A COMBINED PRECONDITIONING STRATEGY FOR NONSYMMETRIC SYSTEMS

BLANCA AYUSO DE DIOS, ANDREW T. BARKER, AND PANAYOT S. VASSILEVSKI

Abstract. We present and analyze a class of nonsymmetric preconditioners within a normal (weighted least-squares) matrix form for use in GMRES to solve nonsymmetric matrix problems that typically arise in finite element discretizations. An example of the additive Schwarz method applied to nonsymmetric but definite matrices is presented for which the abstract assumptions are verified. A variable preconditioner, combining the original nonsymmetric one and a weighted least-squares version of it, is shown to be convergent and provides a viable strategy for using nonsymmetric preconditioners in practice. Numerical results are included to assess the theory and the performance of the proposed preconditioners.

1. INTRODUCTION

The numerical approximation of most phenomena in science and technology requires the solution of linear or nonlinear algebraic systems. Preconditioning is one of the main techniques that combined with a proper iterative method allows for reducing substantially the cost of solving those systems. Many efforts are usually devoted to design proper preconditioning strategies that allow for efficient and fast solution of the resulting algebraic systems [19, 20]. The development of preconditioners is very often guided by the properties of the underlying problem and it sometimes can even dictate particular aspects that should be accounted for, when devising the numerical discretization of the continuous problem (as for instance in [12, 1]).

Even for linear problems, the design and analysis of preconditioners for the linear systems is far from being complete. For symmetric and coercive problems, a reasonable discretization yields a linear system $A_0x = b$ with $A_0$ symmetric
and positive definite (s.p.d). In such cases, as it is well known, the spectral information of the matrix itself dictates completely the convergence of the method. Therefore a preconditioner $B_0$ would be uniform (and could be turned into an optimal preconditioner) if it captures completely such spectral information; in other words, if $B_0$ is spectral equivalent to $A_0$.

However, for linear systems $Ax = b$, with nonsymmetric coefficient matrix $A$, the design of effective preconditioners does not admit a general recipe, at least at the present time. Likewise, there is no general iterative solver, and furthermore, there is no general theory that could be used to explain the success of a particular preconditioner when it is indeed efficient. In most cases, the spectral information does not provide significant information that could guide the development of any good preconditioner. Field of values have shown some utility in certain circumstance, but have also shown many limitations [10, 18, 16]. At the moment it might seem that each particular problem has to be studied separately and a problem-dependent, discretization-dependent preconditioning strategy had to be devised. Besides, even when such preconditioning can be designed, its understanding and analysis are tasks that in most cases are out of reach.

In this paper, we focus on a particular situation, where $A$ is nonsymmetric but still positive definite. The motivation and application comes from a nonsymmetric Discontinuous Galerkin discretization of an elliptic problem [8]. In [2, 3], additive and multiplicative Schwarz preconditioners were developed for the solution of the resulting algebraic system. In both works, the authors show that the GMRES convergence theory cannot be applied for explaining the convergence since the preconditioned system does not satisfy the sufficient conditions required by such theory. However, such discretizations are used in practice and have already shown to have some advantages when approximating advection-diffusion problems [14, 5] and more recently, in the design of methods for some more complex nonlinear problems [17, 15]. In [6], the authors introduce a solver methodology based on the idea of subspace correction for this type of discretizations for elliptic problems, providing the analysis of the resulting iterative methods without using any GMRES theory. In this paper we want to examine, in a more general algebraic abstract framework (that in particular will apply to the type of methods discussed above), the issue of providing some convergence theory for a preconditioner based on the classical (but nonsymmetric) Schwarz preconditioner to be used within GMRES. The ultimate goal is to obtain some insight on how to improve and tune the preconditioner.

Here, in a first stage we consider two preconditioners for $A$; a classical additive Schwarz preconditioner $B$, which is nonsymmetric, and a symmetric preconditioner $Z$ that basically uses actions of the additive Schwarz preconditioner $B$ and its adjoint. Both will be shown to have their pros and cons. For the former, the
non-symmetry of $B$ and of $B^{-1}A$, precludes developing any theory from which to extract either some a-priori information on the convergence or to provide some guidelines on how the preconditioner could be improved or even be designed. The latter, while allowing for developing a convergence theory, will be shown to be not the most efficient possible option, although the a-priori information on convergence could be of special value depending on the application. Other symmetrizing strategies for classical Schwarz methods, different from the one introduced here, have been already considered in literature by other authors [13].

At first sight, the underlying message that one might obtain from the analysis of this first part, is that enforcing the symmetry of the preconditioner for a non-symmetric matrix might result in the very end, in a wasted effort. We believe this might be the case in many situations, and we also think it is relevant and important to point it out. At the same time, we do believe that the results obtained for analyzing the preconditioner, are of independent interest (also because of its simplicity), and might provide some basis (as it had happened already here) or insights for further development of solvers for nonsymmetric systems.

In a second stage of the present paper, we introduce a variable preconditioner $B$ that is constructed by considering a linear combination of two given (general) preconditioners $B$ and $Z$ so that in a sense it tries to integrate and exploit the best of each of them. We describe the construction of this variable preconditioner to be used in GMRES, explaining how the coefficients in its definition are determined at each iteration inside GMRES. We show that from the construction of $B$ we immediately can deduce (theoretically) a convergence estimate that guarantees better performance of the resulting solver. In particular, we show that the new preconditioner outperforms the symmetric preconditioner $Z$ and indeed always converges faster.

The theory is illustrated with extensive numerical experiments, in which we also study the performance of all the considered preconditioners. They are all implemented in parallel to fully take advantage of having considered preconditioners based on additive Schwarz methods. In the numerical tests, we do observe that the combined preconditioner requires less GMRES-iterations to achieve convergence than the the classical additive Schwarz $B$. However, in this particular case, each iteration for the combined preconditioner is more costly, which in the end, makes $B$ perform slightly better in terms of execution time. From these observations, it might be inferred that the new combined preconditioner $B$ might be more competitive in settings where each iteration is expensive, so that the savings in iteration count can make up for the high cost per iteration.

Although we have focused on the nonsymmetric but positive definite case, we believe the ideas presented in the paper might be useful and possibly extended to more complex problem, including the indefinite case. This issue will be subject of future research.
The outline of the paper is as follows. Section 2 contains a description of the problem and the original motivation of it. In Section 3, we construct the preconditioner $Z$ and present the convergence analysis. The combined preconditioner is introduced and analyzed in Section 4. Finally in Section 5, we consider a particular application and we provide numerical experiments that verify the developed theory and assess the performance of the preconditioner.

2. Problem formulation

We are interested in preconditioning a given system of linear equations

$$A x = b,$$

with $A$ being non-symmetric but definite and $n$ is assumed to be large. For the applications we have in mind, $A$ comes from a finite element discretization of some partial differential operator and therefore is sparse and structured.

We also assume that $A$ is ill-conditioned and that therefore a good preconditioner $B$ is required to solve efficiently system (2.1) by an iterative method. A simple option is to construct such $B$ as the classical additive Schwarz preconditioner coming from $A$. More precisely, we denote by $I_k$, $k = 1, \ldots, m$, a set of rectangular matrices, such that $I_k$ extends a local vector $v_k$ to a global vector $I_k v_k$ with zero entries outside its index set. Also, let $I_c = P$ be an interpolation matrix that maps a coarse vector $v_0 = v_c$ to a global vector $I_c v_c = P v_c$. Then, the additive Schwarz preconditioner exploits the local matrices $A_k = I_k^T A I_k$, principal submatrices of $A$, and the coarse matrix $A_c$ defined as $A_c = I_c^T A I_c = P^T A P$.

The inverse of the additive Schwarz preconditioner $B$ takes the following familiar form:

$$B^{-1} = PA_c^{-1}P^T + \sum_{k=1}^{N_c} I_k A_k^{-1} I_k^T.$$

Obviously, since $A$ is nonsymmetric the resulting additive Schwarz preconditioner $B$ is also nonsymmetric. Therefore, for the solution of the resulting preconditioned system $B^{-1} A x = B^{-1} b$, one has to use any of the iterative methods for nonsymmetric systems, such as the Generalized Minimal Residual (GMRES). For analyzing the convergence of the resulting iterative method (for the preconditioned system) one has to resort to one of the available and non-optimal GMRES theories. In the Domain Decomposition framework, the GMRES convergence theory of Eisenstat et. al. [9] is generally used. In particular, to derive (a-priori) any conclusion on the performance of the preconditioner $B$, this theory requires some control on the coercivity of $B^{-1} A$ (in some inner product). Therefore, at least in theory, using $B$ directly as a preconditioner for $A$ might not be successful.

Still, we would like to utilize the actions of $B^{-1}$ to define a preconditioner, say $Z$, for $A$, for which some bounds on the rate of convergence can be a-priori determined. In the next section, we show how such a preconditioner $Z$ can be
constructed (and analyzed) by exploiting the fact that although $A$ is nonsymmetric, it is positive definite in some inner product. We also compare numerically, in Section 5, the performance of the constructed preconditioner $Z$ with the original nonsymmetric additive Schwarz $B$. As we will show, even if a theory can be developed for $Z$ it might not be the most efficient option.

We now state our basic assumption regarding the matrix $A$. More specifically, we assume that there is an s.p.d. matrix $A_0$ such that $A$ and $A_0$ are related by the following basic assumption:

Assumption (H0): Let $A \in \mathbb{R}^{n \times n}$ be nonsymmetric but definite and let $A_0 \in \mathbb{R}^{n \times n}$ be s.p.d. We say that the pair of matrices $(A, A_0)$ satisfy Assumption (H0) with constants $(c_0, c_1)$ if they do satisfy the following coercivity and boundedness estimates:

\begin{align}
(2.3) & \quad v^T A v \geq c_0 v^T A_0 v \quad \text{for all } v, \\
(2.4) & \quad w^T A v \leq c_1 \sqrt{v^T A_0 v} \sqrt{w^T A_0 w} \quad \text{for all } v, w.
\end{align}

3. An abstract result

In this section we present the construction and give the analysis of a preconditioner for $A$ that basically only uses the actions of the additive Schwarz method. We start proving a couple of Lemmas that will be required for our subsequent analysis and derivation.

The next Lemma shows that for any pair of matrices $(A, A_0)$ satisfying (H0) with constants $(c_0, c_1)$, the corresponding pair $(A^{-1}, A_0^{-1})$ (consisting of their inverses) also satisfies (H0) with constants $(c_3, c_4)$ that depend only on $c_0$ and $c_1$.

Lemma 3.1. Let $A \in \mathbb{R}^{n \times n}$ be nonsymmetric but definite and let $A_0 \in \mathbb{R}^{n \times n}$ be s.p.d. Let $(A, A_0)$ be a pair of matrices that satisfy assumption (H0) with constants $(c_0, c_1)$ (in particular, $A \in \mathbb{R}^{n \times n}$ is nonsymmetric but definite and $A_0 \in \mathbb{R}^{n \times n}$ is s.p.d.). Then, the pair $(A^{-1}, A_0^{-1})$ also satisfies assumption (H0) with constants $(c_3, c_4)$; that is,

\begin{align}
(3.1) & \quad v^T A^{-1} v \geq c_0 c_3 v^T A_0^{-1} v, \quad \text{for all } v \in \mathbb{R}^n, \\
(3.2) & \quad w^T A^{-1} v \leq \frac{1}{c_0 c_3} \sqrt{v^T A_0^{-1} v} \sqrt{w^T A_0^{-1} w}, \quad \text{for all } v, w \in \mathbb{R}^n.
\end{align}

Proof. We first show the boundedness estimate (3.2). We define the matrix $Y := A_0^{-\frac{1}{2}} AA_0^{-\frac{1}{2}}$. Then, (2.3) and (2.4) imply (or read) that $Y$ satisfy:

\begin{align}
(3.3) & \quad v^T Y v \geq c_0 \|v\|^2, \quad \text{for all } v \in \mathbb{R}^n,
\end{align}
\( w^T Y v \leq c_1 \|v\|\|w\|, \quad \text{for all } v, w \in \mathbb{R}^n. \)

The positivity (3.3) of \( Y \) guarantees the existence of \( Y^{-1} \) and so taking \( v := Y^{-1}w \) in (3.3), and using the symmetry of the standard \( \ell^2 \)-inner product of two vectors together with the Cauchy Schwarz inequality, we find
\[
    c_0 \|Y^{-1}w\|^2 \leq w^T Y^{-T} w = w^T Y^{-1} w \leq \|w\|\|Y^{-1}w\|
\]
which shows that \( \|Y^{-1}w\| \leq \frac{1}{c_0} \|w\| \), that is, the boundedness of \( Y^{-1} \) in the \( \ell^2 \)-norm:
\[
(3.5) \quad \|Y^{-1}\| \leq \frac{1}{c_0}.
\]

In other words we have shown that
\[
    w^T A_0^{-\frac{1}{2}} A^{-1} A_0^{-\frac{1}{2}} v = w^T Y^{-1} v \leq \frac{1}{c_0} \|v\|\|w\| \quad \forall v, w \in \mathbb{R}^n.
\]

Setting now in the above equation \( v := A_0^{-\frac{1}{2}} v \) and \( w := A_0^{-\frac{1}{2}} w \), we reach the desired boundedness estimate (3.2) for \( A^{-1} \) in terms of \( A_0^{-1} \).

The positivity estimate (3.1) can be shown as follows. On the one hand, the boundedness (3.4) of \( Y \) with \( v = v \) and \( w = Y^{-1} v \) gives
\[
    \|v\|^2 = v^T Y (Y^{-1} v) \leq c_1 \|Y^{-1} v\| \|v\| \quad \forall v \in \mathbb{R}^n.
\]
which readily implies
\[
(3.6) \quad \|Y^{-1} v\| \geq \frac{1}{c_1} \|v\| \quad \forall v \in \mathbb{R}^n.
\]

On the other hand, using the positivity estimate (3.3) of \( Y \), we have
\[
v^T Y^{-1} v = (Y^{-1} v)^T Y^T (Y^{-1} v) = (Y^{-1} v)^T Y (Y^{-1} v) \geq c_0 \|Y^{-1} v\|^2 \quad \forall v \in \mathbb{R}^n.
\]
Then, the above relation together with estimate (3.6) give the following positivity estimate for \( Y^{-1} \):
\[
v^T Y^{-1} v \geq \frac{c_0}{c_1^2} \|v\|^2 \quad \forall v \in \mathbb{R}^n.
\]

Now, setting in the above estimate \( v := A_0^{-\frac{1}{2}} v \), we obtain the coercivity relation (3.1) and conclude the proof. \( \Box \)

Next, let \( B_0 \) denote the s.p.d. additive Schwarz preconditioner of \( A_0 \), whose inverse is defined as:
\[
(3.7) \quad B_0^{-1} = PA_c^{(0)-1} P^T + \sum_{k=1}^{N_s} I_k A_k^{(0)-1} I_k^T.
\]

Note that since \( (A, A_0) \) satisfy assumption (H0) with constants \((c_0, c_1)\), this immediately implies that for each \( k = 1, \ldots N_s \) the family of pairs \((A_k, A_k^{(0)})\) with matrices defined by
\[
A_k := I_k^T A I_k \quad \text{and} \quad A_k^{(0)} := I_k^T A_0 I_k \quad k = 1, \ldots N_s,
\]
also satisfy assumption (H0) with the same constants. The same is also true for the coarse pair of matrices \((A_c, A_c^{(0)})\), where \(A_c = P^T A P\) and \(A_c^{(0)} = P^T A_0 P\). Then, applying Lemma 3.1 to each of these pairs, we have that the corresponding pair of their respective inverses and hence the pair with the product matrices \(\left( I_k A_k^{-1} T_k^T, I_k A_k^{(0)} T_k^{-1} I_k^T \right)\) satisfies (H0) with constants \((c_0 c_1^2, c_0^{-1})\) (i.e, (3.1) and (3.2)). The latter implies that the inverses of the additive Schwarz preconditioners \(B^{-1}\) (as defined in (2.2)) and \(B_0^{-1}\) (as defined in (3.7)), are related in the same way (as their individual terms \(I_k A_k^{-1} T_k^T I_k^T\) and \(I_k A_k^{(0)} T_k^{-1} I_k^T\)). That is: \((B^{-1}, B_0^{-1})\) also satisfy (H0) with constants \((c_0 c_1^{-2}, c_0^{-1})\). Applying once more Lemma 3.1, we straightforwardly deduce that the pair \((B, B_0)\) also satisfy (H0), now with constants \((c_0^3 c_1^{-2}, c_1^2 c_0^{-1})\).

Now, since \(B_0\) is the classical s.p.d. additive Schwarz preconditioner for the s.p.d \(A_0\), \(B_0\) and \(A_0\) can be shown to be spectrally equivalent: there exists \(\gamma_0, \gamma_1 > 0\) such that

\[
\gamma_0 v^T B_0 v \leq v^T A_0 v \leq \gamma_1 v^T B_0 v \quad \forall \ v \in \mathbb{R}^n,
\]

where the constants \(\gamma_0\) and \(\gamma_1\) might depend on the parameters of the discretization and the problem. \(B_0\) would be optimal if neither \(\gamma_0\) nor \(\gamma_1\) depend on the discretization parameters (or size of the system \(n\)).

Using this extra information, it is straightforward to deduce that the pair \((B, A_0)\) also satisfy (H0) with constants \((\beta_0, \beta_1)\) that depend only on \(c_0, c_1, \gamma_0\) and \(\gamma_1\). All these observations are summarized in the following Lemma:

**Lemma 3.2.** Let \((A, A_0)\) satisfy assumption (H0) with constants \((c_0, c_1)\). Let \(B\) be the additive Schwarz preconditioner of \(A\) (defined through (2.2)) and let \(B_0\) be the corresponding s.p.d. additive Schwarz preconditioner of \(A_0\) (defined through (3.7)) and assume \(B_0\) is such that

\[
\gamma_0 v^T B_0 v \leq v^T A_0 v \leq \gamma_1 v^T B_0 v, \quad \forall \ v \in \mathbb{R}^n,
\]

for some \(\gamma_0, \gamma_1 > 0\). Then, the pair \((B, A_0)\) also satisfies (H0) with constants \((\beta_0, \beta_1)\):

\[
v^T B v \geq \beta_0 v^T A_0 v, \quad \text{for all } v,
\]

and

\[
w^T B v \leq \beta_1 \sqrt{v^T A_0 v \sqrt{w^T A_0 w}}, \quad \text{for all } v, w.
\]

The constants \(\beta_0\) and \(\beta_1\) are given by

\[
\beta_0 = \frac{c_0^3}{c_1 \gamma_1} \quad \beta_1 = \frac{c_1^2}{c_0 \gamma_0}.
\]

We note that the same results, (3.9)-(3.10), hold for \(B\) replaced with \(B^T\).
With all this relations at hand, we define the s.p.d. matrix
\begin{equation}
Z := BA_0^{-1}B^T.
\end{equation}
that can be used as a preconditioner for $A$ in GMRES. Observe that the actions of $Z^{-1}$ involve actions of both $B^{-1}$ and $B^{-T}$ as well as multiplications with $A_0$ (not $A_0^{-1}$). Therefore, the preconditioner $Z$ is computationally feasible.

We next prove the main result of the section, which guarantees that the preconditioned GMRES method for $A$ with the s.p.d. preconditioner $Z = BA_0^{-1}B^T$ will be convergent with bounds depending only on the constants involved in relations between $A$ and $Z$.

**Theorem 3.1.** Let $(A, A_0)$ satisfy assumption (H0) with constants $(c_0, c_1)$ and let $B \in \mathbb{R}^{n \times n}$ be the additive Schwarz preconditioner for $A$, whose inverse is defined through (2.2). Let $Z := BA_0^{-1}B^T$ be a preconditioner for $A$. Then, the pair $(A, Z)$ also satisfies (H0) with constants $(\alpha_0, \alpha_1)$ defined in (3.14).

Furthermore, the preconditioned GMRES method for $A$ with the s.p.d. preconditioner $Z$ converges with bounds:
\begin{equation}
\|r_m\|_{Z^{-1}} = \|r_m\|_{Z} \leq \left(1 - \sqrt{\frac{\alpha_0}{\alpha_1}}\right) \frac{\|\tau_0\|_Z}{\|\tau_0\|_{Z^{-1}}},
\end{equation}
where $r_m = Z^{-1}r_m = Z^{-1}(b - Ax_m)$ is the preconditioned residual at the $m$-th iteration with $r_0 = Z^{-1}r_0 := Z^{-1}(b - Ax_0)$; $\|\cdot\|_Z$ and $\|\cdot\|_{Z^{-1}}$ are the inner-product norms induced by the s.p.d matrices $Z$ and $Z^{-1}$, respectively.

**Proof.** From Lemma 3.2, we know that $(B, A_0)$ satisfy (H0) with $(\beta_0, \beta_1)$. In particular, the relations (3.9)-(3.10) (used for $B^T$) show that $X := A_0^{-\frac{1}{2}}B^TA_0^{-\frac{1}{2}}$ is well-conditioned. More precisely, we have
\begin{equation}
\beta_0 \|v\|^2 \leq \|Xv\|^2 \leq \beta_1 \|v\|^2 \quad \text{for all } v \in \mathbb{R}^n.
\end{equation}
That is, the s.p.d. matrix $X^TX$ is well-conditioned. The coercivity of $A$ in terms of $A_0$ expressed in (2.3), and $X^TX$ being well-conditioned (or, bounded) imply that $A_0^{\frac{1}{2}}AA_0^{-\frac{1}{2}}$ is coercive also in terms of $X^TX$:
\begin{equation}
v^T A_0^{\frac{1}{2}}AA_0^{-\frac{1}{2}}v \geq c_0 \|v\|^2 \geq \frac{c_0}{\beta_1} v^T X^TXv \quad \text{for all } v \in \mathbb{R}^n.
\end{equation}
Hence, $A$ is coercive in terms of $A_0^{\frac{1}{2}}X^TXA_0^{\frac{1}{2}} = BA_0^{-1}B^T = Z$, which is the first desired result.

Similarly, the boundedness of $A$ in terms of $A_0$, expressed in (2.4), and $X^TX$ being well-conditioned (or coercive) imply that $A_0^{-\frac{1}{2}}AA_0^{-\frac{1}{2}}$ is bounded also in terms of $X^TX$:
\begin{equation}
w^T A_0^{-\frac{1}{2}}AA_0^{-\frac{1}{2}}v \leq c_1 \|w\| \|v\| \leq \frac{c_1}{\beta_0} \sqrt{w^T X^TXw} \sqrt{v^T X^TXv}, \quad \text{for all } v, w \in \mathbb{R}^n,
\end{equation}
which is equivalent to say that $A$ is bounded in terms of $A^\frac{1}{2} X^T X A^\frac{1}{2} = B A_0^{-1} B^T = Z$. This completes the proof that the pair $(A, Z)$ verifies assumption $(H0)$ with constants $(\alpha_0, \alpha_1)$ defined by:

$$
\alpha_0 = \frac{c_0}{\beta_1} = \frac{C_0}{C_1} \cdot \gamma_0 \quad \alpha_1 = \frac{C_1}{\beta_0} = \frac{C_2}{\gamma_0}
$$

A standard application of the GMRES convergence theory [9] gives (3.13), and the proof of the theorem is complete.

4. A combined preconditioner

In this section, we introduce another preconditioner which in a sense combines the best of both preconditioners $B$ and $Z = B A_0^{-1} B^T$. We define its inverse $B^{-1}$ by forming the linear combination

$$
B^{-1} = B^{-1} + \sigma Z^{-1}.
$$

The parameter $\sigma \in \mathbb{R}$ is allowed to change from iteration to iteration inside the GMRES iterative solver. Therefore, $B$ can be regarded as a variable-step, flexible, preconditioner.

Observe that for $\sigma \geq 0$, by virtue of the analysis of the previous section, the pair $(B^{-1}, A_0^{-1})$ verifies assumption $(H0)$; i.e, $B^{-1}$ is coercive and bounded in $A_0^{-1}$ norm. We now describe the (practical) construction of the preconditioner $B^{-1}$, but considering a more general form:

$$
B^{-1} = \alpha B^{-1} + \sigma Z^{-1}
$$

(4.1)

without assuming the coefficients $\alpha$ and $\sigma$ to have nonnegative sign. At the end of the section we provide the convergence result for $B^{-1}$ which asserts faster convergence within GMRES than the one obtained with the preconditioner $Z^{-1}$.

4.1. Construction of the variable preconditioner. We consider the system of equations (2.1) that we solve by the preconditioned GMRES method with preconditioner $B^{-1}$ as defined in (4.1). We now explain how the coefficients are $\alpha$ and $\sigma$ set inside the GMRES iteration. Let $\| \cdot \| = \sqrt{(\cdot, \cdot)}$ and $\| \cdot \|_* = \sqrt{(\cdot, \cdot)_*}$ be two inner product norms, to be specified and chosen later on, and whose role will become clear in the process.

For $m \geq 0$, we denote by $x_m$ the $m^{th}$-iterate and by $r_m = b - A x_m$ the residual. At the $(m + 1)^{th}$ iteration of GMRES, we construct the new search direction $d_{m+1}$ based not only on the previous search directions $\{d_j\}_{j=0}^m$ but also on the two preconditioned residuals $B^{-1} r_m$ and $Z^{-1} r_m$, as follows:

$$
\beta_{m+1} d_{m+1} = \beta_{m+\frac{1}{2}} B^{-1} r_m + \beta_{m+\frac{2}{2}} Z^{-1} r_m + \sum_{j=0}^m \beta_j d_j.
$$
Here, the coefficients $\beta_j$, $j = 0, 1, \ldots, m$ are chosen such that
\[(d_{m+1}, d_j)_* = 0 \quad \text{for} \quad j < m + 1 \quad \text{and} \quad \|d_{m+1}\|_* = \sqrt{(d_{m+1}, d_{m+1})}_* = 1.\]

It is clear then, that the coefficients $\beta_{m+s}$, $s = 1, 2$ can be considered arbitrary parameters at this point. For any such fixed pair in GMRES, the next iterate $x_{m+1}$ is then computed by minimizing the residual:
\[
\|b - Ax_{m+1}\| = \|b - A(x_m + \sum_{j=0}^{m+1} \alpha_j d_j)\| \longrightarrow \min
\]
with respect to the coefficients $\{\alpha_j\}_{j=0}^{m+1}$. Notice that out of the two coefficients $\beta_{m+s}$, $s = 1, 2$, only their ratio
\[
\sigma = \sigma_{m+1} \equiv \frac{\beta_{m+2}}{\beta_{m+1}},
\]
can be considered as a free parameter (the rest is compensated by the $\alpha_{m+1}$-coefficient).

In practice, we proceed as follows. At step $m+1$, based on the previous search directions $\{d_j\}_{j=0}^m$ and the preconditioned residuals $B^{-1}r_m$ and $Z^{-1}r_m$, we have to solve the minimization problem:
\[
\|b - A(x_m + \sum_{j=0}^{m} \alpha_j d_j + \alpha_{m+1} B^{-1}r_m + \alpha_{m+2} Z^{-1}r_m)\| \longrightarrow \min,
\]
with respect to the coefficients $\{\alpha_j\}_{j=0}^{m}$, and $\alpha_{m+s}$, $s = 1, 2$. As we show next the solution of such minimization problem can be reduced to the solution of a $2 \times 2$ system, by choosing appropriately the inner product $(\cdot, \cdot)_*$.

Consider the quadratic functional $J(\alpha)$
\[
J(\alpha) \equiv \left\|r_m - A\left(\sum_{j=0}^{m} \alpha_j d_j + \alpha_{m+1} B^{-1}r_m + \alpha_{m+2} Z^{-1}r_m\right)\right\|^2 \longrightarrow \min.
\]
as a function of the coefficients $\alpha = (\alpha_r)$. Set the inner product $(\cdot, \cdot)_* = (A(\cdot), A(\cdot))$, which is equivalent to assuming that the search directions $\{d_j\}$ are $(A(\cdot), A(\cdot))$ orthogonal. Then, we have
\[
0 = \frac{1}{2} \frac{\partial J}{\partial \alpha_j} = \alpha_j - (r_m - \alpha_{m+1} AB^{-1}r_m - \alpha_{m+2} AZ^{-1}r_m, A d_j), \quad j \leq m,
\]
which gives
\[
(4.2) \quad \alpha_j = (r_m, A d_j) - \alpha_{m+1} (AB^{-1}r_m, A d_j) - \alpha_{m+2} (AZ^{-1}r_m, A d_j), \quad j \leq m.
\]
Setting now the partial derivatives of $J$ w.r.t. $\alpha_{m+\frac{1}{3}}$ to zero, we get (4.3)

\[
\alpha_{m+\frac{1}{3}} \parallel AB^{-1}r_m \parallel^2 - (r_m - A\left( \sum_{j=0}^{m} \alpha_j d_j + \alpha_{m+\frac{2}{3}} Z^{-1}r_m \right), AB^{-1}r_m) = 0,
\]

\[
\alpha_{m+\frac{2}{3}} \parallel AZ^{-1}r_m \parallel^2 - (r_m - A\left( \sum_{j=0}^{m} \alpha_j d_j + \alpha_{m+\frac{1}{3}} B^{-1}r_m \right), AZ^{-1}r_m) = 0.
\]

Substituting $\alpha_j$ from (4.2) into (4.3) we end up with a system of two equations where the only two unknowns are the coefficients $\alpha_{m+\frac{s}{3}}$ with $s = 1, 2$:

\[
\begin{align*}
\bullet \quad & \alpha_{m+\frac{1}{3}} \left( \parallel AB^{-1}r_m \parallel^2 - \sum_j (AB^{-1}r_m, Ad_j)^2 \right) \\
& + \alpha_{m+\frac{2}{3}} \left( (AZ^{-1}r_m, AB^{-1}r_m) - \sum_j (AZ^{-1}r_m, Ad_j)(AB^{-1}r_m, Ad_j) \right) \\
& = (r_m, AB^{-1}r_m) - \sum_j (r_m, Ad_j)(AB^{-1}r_m, Ad_j) \\
& = (r_m, AB^{-1}r_m) - \sum_j (AB^{-1}r_m, Ad_j)Ad_j,
\end{align*}
\]

\[
\begin{align*}
\bullet \quad & \alpha_{m+\frac{2}{3}} \left( (AZ^{-1}r_m, AB^{-1}r_m) - \sum_j (AB^{-1}r_m, Ad_j)(AZ^{-1}r_m, Ad_j) \right) \\
& + \alpha_{m+\frac{1}{3}} \left( \parallel AZ^{-1}r_m \parallel^2 - \sum_j (AZ^{-1}r_m, Ad_j)^2 \right) \\
& = (r_m, AZ^{-1}r_m) - \sum_j (r_m, Ad_j)(AZ^{-1}r_m, Ad_j) = \\
& = (r_m, AZ^{-1}r_m) - \sum_j (AZ^{-1}r_m, Ad_j)Ad_j.
\end{align*}
\]

To show the solvability of the above system for $\alpha_{m+\frac{s}{3}}$ with $s = 1, 2$ (which will imply that the preconditioner $B^{-1}$ is well defined), we use the following lemma.

**Lemma 4.1.** Let $(H, (.,.))$ be a Hilbert space with inner product $(.,.)$ and let $h, f, g \in H$. Let $S$ be a finite dimensional subspace of $H$ spanned by an orthonormal system $\{p_j\}_{j=1}^{m}$, i.e., $(p_i, p_j) = \delta_{i,j}$. Let $\pi = \pi_S : H \rightarrow S$ be the orthogonal projection on $S$, with respect to the inner product $(.,.)$. Then, the best approximation to $h$ from elements from $S$ augmented by the two vectors $f$ and $g$ is given as the solution of the least squares (or minimization) problem

\[
\begin{equation}
\min_{r=\frac{1}{3}, \frac{2}{3}, \ldots, m} \left\| h - \alpha_{\frac{1}{3}} f - \alpha_{\frac{2}{3}} g - \sum_j \alpha_j p_j \right\| \rightarrow \min
\end{equation}
\]

over the coefficients $\{\alpha_r\}$, $r = \frac{1}{3}, \frac{2}{3}, 1, \ldots, m$. Solving problem (4.5) is equivalent to solve the following two-by-two system:

\[
\begin{pmatrix}
\parallel (I - \pi)f \parallel^2 & ((I - \pi)f, (I - \pi)g) \\
((I - \pi)f, (I - \pi)g) & \parallel (I - \pi)g \parallel^2
\end{pmatrix}
\begin{pmatrix}
\alpha_{\frac{1}{3}} \\
\alpha_{\frac{2}{3}}
\end{pmatrix}
= \begin{pmatrix}
(h, (I - \pi)f) \\
(h, (I - \pi)g)
\end{pmatrix},
\]
which has a unique solution provided \((I - \pi)f\) and \((I - \pi)g\) are linearly independent. If \((I - \pi)f\) and \((I - \pi)g\) are linearly dependent, there is also a solution, since the r.h.s. in (4.6) is compatible.

The remaining coefficients \(\{\alpha_j\}\) are computed from \(\pi(h - \alpha_1 f - \alpha_2 g) = \sum \alpha_j p_j\), that is:

\[
\alpha_j = (h - \alpha_1 f - \alpha_2 g, p_j).
\]

**Proof.** It is clear that the least-squares problem (4.5) reduces to finding the best approximation to \((I - \pi)h\) from the space spanned by the two vectors \((I - \pi)f\) and \((I - \pi)g\). Indeed, we can rewrite (4.5) as

\[
\|(I - \pi)h - \alpha_1 (I - \pi)f - \alpha_2 (I - \pi)g - \sum_j \alpha_j' p_j\| \mapsto \min.
\]

Since the last component \(\sum_j \alpha_j' p_j\) is orthogonal to \((I - \pi)\left(h - \alpha_1 f - \alpha_2 g\right)\), the above minimum equals

\[
\min_{\alpha_1, \alpha_2} \min_{\alpha_j} \left\| (I - \pi)h - \alpha_1 (I - \pi)f - \alpha_2 (I - \pi)g - \sum_j \alpha_j' p_j \right\|
\]

\[
= \min_{\alpha_1, \alpha_2} \min_{\alpha_j} \left( \| (I - \pi)h - \alpha_1 (I - \pi)f - \alpha_2 (I - \pi)g \|^2 + \| \sum_j \alpha_j' p_j \|^2 \right)^{\frac{1}{2}}
\]

\[
= \min_{\alpha_1, \alpha_2} \| (I - \pi)h - \alpha_1 (I - \pi)f - \alpha_2 (I - \pi)g \|
\]

\[
= \min_{\alpha_1, \alpha_2} \left( \| h - \alpha_1 f - \alpha_2 g \|^2 - \| \pi h \|^2 \right)^{\frac{1}{2}}.
\]

The last problem leads exactly to the Gram system (4.6). This completes the proof. \(\square\)

We now apply last Lemma to our case, to show that the system (4.4) has a solution and hence \(B^{-1}\) is well defined. We set \(h = A^{-1} r_m, f = B^{-1} r_m, g = Z^{-1} r_m\) and \(\{p_j\} = \{d_j\}_{j=0}^m\) for the vector space with inner-product \((\cdot, \cdot)_* = (A(\cdot), A(\cdot))\). Using then that the \(\{d_j\}\) are \((\cdot, \cdot)_*\)-orthonormal, we conclude by applying Lemma 4.1 that the system (4.4) is solvable.

Once the coefficients \(\alpha_{m+\frac{s}{2}}, s = 1, 2\) are determined, we compute the new direction \(d_{m+1}\) from

\[
\beta_{m+1} d_{m+1} = \alpha_{m+\frac{1}{2}} B^{-1} r_m + \alpha_{m+\frac{3}{2}} Z^{-1} r_m - \sum_{j=0}^m \beta_j d_j,
\]

by choosing the coefficients \(\{\beta_j\}_{j=0}^m\) to satisfy the required orthogonality conditions

\[
(d_{m+1}, d_j)_* = (Ad_{m+1}, Ad_j) = 0 \quad \text{for} \ j < m + 1,
\]
which gives (assuming by induction that \((d_j, d_k)_* = \delta_{j,k}\))

\[
\beta_j = \left(\alpha_{m+\frac{1}{4}} B^{-1} r_m + \alpha_{m+\frac{3}{4}} Z^{-1} r_m, d_j\right)_* \quad \text{for } j \leq m.
\]

The last coefficient, \(\beta_{m+1}\), is computed so that \(\|d_{m+1}\|_* = 1\).

4.2. Convergence. We close the section, by giving a result that provides an estimate for the convergence of the variable preconditioned GMRES.

**Theorem 4.1.** Let \((A, A_0)\) satisfy assumption \((H0)\) with constants \((c_0, c_1)\) and let \(B \in \mathbb{R}^{n \times n}\) be the additive Schwarz preconditioner for \(A\), whose inverse is defined through (2.2). Let \(Z := BA_0^{-1} B^T\) be a preconditioner for \(A\). Let \(B\) be the variable preconditioner with inverse defined through (4.1) with coefficients determined inside the GMRES iteration by minimization of the residual. Then, the variable preconditioned GMRES method for \(A\) converges faster than the preconditioned GMRES method with preconditioner \(Z\).

**Proof.** The proof of the Theorem follows by the definition of \(B^{-1}\) (as explained before). From its construction it is straightforward to infer the following comparative convergence estimate

\[
\|r_{m+1}\| \leq \min_{\alpha, \sigma} \|r_m - A(\alpha B^{-1} + \sigma Z^{-1})r_m\| \leq \min_{\sigma} \|r_m - \sigma AZ^{-1}r_m\|.
\]

By choosing now \(\|\cdot\|\) as the norm \(\|v\|_{Z^{-1}} = \sqrt{v^T Z^{-1} v}\), the combined preconditioned GMRES method will converge faster than the corresponding GMRES with preconditioner \(Z\) (that satisfies estimate (3.13) as provided in Theorem 3.1).

5. Applications and numerical results

In this section we present an application of the results and framework presented in the previous sections, that will allow us to verify the developed theory and will also assess the performance of the different preconditioners.

The application we consider comes from a nonsymmetric Discontinuous Galerkin discretization of an elliptic problem. In [2, 3], additive and multiplicative Schwarz preconditioners were developed for the solution of the above algebraic system. In both works, the authors show that the GMRES convergence theory cannot be applied for explaining the observed convergence since the preconditioned system does not satisfy the sufficient conditions for such theory. Here we aim at comparing the performance of the different preconditioners introduced in the previous sections, for such discretizations.

More precisely, we consider the following model problem in \(\Omega = [0,1]^2\):

\[
-\Delta u^* = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial \Omega,
\]

where the right hand side \(f\) is chosen so the exact solution is \(u^* = \sin(\pi x) \sin(\pi y)\), and we focus on the Incomplete Interior Penalty Discontinuous Galerkin (IIPG) [8] discretization of the above model problem, with linear discontinuous finite
element space (denoted by $V^{DG}$) on a shape regular triangulation of $\Omega$, denoted by $T_h$. The resulting method reads: Find $u \in V^{DG}$ such that

$$a_h(u, v) = \int_{\Omega} f v \, dx \quad \forall \ v \in V^{DG}.$$  

The bilinear form of the IIPG method is given by

$$a_h(u, v) = \sum_{K \in T_h} \int_K \nabla u \cdot \nabla v \, dx - \sum_{e \in E_h} \int_e \{ \{ \nabla u \} \} \cdot [ \{ v \} ] \, ds + \sum_{e \in E_h} \frac{\eta}{|e|} \int_e [u] \cdot [v] \, ds,$$

for all $u, v \in V^{DG}$. Here, $K \in T_h$ refer to an element of the triangulation, $e \subset \partial K$ denotes an edge of the element and we have denote by $E_h$ the set of all such edges or skeleton of the partition $T_h$. We have used the standard definition of the average and jump operators from [4], and the penalty parameter $\eta$ is set to 5 in all the experiments. We denote by $A_h$ the matrix representation of the operator associated to the bilinear form (5.1), with standard Lagrange basis functions. Similarly, $u$ and $f$ denote the vector representations (in the same basis) of the solution (that we aim to compute) and the right hand side. In the end, the solution process amount to solve the nonsymmetric system:

$$A_h u = f.$$

The preconditioners we use are based on the standard two–level overlapping domain decomposition additive Schwarz preconditioner which we denote by $B^{-1}$. To define it, we consider an overlapping partition of $\Omega$ into rectangular subdomains $\Omega_k$ which overlap each other by an amount equal to the fine discretization size $h$. Then

$$B^{-1} = I_H A_H^{-1} I_H^T + \sum_{k=1}^{N_s} I_k A_k^{-1} I_k^T,$$

Here the $A_k$ operators are restrictions of the original operator $A_h$ to the finite element space $V_k$ that is only supported on $\Omega_k$, that is they correspond to the bilinear forms,

$$a_k(u, v) = a_h(u, v), \quad \forall u, v \in V_k,$$

as in [11, 7]. Since $V_k \subset V_h$, the operators $I_k$ are standard injection. The operator $A_H$ corresponds to the bilinear form (5.1) on a coarser discretization of the original domain $\Omega$, where we label the coarse discretization size $H$. We assume that the fine mesh is a refinement of the coarse mesh used to represent $A_H$ so that $I_H$ is the natural injection on nested grids. The penalty parameter on the coarse grid is taken to be $5H/h$ in order to account for the difference of scales in the edge lengths in the penalty terms (see [2, 11], for further details). We implement these preconditioners on a parallel machine so that each subdomain is assigned to a processor and the subdomain solves can be done in parallel.

Another preconditioner we consider is

$$Z^{-1} = B^{-T} A_0 B^{-1},$$
as outlined in the analysis above. Here $A_0$ is a symmetric operator corresponding to the bilinear form

$$a_0(u, v) = \sum_K \int_K \nabla u \cdot \nabla v \, dx + \sum_e \eta_0 \left| e \right| \int_e [u] \cdot [v] \, ds,$$

where the penalty parameter $\eta_0 = 5$ is the same as it is in the full bilinear form (5.1).

In what follows we consider four different preconditioning techniques for the nonsymmetric system (5.2), namely

1. the standard additive Schwarz preconditioner $B^{-1}$ used in a right–preconditioned GMRES algorithm, for comparison with the other options,
2. the preconditioner $Z^{-1} = B^{-T}A_0B^{-1}$ again used in a right–preconditioned GMRES,
3. the flexible two–preconditioner GMRES variant with the two preconditioners $B^{-1}$ and $Z^{-1}$,
4. and the two–preconditioner GMRES variant with $B^{-1}$ and $B^{-T}$ as the two preconditioners.

We note that in the third case, if we have applied $B^{-1}$ to a vector $u$ we can construct $Z^{-1}u$ by applying $B^{-T}A_0$ to save ourselves one preconditioner application.
Table 3. Time to solution for $h = 2^{-7}, H = 2^{-5}$ with a nearly exact coarse solver and no restart.

<table>
<thead>
<tr>
<th>$N_s$</th>
<th>$B^{-1}$ and $Z^{-1}$</th>
<th>$B^{-1}$ and $B^{-T}$</th>
<th>$Z^{-1}$</th>
<th>$B^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.74</td>
<td>1.79</td>
<td>3.12</td>
<td>0.84</td>
</tr>
<tr>
<td>8</td>
<td>1.39</td>
<td>1.57</td>
<td>3.22</td>
<td>0.74</td>
</tr>
<tr>
<td>16</td>
<td>0.70</td>
<td>0.75</td>
<td>1.58</td>
<td>0.34</td>
</tr>
<tr>
<td>32</td>
<td>1.66</td>
<td>1.83</td>
<td>4.63</td>
<td>0.94</td>
</tr>
<tr>
<td>64</td>
<td>1.44</td>
<td>1.66</td>
<td>4.65</td>
<td>0.88</td>
</tr>
<tr>
<td>128</td>
<td>2.70</td>
<td>2.98</td>
<td>8.59</td>
<td>1.76</td>
</tr>
</tbody>
</table>

Table 4. Iterations to convergence for $h = 2^{-7}, H = 2^{-5}$ with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>$N_s$</th>
<th>$B^{-1}$ and $Z^{-1}$</th>
<th>$B^{-1}$ and $B^{-T}$</th>
<th>$B^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>15</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>32</td>
<td>16</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>64</td>
<td>16</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>128</td>
<td>16</td>
<td>18</td>
<td>19</td>
</tr>
</tbody>
</table>

The number of GMRES iterations necessary to reduce the relative residual by $10^{-6}$ for our four different preconditioning approaches using various numbers of subdomains $N_s$ for a fixed problem is given in Table 1. Here we solve the coarse problem involving $A_H^{-1}$ to a tolerance of $10^{-10}$ so that this solve is nearly exact in order to satisfy the theory more closely. The preconditioning techniques are seen to be scalable in the sense that the number of iterations does not increase as $N_s$ increases for all four methods. Similarly, we show the convergence rates in Table 2, where the convergence rate is defined as the factor by which the true residual is reduced in the last iteration. To get an idea of computational cost, we show the time to solution in seconds for the four approaches in Table 3. We conclude that the $Z^{-1}$ preconditioner is not competitive because it is the most expensive in terms of time per iteration and it also requires the most iterations. For this reason we do not consider it further in these numerical results. The two preconditioning techniques that use the two–preconditioner GMRES variant are seen to be effective in convergence rate but to be somewhat more expensive than the classical $B^{-1}$ preconditioner, as we might expect.

In practice for parallel computing applications the coarse solve in (5.3) would not be done exactly. Another modification that is often made in practice is to restart GMRES after several iterations. In Tables 4, 5, and 6 we repeat the previous experiment where the relative residual tolerance for the coarse solves is
Table 5. Convergence rate for $h = 2^{-7}$, $H = 2^{-5}$ with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>$N_s$</th>
<th>$B^{-1}$ and $Z^{-1}$</th>
<th>$B^{-1}$ and $B^{-T}$</th>
<th>$B^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.19</td>
<td>0.40</td>
<td>0.49</td>
</tr>
<tr>
<td>8</td>
<td>0.36</td>
<td>0.42</td>
<td>0.45</td>
</tr>
<tr>
<td>16</td>
<td>0.34</td>
<td>0.38</td>
<td>0.44</td>
</tr>
<tr>
<td>32</td>
<td>0.32</td>
<td>0.39</td>
<td>0.44</td>
</tr>
<tr>
<td>64</td>
<td>0.23</td>
<td>0.32</td>
<td>0.39</td>
</tr>
<tr>
<td>128</td>
<td>0.25</td>
<td>0.42</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Table 6. Time to solution for $h = 2^{-7}$, $H = 2^{-5}$ with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>$N_s$</th>
<th>$B^{-1}$ and $Z^{-1}$</th>
<th>$B^{-1}$ and $B^{-T}$</th>
<th>$B^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.59</td>
<td>1.73</td>
<td>0.99</td>
</tr>
<tr>
<td>8</td>
<td>1.29</td>
<td>1.35</td>
<td>0.72</td>
</tr>
<tr>
<td>16</td>
<td>0.59</td>
<td>0.62</td>
<td>0.35</td>
</tr>
<tr>
<td>32</td>
<td>0.73</td>
<td>0.79</td>
<td>0.46</td>
</tr>
<tr>
<td>64</td>
<td>0.67</td>
<td>0.71</td>
<td>0.44</td>
</tr>
<tr>
<td>128</td>
<td>0.89</td>
<td>0.95</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Set to $10^{-4}$ and GMRES is restarted every 10 iterations. (These tables should be compared to Tables 1, 2, and 3 respectively.) We see that the convergence behavior is quite similar and the computational cost is lower, suggesting that these common modifications are also useful and effective for our preconditioning strategies.

To see how these methods scale to larger problems, we consider a much finer mesh in Tables 7, 8, and 9, while keeping the mesh size for the coarse solve quite coarse. The scalability of the preconditioners in terms of iterations is still present, the preconditioner still performs well, and in these cases we can see fairly good parallel scalability in the sense that for a fixed problem size, doubling the number of processors in the parallel solve roughly cuts the execution time in half for all our preconditioning strategies.

In all the results we have presented so far, the flexible two-preconditioner GMRES variant has performed slightly better than the classical $B^{-1}$ preconditioner in terms of number of iterations to convergence, but the increased computational cost per iteration of this GMRES variant has ended up making the classical preconditioner perform better in execution time. This suggests that the new method may be competitive in settings where each iteration is very expensive, so that the savings in iteration count can make up for the increased cost per iteration. To examine this setting we consider a problem in Tables 10 and 11.
Table 7. Iterations to convergence for \( h = 2^{-10}, H = 2^{-6} \) with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>( N_s )</th>
<th>( B^{-1} ) and ( Z^{-1} )</th>
<th>( B^{-1} ) and ( B^{-T} )</th>
<th>( B^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>30</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>64</td>
<td>30</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>128</td>
<td>30</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>256</td>
<td>30</td>
<td>31</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 8. Convergence rate for \( h = 2^{-10}, H = 2^{-6} \) with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>( N_s )</th>
<th>( B^{-1} ) and ( Z^{-1} )</th>
<th>( B^{-1} ) and ( B^{-T} )</th>
<th>( B^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>0.49</td>
<td>0.67</td>
<td>0.64</td>
</tr>
<tr>
<td>64</td>
<td>0.56</td>
<td>0.68</td>
<td>0.66</td>
</tr>
<tr>
<td>128</td>
<td>0.52</td>
<td>0.64</td>
<td>0.66</td>
</tr>
<tr>
<td>256</td>
<td>0.51</td>
<td>0.68</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Table 9. Time to solution for \( h = 2^{-10}, H = 2^{-6} \) with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>( N_s )</th>
<th>( B^{-1} ) and ( Z^{-1} )</th>
<th>( B^{-1} ) and ( B^{-T} )</th>
<th>( B^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>94.32</td>
<td>95.55</td>
<td>40.96</td>
</tr>
<tr>
<td>64</td>
<td>37.89</td>
<td>37.59</td>
<td>16.66</td>
</tr>
<tr>
<td>128</td>
<td>15.55</td>
<td>15.60</td>
<td>7.82</td>
</tr>
<tr>
<td>256</td>
<td>7.76</td>
<td>8.81</td>
<td>5.50</td>
</tr>
</tbody>
</table>

Table 10. Iterations to convergence for \( h = 2^{-10}, H = 2^{-9} \) with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>( N_s )</th>
<th>( B^{-1} ) and ( Z^{-1} )</th>
<th>( B^{-1} ) and ( B^{-T} )</th>
<th>( B^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>19</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td>64</td>
<td>17</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>128</td>
<td>15</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>256</td>
<td>15</td>
<td>16</td>
<td>20</td>
</tr>
</tbody>
</table>

where the coarse grid solve is done on a relatively fine grid and is therefore quite expensive. The results in this somewhat artificial setting do show that the new methods are competitive with the classical preconditioning techniques in terms of computational cost.
Table 11. Time to solution for $h = 2^{-10}, H = 2^{-9}$ with an inexact coarse solver and restarting every 10 iterations.

<table>
<thead>
<tr>
<th>$N_s$</th>
<th>$B^{-1}$ and $Z^{-1}$</th>
<th>$B^{-1}$ and $B^{-T}$</th>
<th>$B^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>277.50</td>
<td>280.69</td>
<td>352.85</td>
</tr>
<tr>
<td>64</td>
<td>170.94</td>
<td>187.29</td>
<td>133.73</td>
</tr>
<tr>
<td>128</td>
<td>79.01</td>
<td>93.61</td>
<td>99.50</td>
</tr>
<tr>
<td>256</td>
<td>40.92</td>
<td>45.41</td>
<td>39.59</td>
</tr>
</tbody>
</table>

REFERENCES


Blanca Ayuso de Dios  
CENTRE DE RECERCA MATEMÁTICA  
UAB SCIENCE FACULTY  
CAMPUS DE BELLATERRA  
08193 BELLATERRA, BARCELONA, SPAIN  
E-mail address: bayuso@crm.cat

Andrew T. Barker  
DEPARTMENT OF MATHEMATICS  
AND  
CENTER FOR COMPUTATION AND TECHNOLOGY  
LOUISIANA STATE UNIVERSITY  
BATON ROUGE, LA 70803–4918, U.S.A.  
E-mail address: andrewb@math.lsu.edu

Panayot S. Vassilevski  
CENTER FOR APPLIED SCIENTIFIC COMPUTING  
LAWRENCE LIVERMORE NATIONAL LABORATORY  
7000 EAST AVENUE, MAIL STOP L-560  
LIVERMORE, CA 94550, U.S.A.  
E-mail address: panayot@llnl.gov