Further comparison of additive and multiplicative coarse grid correction

Yvan Notay [*] and Artem Napov [†]

Service de Métrologie Nucléaire
Université Libre de Bruxelles (C.P. 165/84)
50, Av. F.D. Roosevelt, B-1050 Brussels, Belgium.

Postprint version of the paper appeared in
Applied Numerical Mathematics (vol. 65, pp. 53–62, 2013)
`http://dx.doi.org/10.1016/j.apnum.2012.12.001`

Abstract

We consider the situation where a basic preconditioner is improved with a coarse grid correction. The latter can be implemented either additively (like in the standard additive Schwarz method) or multiplicatively (like in the balancing preconditioner). In a previous study, Nabben and Vuik compare both variants, and state that a theoretical comparison of the condition numbers is not possible: whereas it is admitted that the condition number is in most cases smaller with the multiplicative variant, they provide an example for which the converse is true. Here we show that the multiplicative variant has in fact always lower condition number when the basic preconditioner is appropriately scaled. On the other hand, we also show, again assuming an appropriate scaling, that the condition number of the additive variant is at worst a modest multiple of that of the multiplicative variant. Hence both approaches are qualitatively equivalent. Eventually, we show with some examples that both the upper and lower bounds on the condition number of the additive variant are sharp: it can be in some cases equal to the condition number of the multiplicative variant, and in other cases arbitrarily close to the aforementioned modest multiple of this latter value.

Key words. multigrid, domain decomposition, convergence analysis, coarse grid correction, preconditioning

AMS subject classification. 65F10, 65F08, 65N55

---

[*] Supported by the Belgian F.R.S–FNRS ("Directeur de recherches"); *ynotay@ulb.ac.be*

[†] Supported by the Belgian F.R.S–FNRS ("Aspirant"); *anapov@ulb.ac.be*

1 Introduction

We consider the iterative solution of large sparse $n \times n$ linear systems

$$A\,\mathbf{u} \;=\; \mathbf{b} \tag{1.1}$$

with symmetric positive definite (SPD) coefficient matrix A. It is common to perform this solution by combining the conjugate gradient method [13] with some SPD preconditioner M that approximates the system matrix while remaining cheap to construct and invert (see, e.g., [1, 10, 25] for examples). The convergence rate then mainly depends on the spectral condition number

$$\kappa(M^{-1}A) \;=\; \frac{\lambda_{\max}(M^{-1}A)}{\lambda_{\min}(M^{-1}A)} \;,$$

where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ stand for the smallest and the largest eigenvalue, respectively.

In practical applications, A has often some very small eigenvalues, and standard preconditioners may fail to move them sufficiently away from 0; i.e., $M^{-1}A$ has still some fairly small eigenvalues. This may be cured by supplementing the action of the preconditioner with a coarse grid correction step, in which an approximate solution of the residual equation is computed on a coarser grid. This combination of basic preconditioner and coarse grid correction is at the root of multigrid methods (e.g., [12, 28]), where the basic preconditioner is called *smoother*. In the context of these methods, the number of coarse unknowns is generally large and the coarse grid system is therefore solved only approximately, combining again a smoothing iteration and a coarse grid correction (the latter is thus used recursively).

Coarse grid correction is also used in a number of other methods. Often, the (first) coarse grid has then sufficiently few unknowns to make affordable an exact solution of the corresponding system (no need for recursive use). This includes domain decomposition methods (e.g., [24, 26, 27]), where the coarse grid has usually only a few unknowns per subdomain, and also approaches where one attempts to approximate directly the eigenvectors of $M^{-1}A$ corresponding to small eigenvalues [9, 30, 31, 32]; see also [22, 23].

To define a coarse grid correction one needs to set up a $n \times n_c$ rectangular matrix which "prolongates" on the fine grid a vector defined on a coarse grid with n_c unknowns. Ideally, the range of P should span a subspace containing good approximations of the eigenmodes that are converging slowly with the basic preconditioner M. Given P, the coarse grid correction step is implemented with

$$B_c \;=\; P(P^TA\,P)^{-1}P^T \;.$$

There are basically two ways to combine it with the preconditioner M at hand. One is *additive* and amounts to use as inverse preconditioner (that is, as approximation to A^{-1})

$$B_{a\,,\omega} \;=\; \omega\,M^{-1} \;+\; P(P^TA\,P)^{-1}P^T \;, \tag{1.2}$$

where $\omega > 0$ is scaling parameter. This approach is followed in, e.g., [22, 23] and in standard two-level additive Schwarz methods [3, 5, 6, 24, 26, 27]. Often, the scaling is applied instead to the coarse grid correction term. This is however unimportant when $B_{a\,,\omega}$ is used as preconditioner for the conjugate gradient method since then only the relative scaling of both terms matters.

On the other hand, a *multiplicative* correction is obtained by first considering a stationary iteration with the basic preconditioner followed by a coarse grid correction step. The resulting iteration matrix is

$$T_{m\,,\omega} \;=\; (I - P(P^TA\,P)^{-1}P^T\,A)\,(I - \omega\,M^{-1}A)\;,$$

and the equivalent inverse preconditioner is the matrix $B_{m\,,\omega}$ such that $I - B_{m\,,\omega}\,A = T_{m\,,\omega}$; this yields

$$B_{m\,,\omega} \;=\; P(P^TA\,P)^{-1}P^T \;+\; \omega\,M^{-1} \;-\; \omega\,P(P^TA\,P)^{-1}P^T\,A\,M^{-1}\;. \tag{1.3}$$

This preconditioner is used, e.g., in the two-level hybrid Schwarz method as defined in [26, Algorithm 2.3.5]. It is nonsymmetric and can therefore not be used with the conjugate gradient method. A symmetric version is obtained by performing the coarse grid correction twice; that is, defining the inverse preconditioner $B_{m_s\,,\omega}$ such that $I - B_{m_s\,,\omega}\,A = T_{m_s\,,\omega}$ with

$$T_{m_s\,,\omega} \;=\; (I - P(P^TA\,P)^{-1}P^T\,A)\,(I - \omega\,M^{-1}A)\,(I - P(P^TA\,P)^{-1}P^T\,A)\;.$$

This yields

$$B_{m_s\,,\omega} \;=\; P(P^TA\,P)^{-1}P^T \;+\; \omega\,(I - P(P^TA\,P)^{-1}P^T\,A)\,M^{-1}\,(I - A\,P(P^TA\,P)^{-1}P^T)\;, \tag{1.4}$$

which is used in the two-level hybrid Schwarz method as defined in [27, eq. (2.11)], and in the balancing Neumann-Neumann domain decomposition methods [14, 16, 26, 27]. Because the form (1.4) first appeared in the latter context, this variant is sometimes called balancing preconditioner. $B_{m_s\,,\omega}$ is in general more costly to apply than $B_{m\,,\omega}$ but, as pointed out in [27, Section 2.5.2], if one uses the conjugate gradient method with $\mathbf{u}_0 = P(P^TA\,P)^{-1}P^T\mathbf{b}$ as initial approximation, all residual vectors are kept orthogonal to the range of P and hence the application of $B_{m_s\,,\omega}$ requires in practice only the multiplication by $\omega\,(I - P(P^TA\,P)^{-1}P^T\,A)\,M^{-1}$.

An important remark here is that $B_{m\,,\omega}\,A$ and $B_{m_s\,,\omega}\,A$ have identical eigenvalues. Indeed, for any pair of square matrices F and G , $F\,G$ and $G\,F$ have same set of eigenvalues (see, e.g., [21, Lemma A.1] for a proof covering the case where both F and G are singular). Hence, $T_{m_s\,,\omega} = T_{m\,,\omega}(I - P(P^TA\,P)^{-1}P^T\,A)$ has same eigenvalues as $(I - P(P^TA\,P)^{-1}P^T\,A)\,T_{m\,,\omega} = T_{m\,,\omega}$. Since $B_{m_s\,,\omega}$ is positive definite, the eigenvalues of $B_{m\,,\omega}\,A$ are therefore real and positive, and one has in particular

$$\lambda_{\max}(B_{m\,,\omega}\,A) \;=\; \lambda_{\max}(B_{m_s\,,\omega}\,A) \quad , \quad \lambda_{\min}(B_{m\,,\omega}\,A) \;=\; \lambda_{\min}(B_{m_s\,,\omega}\,A)$$

and

$$\kappa(B_{m,\omega} A) \;=\; \kappa(B_{m_s,\omega} A) \; .$$

In the following, we formulate our results with respect to the eigenvalues of $B_{m_s,\omega} A$, but one should keep in mind that they apply verbatim to $B_{m,\omega} A$ as well.

In the context of multigrid methods, the symmetrization is generally performed differently, applying smoothing iterations both before and after the coarse grid correction step. If M_1 is the pre-smoother and M_2 the post-smoother, the iteration matrix for a two-level method is then

$$T_{mg} \;=\; (I - M_2^{-1}A)\,(I - P(P^TA\,P)^{-1}P^T A)\,(I - M_1^{-1}A) \; , \tag{1.5}$$

and the corresponding inverse preconditioner B_{mg}, defined from $I - B_{mg}A = T_{mg}$, is symmetric when $M_1 = M_2^T$. Here, one may note that T_{mg} has same eigenvalues as

$$(I - P(P^TA\,P)^{-1}P^T A)\,(I - M_1^{-1}A)\,(I - M_2^{-1}A) \;=\; (I - P(P^TA\,P)^{-1}P^T A)\,(I - M^{-1}A) \; ,$$

where M is the matrix such that

$$(I - M^{-1}A) \;=\; (I - M_1^{-1}A)\,(I - M_2^{-1}A) \; ; \tag{1.6}$$

that is, M is the equivalent preconditioner which brings in one step the effect of post-smoothing followed by pre-smoothing. It follows that $B_{mg}A$ has same eigenvalues as $B_{m,1} A$ (and $B_{m_s,1}A$) with M defined in this way. Therefore, the analysis below also indirectly applies to the standard form of multigrid preconditioning. Note that this requires M SPD, which is in fact a natural condition to ensure that B_{mg} is SPD, see, e.g., [7, 20].

Now, the purpose of this work is to compare the condition numbers $\kappa(B_{a,\omega} A)$ and $\kappa(B_{m_s,\omega} A)$ associated with the additive and multiplicative variants. The theoretical analyses in [11] and [29, Section 5.7] allow to bound the convergence of the multiplicative method as a function of the condition number of the corresponding additive preconditioner. However, these analyses involve extra factors; i.e., they are not accurate enough to tell which method is better. From a practical viewpoint, it is observed in [26, 27] that the multiplicative variant has often better convergence properties, and it is suggested in [15], that the multiplicative variant has always lower condition number. However, in [19] an example is provided showing that the converse can be true. Here we solve the issue by taking into account the scaling parameter ω. It is indeed clear that multiplicative iteration matrices like $T_{m,\omega}$ or $T_{m_s,\omega}$ can be effective only if the basic preconditioner is properly scaled. In fact, we prove that the multiplicative variant has always lower condition number if $\omega\,\lambda_{\max}(M^{-1}A)$ is equal to or slightly larger than 1.

On the other hand, to our knowledge, there is so far no general bound on the condition number of the additive variant that would depend only on the condition number of the multiplicative variant. Hence one could not guarantee anything for an additive implementation using solely the convergence analysis of a method based on the multiplicative implementation (like, e.g., the analysis of the balancing Neumann-Neumann domain decomposition method [27]). In this paper, we show on the contrary that $\kappa(B_{a,\omega} A)$ is at worst a modest multiple of $\kappa(B_{m_s,\omega}A)$. Here again, this requires a proper scaling of the basic preconditioner, namely that $\omega\,\lambda_{\max}(M^{-1}A)\approx 1$. In particular, the choice $\omega=\lambda_{\max}(M^{-1}A)^{-1}$ yields $\kappa(B_{a,\omega}A)\leq 4\,\kappa(B_{m_s,\omega}A)$.

We have thus a two-sided bound on $\kappa(B_{a,\omega}A)$. Eventually, we provide examples showing that $\kappa(B_{a,\omega}A)$ may be arbitrarily close to either limit, and hence that both upper and lower bounds are sharp.

Note that, besides the additive and multiplicative implementations referred above, a coarse grid correction may also be used for *deflation* [9, 17, 30, 31, 32]. This amounts to decompose the solution of the linear systems in two components, one in the range of P and one in a complementary subspace. Since the former is easy to compute, the preconditioned system can be *deflated*; that is, the iterative solution process is run in a restricted subspace, in which it has better (effective [1]) condition number. We do not discuss this approach here because we have little to add to the extensive comparison by Nabben and Vuik with the additive and multiplicative variants [17, 18]. In fact, it is shown in [18] that the conjugate gradient method combined with either the deflation or the symmetrized multiplicative preconditioner $B_{m_s,\omega}$ produces identical iterates if the latter is used with the special starting vector $\mathbf{u}_0=P(P^TA\,P)^{-1}P^T\mathbf{b}$ mentioned above, whereas the zero vector is used as initial approximation for deflation. Moreover, the spectrum of $B_{m_s,\omega}A$ is the spectrum of the deflated system with, in addition, the eigenvalue 1. Hence both approaches are very close to each other. Regarding the additive variant, the main result in [17] is a proof that deflation always leads to lower condition number. A detailed analysis is also given for the special case where the range of P coincides with an invariant subspace of A.

The paper is organized as follows. Our general analysis is developed in §2, and is supplemented in §3 with the detailed investigation of two particular examples. Concluding remarks are given in §4.

2 Analysis

The following constant plays an important role in our analysis:

$$\mu \;=\; \lambda_{\max}(M^{-1}A)\;\cdot\;\max_{\mathbf{z}\in\Re^n\setminus\{0\}}\frac{\mathbf{z}^T M\big(I-P(P^TM\,P)^{-1}P^T M\big)\,\mathbf{z}}{\mathbf{z}^T A\,\mathbf{z}}\;. \tag{2.1}$$

Note that, equivalently, μ may be defined as the smallest number such that the following weak approximation condition holds:

$$\forall\mathbf{u}\in\Re^n\ \exists\mathbf{v}\in\Re^n\ \text{ such that }\ \|\mathbf{u}-P\,\mathbf{v}\|_M^2\;\leq\;\frac{\mu}{\lambda_{\max}(M^{-1}A)}\,\|\mathbf{v}\|_A^2\;. \tag{2.2}$$

---

[1] the deflated system has n_c zero eigenvalues, which plays however no role in the solution process

It is also worth noting that $\mu \geq 1$ since

$$\begin{aligned} \mu &= \max_{\mathbf{z}} \frac{\mathbf{z}^T A\,\mathbf{z}}{\mathbf{z}^T M\,\mathbf{z}} \cdot \max_{\mathbf{z}} \frac{\mathbf{z}^T M\big(I - P(P^T M\,P)^{-1}P^T\big)\,\mathbf{z}}{\mathbf{z}^T A\,\mathbf{z}} \\ &\geq \max_{\mathbf{z}} \frac{\mathbf{z}^T M\big(I - P(P^T M\,P)^{-1}P^T\big)\,\mathbf{z}}{\mathbf{z}^T M\,\mathbf{z}} \\ &= 1\,. \end{aligned}$$

Theorem 2.1 contains our analysis of the multiplicative scheme. The role of μ is well known in this context, and the stated results are straightforward corollaries of Theorem 2.1 in [21], which, in the symmetric case, is itself only a slight extension of the analysis in [8] (see also Section 3.2 in [29]); the extension allows to cover arbitrary positive definite matrix M whereas the original result was restricted to matrices M defined via (1.6) with $M_1 = M_2^T$.

Theorem 2.1 (Multiplicative variant) *Let A, M be $n \times n$ SPD matrices and let P be a $n \times n_c$ matrix of rank $n_c < n$. Let $B_{m_s,\omega}$ and μ be defined by* (1.4) *and* (2.1)*, respectively.*

Setting $\lambda_M = \lambda_{\max}(M^{-1}A)$, there holds

$$1 \;\leq\; \lambda_{\max}(B_{m_s,\omega}\,A) \;\leq\; \max(1\,,\,\omega\,\lambda_M)\;, \tag{2.3}$$

$$\lambda_{\min}(B_{m_s,\omega}\,A) \;=\; \min\left(1\,,\,\frac{\omega\,\lambda_M}{\mu}\right) \tag{2.4}$$

and

$$\max\left(\frac{\mu}{\omega\,\lambda_M}\,,\,1\right) \;\leq\; \kappa(B_{m_s,\omega}\,A) \;\leq\; \max\left(\frac{\mu}{\min(1,\omega\,\lambda_M)}\,,\,\omega\,\lambda_M\right)\;. \tag{2.5}$$

Moreover, the upper bound (2.5) *is minimal for any ω such that*

$$1 \;\leq\; \omega\,\lambda_M \;\leq\; \mu\;, \tag{2.6}$$

and in this case one has

$$\frac{\mu}{\omega\,\lambda_M} \;\leq\; \kappa(B_{m_s,\omega}\,A) \;\leq\; \mu\;. \tag{2.7}$$

Proof. By Theorem 2.1 in [21], $B_{m,\omega}\,A$ as defined by (1.3) has n_c times the eigenvalue 1, and the remaining $n - n_c$ eigenvalues are the inverse of the nonzero eigenvalues of $\omega^{-1}\,A^{-1}\,M\big(I - P(P^T M\,P)^{-1}P^T\,M\big)$. Since $B_{m_s,\omega}\,A$ and $B_{m,\omega}\,A$ have the same eigenvalues, this yields (2.4) and the left inequality (2.3); the right inequality (2.3) is proved in [21, Corollary 2.1]. Combining (2.4) and (2.3) gives

$$\begin{aligned} \max\left(\frac{\mu}{\omega\,\lambda_M}\,,\,1\right) \;\leq\; \kappa(B_{m_s,\omega}\,A) &\;\leq\; \max(1\,,\,\omega\lambda_M) \cdot \max\left(1\,,\,\frac{\mu}{\omega\,\lambda_M}\right) \\ &\;=\; \max\left(1\,,\,\omega\,\lambda_M\,,\,\frac{\mu}{\min(1,\omega\,\lambda_M)}\right)\;, \end{aligned}$$

hence (2.5) since $\mu \geq 1$. The last result (2.7) is straightforward to check from (2.5). ■

The next theorem contains our results for the additive scheme. It is novel in several aspect. Firstly, additive schemes are seemingly for the first time analyzed in term of the constant μ defined in (2.1) or (2.2). Next we provide both upper and lower bounds, showing that the role of this constant is as crucial as it is for the multiplicative scheme.

Theorem 2.2 (Additive variant) *Let A, M be $n\times n$ SPD matrices and let P be a $n\times n_c$ matrix of rank $n_c < n$. Let $B_{a,\omega}$ and μ defined by (1.2) and (2.1), respectively.*

Setting $\lambda_M = \lambda_{\max}(M^{-1}A)$, there holds

$$\max(1\,,\,\omega\,\lambda_M) \;\leq\; \lambda_{\max}(B_{a,\omega}\,A) \;\leq\; 1+\omega\,\lambda_M \ , \tag{2.8}$$

$$\frac{\omega\,\lambda_M}{\mu} \;\geq\; \lambda_{\min}(B_{a,\omega}\,A) \;\geq\; \frac{1}{1+\omega\,\lambda_M}\,\frac{\omega\,\lambda_M}{\mu} \tag{2.9}$$

and

$$\frac{\mu}{\min(1,\omega\,\lambda_M)} \;\leq\; \kappa(B_{a,\omega}\,A) \;\leq\; \mu\,\frac{(1+\omega\,\lambda_M)^2}{\omega\,\lambda_M}\ . \tag{2.10}$$

In particular,

$$\mu \;\leq\; \min_{\omega}\;\kappa(B_{a,\omega}\,A) \;\leq\; 4\,\mu\ . \tag{2.11}$$

Proof. $\lambda_{\max}(B_{a,\omega}\,A)$ cannot be smaller than the maximum of $\lambda_{\max}(M^{-1}A)$ and $\lambda_{\max}(P\,A_c^{-1}P^T A)$ $(=1)$, and cannot be larger than their sum, hence (2.8). To prove (2.9), let $A_c = P^T A\,P$, $M_c = P^T M\,P$ and $S_c = A_c + \omega^{-1}M_c$. One has

$$\begin{aligned}
&\left(\omega M^{-1} + P\,A_c^{-1}P^T\right)\left(\omega^{-1}M - \omega^{-2}M\,P\,S_c^{-1}P^T M\right)\\
&\qquad\qquad = \; I + \omega^{-1}P(A_c^{-1} - S_c^{-1} - \omega^{-1}A_c^{-1}M_c S_c^{-1})P^T M\\
&\qquad\qquad = \; I + \omega^{-1}P\,A_c^{-1}(S_c - A_c - \omega^{-1}M_c)S_c^{-1}P^T M\\
&\qquad\qquad = \; I\ ,
\end{aligned}$$

hence $B_{a,\omega}^{-1} = \omega^{-1}M - \omega^{-2}M\,P\,S_c^{-1}P^T M$ (which could also be concluded from the Sherman-Morrison-Woodbury formula, see, e.g, [10, p. 50]). Moreover,

$$\omega^{-1}(\omega M_c^{-1} - S_c^{-1})\,S_c\,A_c^{-1}\,M_c \;=\; \omega^{-1}(\omega M_c^{-1}S_c - I)\,A_c^{-1}\,M_c \;=\; I\ ,$$

showing that

$$\omega M_c^{-1} - S_c^{-1} \;=\; \omega\,(S_c\,A_c^{-1}\,M_c)^{-1} \;=\; \omega\,\left(M_c + \omega^{-1}M_c A_c^{-1}M_c\right)^{-1}\ .$$

Therefore, plugging into the above expression of $B_{a,\omega}^{-1}$ the expression of S_c^{-1} that can be deduced from this relation, one obtains

$$B_{a,\omega}^{-1} = \omega^{-1}\left(M - M\,P\,M_c^{-1}P^T M + M\,P\,(M_c + \omega^{-1}M_c A_c^{-1}M_c)^{-1}P^T M\right)\ .$$

Hence, for any $\mathbf{z} \in \Re^n \backslash \{0\}$,

$$\begin{aligned}
\frac{\mathbf{z}^T \omega^{-1} M\big(I - P\, M_c^{-1} P^T\big)\, \mathbf{z}}{\mathbf{z}^T A\, \mathbf{z}} &\leq \frac{\mathbf{z}^T B_{a\,,\omega}^{-1}\, \mathbf{z}}{\mathbf{z}^T A\, \mathbf{z}} \\
&\leq \frac{\mathbf{z}^T \omega^{-1} M\big(I - P\, M_c^{-1} P^T M\big)\, \mathbf{z} \;+\; \mathbf{z}^T M\, P M_c^{-1} A_c M_c^{-1} P^T M\, \mathbf{z}}{\mathbf{z}^T A\, \mathbf{z}}
\end{aligned}$$

and therefore, since $A_c = P^T A\, P$,

$$\frac{\mu}{\omega\, \lambda_M} \;\leq\; \lambda_{\max}(A^{-1} B_{a\,,\omega}^{-1}) \;\leq\; \frac{\mu}{\omega\, \lambda_M} \;+\; \|P\, M_c^{-1} P^T M\|_A^2 \;.$$

Since $\lambda_{\max}(A^{-1} B_{a\,,\omega}^{-1}) = \lambda_{\min}(B_{a\,,\omega}\, A)^{-1}$, this shows the left inequality (2.9). The right one also follows because $P\, M_c^{-1} P^T M$ is a projector, hence $\|P\, M_c^{-1} P^T M\|_A = \|I - P\, M_c^{-1} P^T M\|_A$ by Kato's Lemma (e.g., [29, Lemma 3.6]), whereas, letting $\pi = I - P\, M_c^{-1} P^T M$,

$$\begin{aligned}
\mu &= \lambda_M \;\cdot\; \max_{\mathbf{z}} \frac{\mathbf{z}^T \pi^T M\, \pi\, \mathbf{z}}{\mathbf{z}^T A\, \mathbf{z}} \\
&= \lambda_M \;\cdot\; \max_{\mathbf{z}} \frac{\mathbf{z}^T \pi^T M\, \pi\ \mathbf{z}}{\mathbf{z}^T \pi^T A \pi\ \mathbf{z}}\, \frac{\mathbf{z}^T \pi^T A \pi\ \mathbf{z}}{\mathbf{z}^T A\, \mathbf{z}} \\
&\geq \lambda_M \;\cdot\; \min_{\mathbf{z}} \frac{\mathbf{z}^T \pi^T M\, \pi\ \mathbf{z}}{\mathbf{z}^T \pi^T A \pi\ \mathbf{z}} \;\cdot\; \max_{\mathbf{z}} \frac{\mathbf{z}^T \pi^T A \pi\ \mathbf{z}}{\mathbf{z}^T A\, \mathbf{z}} \\
&\geq \|\pi\|_A^2 \;.
\end{aligned}$$

Eventually, the inequalities (2.10) are straightforward consequences of (2.8), (2.9), whereas the last result (2.7) is straightforward to check from (2.10). ■

In Theorem 2.3 below, we eventually compare both variants. This theorem directly follows from the two previous ones and highlights the main results of our analysis: on the one hand, the condition number of the additive method is at worst a modest multiple of the condition number of the multiplicative method; on the other hand, with an appropriate scaling, the multiplicative method is always faster.

Theorem 2.3 (Comparison) *Let* A , M *be* $n \times n$ *SPD matrices and let* P *be a* $n \times n_c$ *matrix of rank* $n_c < n$. *Let* $B_{a\,,\omega_a}$, $B_{m_s\,,\omega_m}$ *be given by* (1.2), (1.4), *with* $\omega = \omega_a$ *and* $\omega = \omega_m$, *respectively.*

If $\omega_m = \omega_a$, *there holds*

$$\kappa(B_{a\,,\omega_a} A) \;\leq\; \kappa(B_{m_s\,,\omega_m} A)\; (1 + \omega_a\, \lambda_M)^2 \tag{2.12}$$

and if $\omega_m\, \lambda_M \leq \mu$, *one has*

$$\kappa(B_{a\,,\omega_a} A) \;\geq\; \kappa(B_{m_s\,,\omega_m} A)\;. \tag{2.13}$$

for any $\omega_a > 0$.

Proof. The inequality (2.12) is obtained by combining the right inequality (2.10) with the left inequality (2.5), and (2.13) is obtained by combining the left inequality (2.10) with the right inequality (2.5) (the given condition on ω_m ensuring that the maximum in the right hand side of (2.5) is the first of the two terms, which is just the lower bound in (2.10)). ■

We first comment on the condition (2.6) to have minimal upper bound on the condition number of the multiplicative method. It also guarantees, see (2.13), that the condition number of this method is always better than that of the additive one. In a number of applications this condition is easy to met. With, incomplete LU factorization preconditioning [22, 23], the largest eigenvalue of the preconditioned system $\lambda_{\max}(M^{-1}A)$ is slightly larger than 1 and hence, for realistic values of μ, (2.6) holds with $\omega = 1$; i.e., no additional scaling is needed. With additive Schwarz preconditioning [24, 26, 27], a sharp upper bound on λ_M is known and hence it is not difficult to select ω such that $\omega\,\lambda_M \approx 1$. On the other hand, when applying our analysis to symmetric two-grid schemes defined from (1.5) (with $M_1 = M_2^T$), one is restricted to $\omega = 1$: it is not possible to rescale M because it is only implicitly defined from (1.6). However, for any M such that (1.6) holds with $M_1 = M_2^T$, one has $I - A^{1/2}M^{-1}A^{1/2} = (I - A^{1/2}M_1^{-1}A^{1/2})(I - A^{1/2}M_1^{-T}A^{1/2})$. Hence $I - A^{1/2}M^{-1}A^{1/2}$ is nonnegative definite, entailing that $\lambda_M \le 1$, with, in fact, near equality in most practical situations. Then, (2.6) is also (approximately) satisfied, although only $\omega = 1$ is feasible.

Now, considering also the lower bound (2.7) on $\kappa(B_{m_s,\omega}A)$, it seems advantageous to further select ω such that $\omega\,\lambda_M$ approaches μ. Then, $\kappa(B_{m_s,\omega}A)$ may indeed be significantly smaller than its upper bound. However, this remark is of little practical interest because how large is μ is often unknown. Moreover, it follows from Statement 4 of Theorem 2.1 in [21] that $\lambda_{\max}(B_{m_s,\omega}A)$ is in fact the maximum between 1 and the largest eigenvalue of $\omega\,A^{1/2}M^{-1}A^{1/2}$ projected orthogonally onto a given $n - n_c$ dimensional subspace. If $M^{-1}A$ has many eigenvalues close to the largest one, the upper bounds in (2.3), (2.5) and (2.7) are then likely very tight. Hence there is little hope to really improve the conditioning by raising ω up to the upper limit indicated in (2.6).

Regarding the additive method, the lower bound (2.10) on $\kappa(B_{a,\omega}A)$ is optimal for any $\omega \ge \lambda_M^{-1}$, whereas the upper bound is minimized when $\omega = \lambda_M^{-1}$. In the latter case, the condition number of the additive preconditioner is at most four times that of the multiplicative variant, which is itself equal to its optimal upper bound μ. As seen on an example below, the scaling factor that effectively minimizes $\kappa(B_{a,\omega}A)$ may however differ from the best theoretical value $\omega = \lambda_M^{-1}$. The additive variant is actually somehow sensitive to the scaling of the basic preconditioner, in fact more than the multiplicative variant for which only the condition (2.6) is important. This is a further advantage of the latter, besides a lower condition number, which is guaranteed as soon as $\omega\,\lambda_M \le \mu$.

3 Examples

Regarding $B_{m_s,\omega}A$, our analysis in Theorem 2.1 produces an identity for $\lambda_{\min}$ and an upper bound for $\lambda_{\max}$; the latter, according to the discussion above, is expected to be accurate in practical situations even in cases where the interval defined in (2.3) would not be tight. On the other hand, more uncertainty is left on the eigenvalues of $B_{a,\omega} A$; in particular, there is a factor of $(1+\omega\,\lambda_M)^2$ between the upper bound (2.12) and the lower bound (2.13) on the condition number. In this sections we show that our analysis is nevertheless sharp in the sense that, in some cases (as in Example 1) $\kappa(B_{a,\omega} A)$ is indeed equal to its lower bound, whereas in other cases (as in Example 2) it can be arbitrarily close to the upper bound.

The setting of the examples is kept simple to allow the derivation of meaningful expressions for the eigenvalues, but general enough to bring some perspective on realistic problems. In both cases, we consider prolongations of the form

$$P \;=\; \begin{pmatrix} 0 \\ I \end{pmatrix} \;,$$

where the identity block is of size $n_c \times n_c$. This prolongation structure resembles to the one of (generalized) hierarchical basis methods [2, 4, 33]. We further restrict ourselves to $M = I$; that is, the coarse grid correction aims to accelerate simple Richardson iterations. Note that we have then

$$M\big(I - P(P^T M\,P)^{-1}P^T M\big) \;=\; \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \;. \tag{3.1}$$

Example 1

In the first example, we consider a block diagonal system matrix

$$A \;=\; \begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix} \;.$$

This example is thus along the line of the analyzes that consider the case where the columns of P are (linear combinations of) the eigenvectors of A corresponding to the the n_c smallest eigenvalues [9, 17, 18, 22, 23]. Here we are slightly more general, and we just assume, for the sake of simplicity, that $\lambda_{\max}(A_{11}) \geq \lambda_{\max}(A_{22})$; hence

$$\lambda_M \;=\; \lambda_{\max}(A) \;=\; \lambda_{\max}(A_{11}) \;.$$

Note that with $n=2$, $n_c=1$, $\omega=1$, $\lambda_{\max}(A_{11}) = \lambda_{\min}(A_{11}) = 101$ and $\lambda_{\max}(A_{22}) = \lambda_{\min}(A_{22}) = 100$, one recovers as a particular case the example in [19] that demonstrates that additive coarse grid correction may lead to smaller condition number.

Now, since

$$\omega^{-1}\,A^{-1}\,M\big(I - P(P^T M\,P)^{-1}P^T M\big) \;=\; \begin{pmatrix} \omega^{-1}A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \;, \tag{3.2}$$

we have

$$\mu \;=\; \kappa(A_{11})\,.$$

On the other hand, one may check (by direct computation or using [21, Theorem 2.1]) that the eigenvalues of $B_{m_s\,,\omega}\,A$ are the inverse of the nonzero eigenvalues of the matrix (3.2) plus n_c times the eigenvalue 1, which yields

$$\begin{aligned}\lambda_{\max}(B_{m_s\,,\omega}\,A) &= \max(\omega\,\lambda_{\max}(A_{11})\,,\,1) \;=\; \max(\omega\,\lambda_M\,,\,1)\ ,\\ \lambda_{\min}(B_{m_s\,,\omega}\,A) &= \min(\omega\,\lambda_{\min}(A_{11})\,,\,1) \;=\; \min\left(\tfrac{\omega\,\lambda_M}{\mu}\,,\,1\right)\ .\end{aligned}$$

Hence we reproduce (2.4) whereas $\lambda_{\max}$ is equal to its upper bound (2.3). We have then

$$\kappa(B_{m_s\,,\omega}\,A) \;=\; \kappa(A_{11}) \;=\; \mu$$

if

$$\frac{1}{\lambda_{\max}(A_{11})} \;\le\; \omega \;\le\; \frac{1}{\lambda_{\min}(A_{11})}\ ,$$

which is nothing but the condition (2.6).

On the other hand, direct computations yield

$$\begin{aligned}\lambda_{\max}(B_{a\,,\omega}\,A) &= \max(\omega\,\lambda_{\max}(A_{11})\,,\,1+\omega\,\lambda_{\max}(A_{22}))\\ &= \max\left(\omega\,\lambda_M\,,\,1+\omega\,\lambda_M\,\tfrac{\lambda_{\max}(A_{22})}{\lambda_{\max}(A_{11})}\right)\ , &(3.3)\\ \lambda_{\min}(B_a A) &= \min(\omega\,\lambda_{\min}(A_{11})\,,\,1+\omega\,\lambda_{\min}(A_{22}))\\ &= \min\left(\tfrac{\omega\,\lambda_M}{\mu}\,,\,1+\omega\,\lambda_{\min}(A_{22})\right)\ . &(3.4)\end{aligned}$$

These expressions are well in agreement with (2.8), (2.9) (since $\lambda_{\max}(A_{22}) \le \lambda_{\max}(A_{11})$ by assumption, whereas the lower bound in (2.9) is always lower than 1). In particular, $\lambda_{\max}$ is equal to its lower bound (2.8) if the maximum in the right hand side of (3.3) is the first of the two terms, whereas $\lambda_{\min}$ is equal to its upper bound (2.9) if the minimum in the right hand side of (3.4) is also the first of the two terms; then one has:

$$\kappa(B_{a\,,\omega}\,A) \;=\; \kappa(A_{11}) \;=\; \mu\ . \qquad (3.5)$$

The conditions to obtain this are, on the one hand $\omega \ge (\lambda_{\max}(A_{11}) - \lambda_{\max}(A_{22}))^{-1}$ (for having the first term in the right hand side of (3.3) larger than the second one), and on the other hand either $\lambda_{\min}(A_{11}) \le \lambda_{\min}(A_{22})$ or $\omega \le (\lambda_{\min}(A_{11}) - \lambda_{\min}(A_{22}))^{-1}$ (for having the first term in the right hand side of (3.4) smaller than the second one). Considering both requirements altogether, (3.5) thus holds if either

$$\lambda_{\min}(A_{11}) - \lambda_{\min}(A_{22}) \;\le\; 0 \qquad \text{and} \qquad \frac{1}{\lambda_{\max}(A_{11}) - \lambda_{\max}(A_{22})} \;\le\; \omega$$

or

$$\begin{aligned}0 \;\le\; \lambda_{\min}(A_{11}) - \lambda_{\min}(A_{22}) \;&\le\; \lambda_{\max}(A_{11}) - \lambda_{\max}(A_{22})\\ \text{and} \qquad \frac{1}{\lambda_{\max}(A_{11}) - \lambda_{\max}(A_{22})} \;&\le\; \omega \;\le\; \frac{1}{\lambda_{\min}(A_{11}) - \lambda_{\min}(A_{22})}\ .\end{aligned}$$

Hence if $\lambda_{\min}(A_{11}) - \lambda_{\min}(A_{22}) \le \lambda_{\max}(A_{11}) - \lambda_{\max}(A_{22})$, the additive variant is as efficient as the multiplicative one for some range of the scaling parameter ω. This range can be wide, but it does not contain the value $\omega = \lambda_M^{-1}$ that optimizes the upper bound in (2.10). For this latter value,

$$\kappa(B_{a,\omega}\, A) \;=\; \kappa(A_{11}) \left(1 + \frac{\lambda_{\max}(A_{22})}{\lambda_{\max}(A_{11})}\right) .$$

On the other hand, if $\lambda_{\min}(A_{11}) - \lambda_{\min}(A_{22}) > \lambda_{\max}(A_{11}) - \lambda_{\max}(A_{22})$, one may check that $\kappa(B_{a,\omega}\, A) > \kappa(A_{11})$ for any ω.

Example 2

To analyze the next example, the following lemma is useful.

Lemma 3.1 *The matrix*

$$A \;=\; \begin{pmatrix} a\,I & C^T \\ C & b\,I \end{pmatrix}$$

has extremal eigenvalues

$$\begin{aligned} \lambda_{\max}(A) &= \frac{a+b+\sqrt{(a-b)^2+4\,\|C\|^2}}{2} , \\ \lambda_{\min}(A) &= \frac{a+b-\sqrt{(a-b)^2+4\,\|C\|^2}}{2} . \end{aligned}$$

Proof. Let n_1, n_2 be the size of the top left and bottom right blocks, respectively. Without loss of generality (both blocks play a symmetric role), we assume $n_1 > n_2$. From

$$\det\left(\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\right) \;=\; \det(B_{11})\ \det(B_{22} - B_{21}B_{11}^{-1}B_{12}) ,$$

one finds, assuming $\lambda \neq a$,

$$\begin{aligned} \det(A - \lambda\, I) &= \det\big((a-\lambda)I_{n_1}\big)\ \det\big((b-\lambda)I_{n_2} - (a-\lambda)^{-1}C\,C^T\big) \\ &= (a-\lambda)^{n_1-n_2}\ \det\big((a-\lambda)(b-\lambda)I_{n_2} - C\,C^T\big) . \end{aligned}$$

Hence λ is an eigenvalue of A that is not equal to a if and only if

$$(a-\lambda)(b-\lambda) - \nu_i \;=\; 0 ,$$

where ν_i is an eigenvalue of $C\,C^T$. Noting that all these ν_i are nonnegative, one sees that the largest λ ($\neq a$) is obtained by taking the largest of the two roots with the largest ν_i, and that the smallest λ is obtained by taking the smallest of the two roots with again the largest ν_i. Knowing that the largest eigenvalue of C^TC is $\|C\|^2$, the proof is completed by noting that a can be neither smaller than the given expression for $\lambda_{\min}(A)$ nor larger than the given expression for $\lambda_{\max}(A)$. ■

In our second example we consider

$$A \;=\; \begin{pmatrix} (1-\alpha)I & C^T \\ C & (1-\beta)I \end{pmatrix} ,$$

where C is any matrix of appropriate size and where α , β are positive parameters. We also assume A scaled in such a way that $\lambda_M = \lambda_{\max}(M^{-1}A) = \lambda_{\max}(A) = 1$. With the above lemma, one may check that this holds if and only if

$$\|C\|^2 \;=\; \alpha\,\beta \; . \tag{3.6}$$

On the other hand, with this condition, Lemma 3.1 also implies

$$\lambda_{\min}(A) \;=\; 1-\alpha-\beta \; .$$

Because the present study is restricted to symmetric positive definite matrices, we therefore assume $\alpha+\beta<1$. We also restrict ourselves to the case $n_c < n-n_c$, where n_c is the size of the bottom right block.

Now, (3.1) implies that

$$\omega^{-1}\,A^{-1}\,M\big(I-P(P^TM\,P)^{-1}P^TM\big) \;=\; \begin{pmatrix} \omega^{-1}S_A^{-1} & 0 \\ * & 0 \end{pmatrix} , \tag{3.7}$$

where $S_A = (1-\alpha)I-(1-\beta)^{-1}C^TC$ is the Schur complement of A taken with respect to its bottom right block, and where the exact expression of the block denoted by a $*$ is unimportant. It follows that, using (3.6),

$$\mu \;=\; \frac{1}{\lambda_{\min}\left((1-\alpha)I-(1-\beta)^{-1}C^TC\right)} \;=\; \frac{1}{1-\alpha-\frac{\alpha\beta}{1-\beta}} \;=\; \frac{1-\beta}{\lambda_{\min}(A)} \; .$$

Further, one may check (by direct computation or using [21, Theorem 2.1]) that the eigenvalues of $B_{m_s,\omega}\,A$ are the inverse of the nonzero eigenvalues of the matrix (3.7) plus n_c times the eigenvalue 1. Moreover, the condition $n_c < n-n_c$ implies that the smallest eigenvalue of C^TC is equal to zero, hence

$$\lambda_{\max}(B_{m_s,\omega}\,A) \;=\; \max(1\,,\,\omega(1-\alpha))$$

(in agreement with (2.3) since $\lambda_M = 1 \geq 1-\alpha$) and

$$\lambda_{\min}(B_{m_s,\omega}\,A) \;=\; \min\left(1\,,\,\frac{\omega\,\lambda_{\min}(A)}{1-\beta}\right) \;=\; \min\left(1\,,\,\frac{\omega}{\mu}\right)$$

(as expected, we reproduce (2.4)). Therefore, since $\mu(1-\alpha)>1$,

$$\kappa(B_{m_s,\omega}\,A) \;=\; \begin{cases} (1-\alpha)\mu & \text{if } \frac{1}{1-\alpha} \leq \omega\,\lambda_M \leq \mu \; , \\ \frac{\mu}{\omega\,\lambda_M} & \text{if } \omega\,\lambda_M \leq \frac{1}{1-\alpha} \; . \end{cases}$$

Here one can thus obtain a condition number smaller than μ by choosing ω close to its upper limit in (2.6). Note, however, that

$$1-\alpha \;=\; \lambda_{\min}(A)+\beta \;=\; 1-(\mu-1)\lambda_{\min}(A)\ .$$

Then, recall that we wish to consider a realistic situation. A coarse grid correction is useful if $\lambda_{\min}(A)=\kappa(A)^{-1}$ is pretty small, and further one should have $\mu \ll \lambda_{\min}(A)^{-1}$ since otherwise the coarse grid correction would not be efficient. This implies $1-\alpha\approx 1$ and the optimal condition number is not significantly smaller than μ.

On the other hand,

$$B_{a\,,\omega} \;=\; \begin{pmatrix} \omega\, I & \\ & \left(\omega+\frac{1}{1-\beta}\right) I \end{pmatrix}\ ,$$

hence, letting $\xi=\sqrt{\omega\left(\omega+\frac{1}{1-\beta}\right)}$,

$$B_{a\,,\omega}^{1/2} A\, B_{a\,,\omega}^{1/2} \;=\; \begin{pmatrix} \omega(1-\alpha) I & \xi\, C^T \\ \xi\, C & (\omega(1-\beta)+1)\, I \end{pmatrix}\ .$$

Lemma 3.1 then yields, with (3.6),

$$\begin{aligned} 2\,\lambda_{\substack{\max\\ \min}}(B_{a\,,\omega}\,A) \;&=\; \omega(2-\alpha-\beta)+1\pm\sqrt{\left(1+\omega(\alpha-\beta)\right)^2+4\,\xi^2\,\|C\|^2} \\ &=\; \omega(1+\lambda_{\min}(A))+1\pm\sqrt{\left(1+\omega(\alpha-\beta)\right)^2+4\,\omega^2\alpha\,\beta+4\,\omega\,\alpha\,\beta/(1-\beta)} \\ &=\; \omega(1+\lambda_{\min}(A))+1\pm\sqrt{\left(1-\omega(\alpha+\beta)\right)^2+4\,\omega\,\alpha(1+\beta/(1-\beta))} \\ &=\; \omega(1+\lambda_{\min}(A))+1\pm\sqrt{\left(1-\omega+\omega\,\lambda_{\min}(A)\right)^2+4\,\omega\,(1-1/\mu)}\ . \end{aligned}$$

Now, assuming again $\lambda_{\min}(A)\;\ll\;\mu^{-1}$, one obtains

$$\begin{aligned} \lambda_{\substack{\max\\ \min}}(B_{a\,,\omega}\,A) \;&\approx\; \frac{\omega+1\pm\sqrt{(1-\omega)^2+4\,\omega(1-1/\mu)}}{2} \\ &=\; \frac{\omega+1}{2}\left(1\pm\sqrt{1-\frac{4\,\omega}{\mu(\omega+1)^2}}\right)\ . \end{aligned}$$

Further, if μ, despite being much smaller than $\lambda_{\min}(A)^{-1}$, remains relatively large, one has

$$\begin{aligned} \lambda_{\max}(B_{a\,,\omega}\,A) \;&\approx\; \omega+1 \;=\; \omega\,\lambda_M+1\ , \\ \lambda_{\min}(B_{a\,,\omega}\,A) \;&\approx\; \frac{\omega}{\mu(\omega+1)} \;=\; \frac{\omega\,\lambda_M}{\mu\,(\omega\,\lambda_M+1)}. \end{aligned}$$

Hence, the situation is opposite to that in the preceding example: $\lambda_{\max}$ is close to its upper limit in (2.8) and $\lambda_{\min}$ close to its lower limit in (2.9). One has then

$$\kappa(B_{a\,,\omega}\,A) \;\approx\; \mu\,\frac{(\omega\,\lambda_M+1)^2}{\omega\,\lambda_M} \;=\; \begin{cases} \frac{\kappa(B_{m_s\,,\omega}\,A)}{(1-\alpha)\,\omega\,\lambda_M}(\omega\,\lambda_M+1)^2 & \text{if } \frac{1}{1-\alpha}\le\omega\,\lambda_M\le\mu\ , \\ \kappa(B_{m_s\,,\omega}\,A)\,(\omega\,\lambda_M+1)^2 & \text{if } \omega\,\lambda_M\le\frac{1}{1-\alpha}\ ; \end{cases}$$

that is, $\kappa(B_{a\,,\omega}\,A)$ is close to its upper bounds (2.10), (2.12) and in fact can be made arbitrarily close to them by selecting α, β such that $\lambda_{\min}(A)$ is sufficiently small and μ sufficiently large.

4 Conclusions

We have developed an analysis of preconditioning methods that combine a basic preconditioner with a coarse grid correction. This analysis emphasizes the role of proper scaling of the basic preconditioner. With the multiplicative variant, the scaling is appropriate when the largest eigenvalue associated with the basic preconditioner is equal to or slightly larger than 1. With the additive variant, the best scaling is more problem dependent, but, in the absence of additional information, the bound on the condition number is optimal when this largest eigenvalue is equal to one.

When these scaling rules are followed, our analysis proves that the condition numbers for the additive ($\kappa(B_{a\,,\omega}\,A)$) and multiplicative ($\kappa(B_{m_s\,,\omega}\,A)$) variants are related by:

$$\kappa(B_{m_s\,,\omega}\,A) \;\le\; \kappa(B_{a\,,\omega}\,A) \;\le\; 4\,\kappa(B_{m_s\,,\omega}\,A)\ .$$

Hence the condition number is always smaller with the multiplicative variant, but the improvement is at most by a modest factor. Our analysis also shows that these bounds are sharp, as, depending on the case at hand, $\kappa(B_{a\,,\omega}\,A)$ may be either equal to the lower bound or arbitrarily close to the upper bound.

References

[1] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, 1994.

[2] R. E. Bank and T. F. Dupont, *Analysis of a two-level scheme for solving finite element equations*, Tech. Rep. CNA-159, Center for Numerical Analysis, The University of Texas at Austin, Texas, USA, 1980.

[3] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, *The construction of preconditioners for elliptic problems by substructuring. I*, Math. Comp., 47 (1986), pp. 103–134.

[4] E. Chow and P. S. Vassilevski, *Multilevel block factorizations in generalized hierarchical bases*, Numer. Lin. Alg. Appl., 10 (2003), pp. 105–127.

[5] M. Dryja, *An additive schwarz algorithm for two- and three-dimensional finite element elliptic problems*, in Domain Decomposition Methods. Second International Symposium on Domain Decomposition Methods, T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, eds., Philadelphia, PA, 1989, SIAM, pp. 168–172.

[6] M. Dryja and O. B. Widlund, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, eds., Philadelphia, PA, 1990, SIAM, pp. 3–21.

[7] R. D. Falgout and P. S. Vassilevski, *On generalizing the algebraic multigrid framework*, SIAM J. Numer. Anal., 42 (2005), pp. 1669–1693.

[8] R. D. Falgout, P. S. Vassilevski, and L. T. Zikatanov, *On two-grid convergence estimates*, Numer. Lin. Alg. Appl., 12 (2005), pp. 471–494.

[9] J. Frank and C. Vuik, *On the construction of deflation-based preconditioners*, SIAM J. Sci. Comput., 23 (2001), pp. 442–462.

[10] G. H. Golub and C. F. van Loan, *Matrix Computations, 3rd ed.*, The John Hopkins University Press, Baltimore, MD, 1996.

[11] M. Griebel and P. Oswald, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math., 70 (1995), pp. 163–180.

[12] W. Hackbusch, *Multi-grid Methods and Applications*, Springer, Berlin, 1985.

[13] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436.

[14] J. Mandel, *Balancing domain decomposition*, Comm. Numer. Methods Engrg., 9 (1993), pp. 233–241.

[15] ——, *Hybrid domain decomposition with unstructured subdomains*, in Domain Decomposition Methods in Science and Engineering, J. Mandel, C. Farhat, and X.-C. Cai, eds., Providence, RI, 1994, AMS, pp. 103–112.

[16] J. Mandel and M. Brezina, *Balancing domain decomposition for problems with large jumps in coefficients*, Math. Comp., 65 (1996), pp. 1387–1401.

[17] R. Nabben and C. Vuik, *A comparison of deflation and coarse grid correction applied to porous media flow*, SIAM J. Numer. Anal., 42 (2004), pp. 1631–1647.

[18] ——, *A comparison of deflation and the balancing preconditioner*, SIAM J. Sci. Comput., 27 (2006), pp. 1742–1759.

[19] ——, *A comparison of abstract versions of deflation, balancing and additive coarse grid correction preconditioners*, Numer. Lin. Alg. Appl., 15 (2008), pp. 355–372.

[20] Y. Notay, *Algebraic multigrid and algebraic multilevel methods: a theoretical comparison*, Numer. Lin. Alg. Appl., 12 (2005), pp. 419–451.

[21] ——, *Algebraic analysis of two-grid methods: the nonsymmetric case*, Numer. Lin. Alg. Appl., 17 (2010), pp. 73–97.

[22] Y. Notay and A. Van de Velde, *Coarse-grid acceleration of parallel incomplete factorization preconditioners*, in Iterative Methods in Linear Algebra II, S. Margenov and P. Vassilevski, eds., Series in Computational and Applied Mathematics Vol.3, IMACS, 1996, pp. 106–130.

[23] A. Padiy, O. Axelsson, and B. Polman, *Generalized augmented matrix preconditioning approach and its application to iterative solution of ill-conditioned algebraic systems*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 793–818.

[24] A. Quarteroni and A. Valli, *Domain Decomposition Methods for Partial Differential Equations*, Oxford Science Publications, Oxford, 1999.

[25] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, 2003. Second ed.

[26] B. Smith, P. Bjørstad, and W. Gropp, *Domain Decomposition*, Cambridge University Press, Cambridge, 1996.

[27] A. Toselli and W. Widlund, *Domain Decomposition*, Springer Ser. Comput. Math. 34, Springer-Verlag, Berlin, 2005.

[28] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, London, 2001.

[29] P. S. Vassilevski, *Multilevel Block Factorization Preconditioners*, Springer, New York, 2008.

[30] C. Vuik, A. Segal, L. el Yaakoubi, and E. Dufour, *A comparison of various deflation vectors applied to elliptic problems with discontinuous coefficients*, Appl. Numer. Math., 41 (2002), pp. 219–233.

[31] C. Vuik, A. Segal, and J. A. Meijerink, *An efficient preconditioned cg method for the solution of a class of layered problems with extreme contrasts in the coefficients*, J. Computational Physics, 152 (1999), p. 385403.

[32] C. Vuik, A. Segal, J. A. Meijerink, and G. T. Wijma, *The construction of projection vectors for a deflated ICCG method applied to problems with extreme contrasts in the coefficients*, J. Computational Physics, 172 (2001), pp. 426–450.

[33] H. Yserentant, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.