like in eq. (1.1), the system matrix $L \in \mathbb{R}^{N\times N}$ has a *Kronecker sum* structure, i.e., having the one–dimensional second–order finite–difference Laplacian $L_{1D} \in \mathbb{R}^{\sqrt{N}\times\sqrt{N}}$ corresponding to the stencil $\frac{1}{h^2}[-1\ 2\ -1]$ on a uniform mesh with $\sqrt{N}$ interior points and spacing h, we have

$$L = L_{1D} \oplus L_{1D} \equiv L_{1D} \otimes I + I \otimes L_{1D}. \tag{6.1}$$

Applying Corollary 4.3 *directly* is possible but the prolongation and restriction operators then require the construction of $-A^{-1}B$ (and its hermitian conjugate), which is a dense matrix, posing a significant bottleneck to the efficiency of the multilevel method.

One common practice addressing this issue is to reorder the DoFs, here using the *red-black reordering*. Taking $n = N/2$, both A and D become *diagonal* while B and C inherit the sparsity of L (in our case having four nonzero entries per row/column). This way the prolongation and restriction operators are much more efficient and we can use Algorithm 2.2 with Corollary 4.3 and GMRES acceleration from Section 5 to obtain a direct solver.

Another choice, used more commonly for more challenging problems, is to approximate expensive operators by means of *tensor lifting* [23], i.e., the idea that if the differential operator has a "tensor structure", e.g., it is a tensor product or sum, then we can use this algebraically and approximate the ideal operators by the tensor product or sum of their lower-dimensional discretizations. We demonstrate this heuristic on the model problem above. First, we red-black reorder the DoFs of L_{1D} so that

$$\Pi_{1D}^{\mathrm{RB}} L_{1D} (\Pi_{1D}^{\mathrm{RB}})^T = \begin{bmatrix} A_{1D} & B_{1D} \\ B_{1D}^T & D_{1D} \end{bmatrix},$$

and $A_{1D}, D_{1D} \in \mathbb{R}^{\sqrt{N}/2\times\sqrt{N}/2}$ become diagonal. Then, we apply the algebraic construction of Corollary 4.3 to L_{1D} based on its above blocking and obtain $P_{1D}, R_{1D}, S_{1D}^{-1}$. The approximate prolongation, restriction and smoothing operators for L then follow as

$$P = P_{1D} \otimes P_{1D}, \quad R = R_{1D} \otimes R_{1D} \quad \text{and} \quad S^{-1} = S_{1D}^{-1} \oplus S_{1D}^{-1}. \tag{6.2}$$

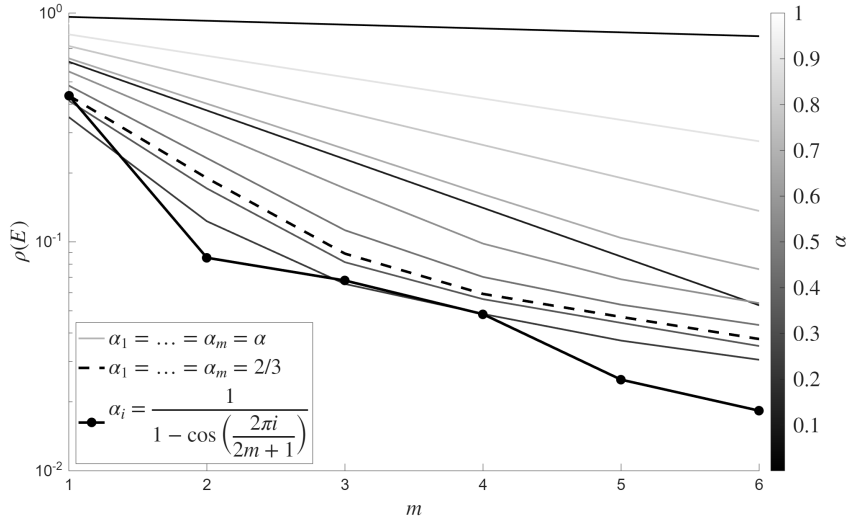

Fig. 2: Spectral radius $\rho(E)$ as a function of the number of smoothing steps m for the two–dimensional Laplacian system on a 16×16 interior finite–difference grid preconditioned with Algorithm 2.2. The right–hand side is taken as $\mathbf{b} \equiv \mathbf{1}$ so that the underlying PDE solution is smooth. Gray curves: constant $\alpha \in \{0.1, \ldots, 1.0\}$; dashed: constant $\alpha = 2/3$; thick curve with filled markers: smoothing parameters given by Corollary 4.3.

Clearly, the results of previous sections do not apply but due to the 1D reordering we can draw analogous conclusions for the sparsity as in the case of the 2D red-black reordering for this model problem, i.e., we get a diagonal smoother and sparse prolongation and restriction operators, with the same number of nonzeros per row/column. The choice of the smoothing parameters α is usually done heuristically, based on considering the smoother alone as in [14], leading to the common choice $\alpha_i = 2/3$ for all i, using LFA as in [21] or based on a lifting strategy similarly to [28], where tedious analysis of the method in 1D yields optimal parameters and these are simply used for higher-dimensional problems. Adopting a similar mindset, we illustrate in Figure 2 that even with the choices in eq. (6.2) instead of eq. (2.1) our parameters $\alpha_1, \ldots, \alpha_m$ perform *near-optimal* compared to other choices of α. This suggests that the algebraic spectral optimization of Corollary 4.3 extends approximately to higher–dimensional tensor–product discretizations while preserving sparse restriction and prolongation operators.

We also want to note that using either of these restriction/prolongation operators, the coarse matrix M_0 becomes denser than L itself, e.g., in our case M_0 has the sparsity pattern of the 9-point stencil and so neither of these techniques give a win-all answer (in terms of efficiency of the multilevel method). The point is that the results of Corollary 4.3 on the optimal parameters can be used also in the tensor-lifting setting. Further testing and justification of these observations remain an active areas of research for us.

6.3. Discontinuous Galerkin 1D. The clustering result of Corollary 4.3 is purely algebraic but similar work has been recently explored also from the geometrical point of view in [27] for the *discontinuous Galerkin* method applied to a simple test

problem, where the author find optimal parameters within a family of prolongation and restriction operators. In [27], the prolongation operator was designed under strict locality and symmetry constraints: it acted *within each DG element* and its sparsity pattern ensured that each fine-grid degree of freedom depends on a *single* coarse element. The interpolation was therefore purely local, and the adjustable parameters controlling its behavior were the discontinuity parameter c combined with the penalty parameter δ_0 and the relaxation parameter α. Under these constraints, the authors showed there exists a *unique* combination of (α, c, δ_0) that produces a perfectly clustered spectrum.

In our setting, we no longer prescribe the interpolation geometry, i.e., these geometric constraints are lifted. Instead, we derive the prolongation operator directly from the algebraic block structure of the reordered fine-grid matrix. As in Section 6.2, we perform an element-wise red-black permutation so that the operator takes the form

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad P_{\mathrm{RB}} = \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}, \quad R_{\mathrm{RB}} = P_{\mathrm{RB}}^{\top}, \quad M_0 = R_{\mathrm{RB}} M P_{\mathrm{RB}},$$

considering *two admissible* $1D$ *red-black reorderings – element-wise* (grouping DoFs by alternating elements) and *interface-wise* (alternating traces across interfaces). These lead to different *local* algebraic stencils for P and for the Jacobi smoother S^{-1}. In all cases, the smoothing relaxations are the $\{\alpha_i\}_{i=1}^m$ from Corollary 4.3.

Element-wise (local stencils). On each two-element patch, the local prolongation block (fine 8 DoFs from 4 coarse DoFs) and the Jacobi smoother uses 2×2 diagonal blocks are

$$P_{\mathrm{loc,el}}(\delta) = \begin{pmatrix} \frac{\delta}{4\delta-2} & \frac{\delta-1}{4\delta-2} & 0 & 0 \\ \frac{\delta-1}{4\delta-2} & \frac{\delta}{4\delta-2} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{\delta}{4\delta-2} & \frac{\delta-1}{4\delta-2} & \frac{\delta}{4\delta-2} & \frac{\delta-1}{4\delta-2} \\ \frac{\delta-1}{4\delta-2} & \frac{\delta}{4\delta-2} & \frac{\delta-1}{4\delta-2} & \frac{\delta}{4\delta-2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad S^{-1}_{\mathrm{loc,el}}(\delta) = \frac{1}{-1+2\delta} \begin{pmatrix} \delta & \delta-1 \\ \delta-1 & \delta \end{pmatrix}. \tag{6.3}$$

Interface-wise (local stencils). On each interface, the local prolongation block and the Jacobi smoother are

$$P_{\mathrm{loc,int}}(\delta) = \begin{pmatrix} \frac{\delta-1}{\delta} & \frac{1}{2\delta} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2\delta} & \frac{\delta-1}{\delta} & \frac{1}{2\delta} & 0 \\ 0 & \frac{1}{2\delta} & \frac{\delta-1}{\delta} & \frac{1}{2\delta} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2\delta} & \frac{\delta-1}{\delta} \end{pmatrix}, \quad S^{-1}_{\mathrm{loc,int}}(\delta) = \frac{1}{\delta} I_2, \tag{6.4}$$

where $\delta > 0$ denotes the DG penalty parameter (analogous to δ_0 in [27]) and assembly uses the usual overlap; the resulting global P are *not* block diagonal.

The definition of R_{RB} and M_0 follows directly from the symmetric two-level formulation introduced in Section 2; it is not an additional assumption but a necessary consequence of the formulation. In particular, the theory developed in Section 3 applies for an *arbitrary number of smoothing steps* m, with the relaxation parameters $\{\alpha_j\}_{j=1}^m$ chosen as in Corollary 4.3. Obtaining such a closed-form description of the spectrum for general m would be extremely cumbersome with a classical Local Fourier Analysis approach, which requires a mode-by-mode treatment and becomes rapidly intractable as the number of smoothing steps increases.

Compared with the interpolation from [27], this algebraic prolongation has a richer coupling pattern: some fine-grid degrees of freedom are interpolated from *two neighboring coarse elements* instead of one. This stronger coupling increases the effective range of the operator and removes the need for parameter tuning. The two-point spectral clustering, which in [27] appeared only for a specific triple (α, c, δ_0), now emerges automatically for *any* choice of δ.

From the DG perspective, this prolongation is again discontinuous: even if the coarse solution is continuous across an interface (that is, the two coarse traces coincide), the mixed rows of $P_{\mathrm{loc}}(\delta)$ generally produce different fine values on the two sides of that interface, except for the constant mode. Hence a continuous coarse function is typically mapped to a *discontinuous* fine one. Together with the results of [27], this reinforces the same qualitative conclusion: when constructing efficient multilevel preconditioners for DG discretizations, the answer to the question *"Should multilevel methods for DG discretizations use discontinuous interpolation operators?"* remains **yes**—even when the interpolation arises purely from the algebraic two-level factorization rather than from an explicit geometric design.

This shows that the clustering phenomenon does not depend on the geometric notion of continuity or discontinuity of the interpolation, but follows directly from the algebraic structure of the symmetric two-level formulation itself.

REFERENCES

[1] O. Axelsson and P. S. Vassilevski. A black box generalized conjugate gradient solver with inner iterations and variable-step preconditioning. *SIAM Journal on Matrix Analysis and Applications*, 12(4):625–644, 1991.

[2] N. S. Bakhvalov. On the convergence of a relaxation method with natural constraints on the elliptic operator. *USSR Computational Mathematics and Mathematical Physics*, 6(5):101–135, 1966.

[3] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[4] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation*, 31(138):333–390, 1977.

[5] A. Brandt. Algebraic multigrid theory: The symmetric case. *Applied Mathematics and Computation*, 19(1–4):23–56, 1986.

[6] A. Brandt, S. F. McCormick, and J. W. Ruge. Algebraic multigrid (AMG) for sparse matrix equations. In D. J. Evans, editor, *Sparsity and Its Applications*, pages 257–284. Cambridge University Press, Cambridge, UK, 1985.

[7] J. Brannick, A. Frommer, K. Kahl, S. Maclachlan, and L. Zikatanov. Adaptive reduction-based multigrid for nearly singular and highly disordered physical systems. *Electronic Transactions on Numerical Analysis*, 37:276–295, 2010.

[8] M. Brezina, T. Manteuffel, S. McCormick, J. Ruge, and G. Sanders. Towards adaptive smoothed aggregation (αSA) for nonsymmetric problems. *SIAM Journal on Scientific Computing*, 32(1):14–39, 2010.

[9] W. L. Briggs, V. E. Henson, and S. F. McCormick. *A Multigrid Tutorial*. SIAM, Philadelphia, PA, USA, 2nd Edition edition, 2000.

[10] B. L. Buzbee, G. H. Golub, and C. W. Nielson. On direct methods for solving Poisson's equations. *SIAM Journal on Numerical Analysis*, 7(4):627–656, 1970.

[11] T. F. Chan and W. L. Wan. Robust multigrid methods for nonsmooth coefficient elliptic linear systems. *Journal of Computational and Applied Mathematics*, 123(1):323–352, 2000.

[12] E. de Sturler and J. Liesen. Block-diagonal and constraint preconditioners for nonsymmetric indefinite linear systems. Part I: Theory. *SIAM Journal on Scientific Computing*, 26(5):1598–1619, 2005.

[13] V. Dolean, P. Jolivet, and F. Nataf. *An Introduction to Domain Decomposition Methods: Algorithms, Theory, and Parallel Implementation*. SIAM, Philadelphia, PA, USA, 2015.

[14] O. G. Ernst and M. J. Gander. Multigrid methods for Helmholtz problems: A convergent scheme in 1D using standard components. In I. Graham, U. Langer, J. M. Melenk, and Mo. Sini, editors, *Direct and Inverse Problems in Wave Propagation and Applications*,

volume 14 of *Radon Series on Computational and Applied Mathematics*, pages 135–186. De Gruyter, 2013.

[15] R. D. Falgout and P. S. Vassilevski. On generalizing the algebraic multigrid framework. *SIAM Journal on Scientific Computing*, 27(4):1233–1259, 2004.

[16] R. P. Fedorenko. A relaxation method for solving elliptic difference equations. *USSR Computational Mathematics and Mathematical Physics*, 1:109–128, 1962.

[17] R. P. Fedorenko. The speed of convergence of one iterative process. *USSR Computational Mathematics and Mathematical Physics*, 4(3):227–235, 1964.

[18] H. Foerster, K. Stüben, and U. Trottenberg. Non-standard multigrid techniques using checkered relaxation and intermediate grids. In Martin H. Schultz, editor, *Elliptic Problem Solvers*, pages 285–300. Academic Press, 1981.

[19] M. J. Gander. 50 Years of time parallel time integration. In T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, editors, *Multiple Shooting and Time Domain Decomposition Methods*, pages 69–113. Springer, Cham, 2015.

[20] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series and Products*. Academic Press, Amsterdam, 7 edition, 2007.

[21] W. Hackbusch. *Multi-Grid Methods and Applications*, volume 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[22] W. Hackbusch. *Hierarchical Matrices: Algorithms and Analysis*, volume 49. Springer Berlin, Heidelberg, 2015.

[23] P. W. Hemker, W. Hoffmann, and M. H. van Raalte. Fourier two-level analysis for discontinuous Galerkin discretization with linear elements. *Numerical Linear Algebra with Applications*, 11(5-6):473–491, 2004.

[24] I. C. F. Ipsen. A note on preconditioning nonsymmetric matrices. *SIAM Journal on Scientific Computing*, 23(3):1050–1051, 2001.

[25] J. Liesen and Z. Strakoš. *Krylov Subspace Methods: Principles and Analysis*. Oxford University Press, Oxford, UK, 2013.

[26] J. P. Lucero Lorca. Towards a multigrid preconditioner interpretation of Hierarchical Poincaré-Steklov solvers. arXiv:2511.00735, 2025.

[27] J. P. Lucero Lorca and M. J. Gander. Should multilevel methods for discontinuous Galerkin discretizations use discontinuous interpolation operators? In S. C. Brenner, E. Chung, A. Klawonn, F. Kwok, J. Xu, and J. Zou, editors, *Domain Decomposition Methods in Science and Engineering XXVI*, pages 273–280. Springer, Cham, 2022.

[28] J. P. Lucero Lorca and M. J. Gander. Optimization of two-level methods for DG discretizations of reaction-diffusion equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 58(6):2351–2386, 2024.

[29] J. P. Lucero Lorca, D. Rosenberg, I. Jankov, C. McCoid, and M. J. Gander. On an efficient line smoother for the p-multigrid γ-cycle. arXiv:2504.10710, 2025.

[30] S. MacLachlan, T. Manteuffel, and S. F. McCormick. Adaptive reduction-based AMG. *Numerical Linear Algebra with Applications*, 13(8):599–620, 2006.

[31] T. A. Manteuffel, S. Münzenmaier, J. Ruge, and B. Southworth. Nonsymmetric reduction-based algebraic multigrid. *SIAM Journal on Scientific Computing*, 41(5):S242–S268, 2019.

[32] T. A. Manteuffel, J. Ruge, and B. Southworth. Nonsymmetric algebraic multigrid based on local approximate ideal restriction (ℓAIR). *SIAM Journal on Scientific Computing*, 40(6):A4105–A4130, 2018.

[33] P.-G. Martinsson. The Hierarchical Poincaré–Steklov (HPS) solver for elliptic PDEs: A tutorial. arXiv:1506.01308, 2015.

[34] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM Journal on Scientific Computing*, 21(6):1969–1972, 2000.

[35] Y. Notay. Flexible conjugate gradients. *SIAM Journal on Scientific Computing*, 22(4):1444–1460, 2000.

[36] L. N. Olson, J. B. Schroder, and R. S. Tuminaro. A general interpolation strategy for algebraic multigrid using energy minimization. *SIAM Journal on Scientific Computing*, 33(2):966–991, 2011.

[37] M. Ries, U. Trottenberg, and G. Winter. A note on MGR methods. *Linear Algebra and Its Applications*, 49:1–26, 1983.

[38] J. W. Ruge and K. Stüben. Algebraic multigrid. In S. F. McCormick, editor, *Multigrid Methods, Frontiers in Applied Mathematics, Vol. 3*, pages 73–130. SIAM, Philadelphia, PA, USA, 1987.

[39] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM Journal on Scientific Computing*, 14(2):461–469, 1993.

[40] M. Sala and R. S. Tuminaro. A new Petrov–Galerkin smoothed aggregation preconditioner

for nonsymmetric linear systems. *SIAM Journal on Scientific Computing*, 31(1):143–166, 2008.

[41] C. Siefert and E. de Sturler. Preconditioners for generalized saddle-point problems. *SIAM Journal on Numerical Analysis*, 44(3):1275–1296, 2006.

[42] P. Wesseling. *An Introduction to Multigrid Methods*. John Wiley & Sons, Chichester, 1992.

[43] T. A. Wiesner, R. S. Tuminaro, W. A. Wall, and M. W. Gee. Multigrid transfers for nonsymmetric systems based on Schur complements and Galerkin projections. *Numerical Linear Algebra with Applications*, 21(3):415–438, 2014.

[44] X. Xu and C.-S. Zhang. On the ideal interpolation operator in algebraic multigrid methods. *SIAM Journal on Numerical Analysis*, 56(3):1693–1710, 2018.