NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
*Numer. Linear Algebra Appl.* 2014; **21**:194–220
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nla.1930

# Theoretical bounds for algebraic multigrid performance: review and analysis

Scott P. MacLachlan [1,*,†] and Luke N. Olson [2]

[1] *Department of Mathematics, Tufts University, Medford, MA 02155, USA*
[2] *Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

SUMMARY

Algebraic multigrid methods continue to grow in robustness as effective solvers for the large and sparse linear systems of equations that arise in many applications. Unlike geometric multigrid approaches, however, the theoretical analysis of algebraic multigrid is less predictive of true performance. Multigrid convergence factors naturally depend on the properties of the relaxation, interpolation, and coarse-grid correction routines used, yet without the tools of Fourier analysis, optimal and practical bounds for algebraic multigrid are not easily quantified. In this paper, we survey bounds from existing literature, with particular focus on the predictive capabilities of the theory, and provide new results relating existing bounds. We highlight the impact of these theoretical observations through several model problems and discuss the role of theoretical bounds on practical performance. Copyright © 2014 John Wiley & Sons, Ltd.

Received 18 May 2013; Revised 6 December 2013; Accepted 6 December 2013

KEY WORDS: algebraic multigrid; convergence bounds; predictive theory

## 1. INTRODUCTION

The successes of multilevel preconditioners in solving linear systems that arise from discretizations of differential equations in the physical sciences have been numerous in recent years, with both geometric and algebraic multigrid (AMG) methods being applied to a wide range of practical problems. Geometric multigrid methods, whose modern development started with the seminal paper by Brandt [1], are traditionally constructed through careful Fourier or finite-element analysis to ensure that coarse-grid correction (CGC) is accurate on error for which relaxation is least effective. Typically, performance is guaranteed through this process, with convergence factors that are bounded independently of the mesh parameters. In contrast, AMG methods attempt to form a multilevel process through matrix and graph properties alone [2–4], often motivated by heuristic arguments aimed to match the performance of geometric multigrid methods for simple (elliptic) problems. Nevertheless, AMG performance similar to that of geometric multigrid has been demonstrated on a range of problems, including those with irregular meshes or significant jumps in material parameters.

A common criticism of existing AMG theory is the apparent lack of practical benefit. By itself, establishing convergence of AMG is valuable; however, a reasonable estimate of convergence speed is much more beneficial, as it can confirm the performance of a given method or guide the AMG setup phase toward improved parameters and components. Thus, our ideal bounds form a predictive theory, giving sharp bounds on the convergence factors obtained in practice, with explicit dependence on parameters of the method. For a general AMG method and broad class of problems, achieving such results is unlikely. This realization, however, motivates our presentation and assessment of the existing AMG framework and offers directions toward a more applied view of the theory.

*Correspondence to: Scott P. MacLachlan, Department of Mathematics, Tufts University, Medford, MA 02155, USA.
†E-mail: scott.maclachlan@tufts.edu

In this paper, we review the current state of two-level theoretical analysis of various AMG approaches. One focus of our analysis is relationships between the various bounds in the literature, particularly with an eye toward sharpness of the bounds. Numerical experiments are presented to highlight the loss of sharpness that appears in commonly used bounds. An important consequence of this point of view is an understanding of how these bounds are (or can be) used to make informed choices about the components and parameters of an AMG algorithm for particular problems. We do not generally consider the important case of multilevel convergence bounds, simply because there are many fewer options for these bounds and they contribute less directly to practical AMG algorithms. A review of such bounds was recently presented in [5], with some insight into the relationship of two-level and multilevel bounds presented in [6]; an excellent general treatment of both two-level and multilevel theory appears in [7]. In general, theory for multilevel methods relies on more difficult conditions than the two-level theory proposed here, including either additional stability criteria for V-cycles [8] or more expensive cycling, as in algebraic multilevel iteration (AMLI)-type methods [9].

A convergence bound that is optimal in terms of both sharpness and computability is likely elusive given the myriad of methods that may be constructed through various combinations of relaxation methods, coarse-grid selection schemes, definitions of interpolation, and forms of CGC cycling. For any given set of choices, the resulting worst-case convergence factor per multigrid cycle may be expressed exactly, as the operator norm of the error-propagation matrix, for the norm in which convergence is measured. We note that this remains only a *bound* on the true (measured) per-cycle convergence factor for any given iteration and initial error, albeit a *sharp* bound, as it is both achieved for a particular error and the expected asymptotic convergence factor of stationary iteration for most initial errors. While this norm can be computed directly as a singular value, such computation is impractical. Conversely, the convergence factor may be easily bounded in terms of the convergence factor of the relaxation operator, but this neglects the important role of CGC. The convergence bounds considered in this paper occupy a middle ground in that they are typically based on upper bounds of the true operator norm, which are careful to account for the acceleration provided by a CGC process.

In the abstract, the focus of multigrid convergence theory is solely on bounding the two-level or multilevel convergence factors of algorithms such as AMG. For $A \in \mathbb{R}^{n\times n}$, if the bound on the error reduction in each iteration of AMG is independent of $n$, then only a fixed number of (stationary) iterations are needed to reach any fixed tolerance required of the solution, independently of problem size. In addition, if the computational complexity of each cycle is bounded by an $\mathcal{O}(n)$ number of operations, then the method is said to scale *optimally*, requiring only $\mathcal{O}(n)$ operations to reach any fixed tolerance. While fundamentally important, we do not explicitly consider complexity here, as it does not naturally enter into the convergence bounds discussed; a more natural approach is to apply these bounds to algorithms with $\mathcal{O}(n)$ per-cycle complexity and to measure reduction versus a reasonable unit of work. We also note that in the absence of such a convergence bound, optimal scalability may still be possible through the use of Krylov acceleration, but we also do not consider this here.

In Section 1.1, we introduce the AMG algorithm and the notation to be used in subsequent sections. We follow in Section 1.2 with a short introduction to AMG theory and further motivation of the goals of this paper. In Section 2, we discuss classical approaches to AMG theory, focusing first on the natural sharp bounds of two-level convergence, then on the split bounds that arise in many contexts. More recent two-level theoretical developments from [10–12] are then discussed in Section 3. The roles of compatible relaxation (CR), reduction-based AMG, and adaptive AMG principles in computing bounds on AMG performance are discussed in Section 4. Examples that highlight the various convergence bounds from these works are discussed in Section 5.

### *1.1. The algebraic multigrid algorithm*

Consider the matrix equation

$$A\mathbf{u} = \mathbf{f}, \tag{1}$$



DOI: 10.1002/nla

with $n$ so-called fine-grid points (or degrees of freedom) and where $A \in \mathbb{R}^{n\times n}$ is assumed to be symmetric and positive definite. In a purely algebraic context, we refer to a grid as a set of unknowns. Thus, the fine-grid equations in (1) are defined over a full index set, $\Omega = \{1, \ldots, n\}$, by $\sum_{j\in\Omega} a_{ij}u_j = f_i$ for each $i \in \Omega$. In this section, we introduce notation through a brief look at the two-level AMG algorithm. Fine-grid vectors and matrices are denoted with no index, while vectors and matrices on the coarse level are denoted with a subscripted $c$.

The solution phase of AMG follows the familiar geometric approach. The two-level multigrid method consists of the steps outlined in Algorithm 1.

**Algorithm 1:** AMG solution phase

| | |
|---|---|
| $\mathbf{u} \leftarrow \hat{\mathcal{G}}(\mathbf{u}, \mathbf{f})$ | {Pre-relaxation on the fine grid, $\Omega$} |
| $\mathbf{r}_c \leftarrow R\mathbf{r}$ | {Restriction of the residual to the coarse-grid, $\Omega_c$} |
| $\mathbf{e}_c \leftarrow A_c^{-1}\mathbf{r}_c$ | {Solution of the coarse-grid problem on $\Omega_c$} |
| $\hat{\mathbf{e}} \leftarrow P\mathbf{e}_c$ | {Interpolation of the CGC to the fine grid} |
| $\mathbf{u} \leftarrow \mathbf{u} + \hat{\mathbf{e}}$ | {Correction of the fine-grid solution} |
| $\mathbf{u} \leftarrow \mathcal{G}(\mathbf{u}, \mathbf{f})$ | {Post-relaxation on the fine grid} |

Throughout the paper, we assume that a disjoint splitting of the fine-level index set exists: $\Omega = F \cup C$, where $C$ represents the set of indices associated with coarse-level variables and $F$ is its complement of variables belonging only on the fine level. We thus have $\Omega_c = C$. Next, we define grid-transfer operators,

$$P : \Omega_c \to \Omega \quad \text{and} \quad R : \Omega \to \Omega_c,$$

as interpolation and restriction, respectively. Here, we assume both operators to be full rank and to be associated through the Galerkin relationship $R = P^T$. Interpolation may take any form, although we highlight some theoretical approaches that only apply when $P$ takes the 'classical' form $P = \begin{bmatrix} W_{fc} \\ I \end{bmatrix}$. Once $P$ is defined, we assume that the coarse-grid operator is always constructed through the variational principle: $A_c = P^T AP$. These choices are motivated by the assumption that $A$ is symmetric and positive definite, so that it defines both an inner product $\langle u, v\rangle_A = \langle Au, v\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product and a norm $\|u\|_A = \langle u, u\rangle_A^{\frac{1}{2}}$. As a result, the solution of (1) is equivalent to a minimization process: $u = \text{argmin}_v \langle Av, v\rangle - 2\langle f, v\rangle$. Letting $D = \text{diag}(A)$, we define the following additional notation for discrete inner products:

$$\langle \mathbf{u}, \mathbf{v}\rangle_D = \langle D\mathbf{u}, \mathbf{v}\rangle,$$
$$\langle \mathbf{u}, \mathbf{v}\rangle_{AD^{-1}A} = \langle D^{-1}A\mathbf{u}, A\mathbf{v}\rangle.$$

For each inner product, the associated norm is also defined using the same subscript. Additionally, we define $G$ as the post-relaxation error-propagation operator on the fine grid,

$$\mathbf{u} \leftarrow G\mathbf{u} + (I - G)A^{-1}\mathbf{f} \qquad \text{or} \quad \mathbf{e} \leftarrow G\mathbf{e}.$$

Also, let $\mathcal{G}$ denote the associated affine update operator for the solution so that

$$\mathbf{u} \leftarrow \mathcal{G}(\mathbf{u}, \mathbf{f}),$$

as in Algorithm 1, with similar definitions for pre-relaxation, $\hat{\mathcal{G}}$ (possibly trivial, with $\hat{G} = I$, or different than $\mathcal{G}$). In general, we write $G = I - MA$ for some matrix $M$; as an example, $G = I - \omega D^{-1}A$ gives weighted Jacobi relaxation. In what follows, we assume that both relaxation processes are norm convergent (in the $A$-norm), meaning that $\|G\|_A < 1$.

An essential part of the two-grid analysis is the CGC error-propagation operator, $T$. A single two-level correction step is defined by $\mathbf{u} \leftarrow \mathbf{u} + P\mathbf{e}_c$, where $\mathbf{e}_c$ is the error correction from the coarse



DOI: 10.1002/nla

problem. Because $\mathbf{e}_c$ is defined by $A_c\mathbf{e}_c = P^T(\mathbf{f} - A\mathbf{u})$, the error propagation in the CGC step is written as

$$\mathbf{e} \leftarrow \left(I - P\left(P^T AP\right)^{-1} P^T A\right)\mathbf{e}, \tag{2}$$

which defines the error-propagation operator, $T$. Because $A$ is symmetric and positive definite, $T = I - PA_c^{-1}P^T A$ is an orthogonal projection. This observation characterizes the error after CGC as minimized in the energy norm over the range of interpolation.

We summarize the critical assumptions made earlier, which we use uniformly throughout this paper.

*Assumption 1*
Assume $A$ is symmetric and positive-definite, $P$ is full rank, $G$ is a norm-convergent relaxation operator ($\|G\|_A < 1$), and $T = I - P(P^T AP)^{-1}P^T A$.

### *1.2. Algebraic multigrid theory*

Early efforts to develop a theoretical framework for multigrid focus on geometric-based methods—that is, methods that rely on the structure or geometry of the problem to define a coarse level discretization—and rely on techniques such as local Fourier (mode) analysis [1, 13] in the development of convergence bounds. Subsequent contributions in [14] and [15] concentrate on bounding the convergence of the multigrid method by analyzing only the performance of the relaxation scheme. As a result, observations from these works and results initiated in [16–18] were central to forming theory based on the algebraic properties of the resulting matrix problems for these geometric methods. Simultaneously, related ideas were developed in an algebraic-based multigrid context in [19] (see Section 2).

In this paper, we focus on the development of *algebraic* theory for AMG, by which, we mean approaches that are quantified by properties of the matrices involved and not by reliance on the underlying PDEs, discretization, or mesh properties. It is important to note that this is not the only approach to delineate the theory; indeed, several theoretical developments occupy a middle ground, such as the work of [8,20,21], where algebraic conditions are verified by finite-element theory. This suggests that alternative perspectives in reviewing the theoretical development may be valuable—for example, by highlighting the theory based on regularity arguments. In this paper, we concentrate specifically on the algebraic distinction set forth by the early works noted above.

The theory of [18] is further generalized in [22]. Together, [18, 19, 22] formed a basis for much of the classical theoretical analysis of AMG and contribute to much of the motivation presented in [4]. Since [4], the AMG method has evolved in a number of directions too large to fully review, including algorithmic advances in parallelism [23, 24] and application-focused methods [25–27]. Progress in generalizing the algorithm and improving convergence and robustness has largely been made in the scope of defining alternative interpolation operators [10, 28–31], of introducing different coarsening strategies [32–36], and on methods that automatically adapt the construction of the multilevel hierarchy to the problem [31, 37]. These developments have rapidly improved the viability of AMG as a commonplace solution technique, and their motivation is often deeply rooted in the underlying theory.

A notable theoretical contribution within this development is presented in [10], where an AMG method called AMGe is formulated using interpolation based on local finite-element stiffness matrices. Heuristics are developed in terms of the spectrum of the operator, $A$, to ensure acceptable convergence of the two-grid scheme. These heuristics guide both the relaxation process and the definition of interpolation. In particular, two new measures are presented, which quantify the impact of the choice of interpolation, $P$, on the CGC process. We detail the AMGe framework more in Section 3 and place these theoretical results in the context of earlier theory related to classical AMG.

The complementary nature of relaxation and CGC has long been recognized as the key to a successful AMG method. Revisiting the underlying process in [33, 34] leads to both insightful algorithmic advances and novel contributions to the nature of AMG theory. Recently, in [11, 12], the



DOI: 10.1002/nla

theoretical structure of various methods was summarized and generalized, motivating a direction for further algorithmic and theoretical development of AMG.

In this paper, we highlight the connection between several advances in the theoretical development of AMG, by reviewing the theoretical convergence bounds in the literature and by analyzing their relationships. In particular, we establish the loss of sharpness among various bounds on AMG performance and discuss the impact and computability of the various theoretical measures introduced. We also prove new relationships between these bounds, which further highlights the sharpness in the theory and motivates the different approaches. Furthermore, we consider implications of these bounds through several intuitive examples and aim to highlight directions of research that may lead toward more useful theoretical tools for practical application of AMG.

In the following, we limit our discussion of convergence to the two-level V(0,1)-cycle, because two-level V(1,0) and V(1,1) cycling yields similar theory and bounds, as shown in Lemma 4.1 of [17]. In particular, if we consider relaxation schemes whose error-propagation operators are of the form $G = I - MA$ (including, most notably, Gauss–Seidel and Jacobi relaxations), then the error-propagation operation of the V(0,1)-cycle is the $A$-adjoint of the operator for the V(1,0)-cycle with relaxation in reverse ordering, as summarized in the following theorem.

*Theorem 1 (Lemma 4.1 of* [17]*)*
Under Assumption 1, the $A$-adjoint of $GT$ is $TG^{\dagger}$, with $G^{\dagger} = I - M^T A$.

Because $GT$ and $TG^{\dagger}$ are $A$-adjoints, they have the same $A$-norm, $\|GT\|_A = \|TG^{\dagger}\|_A$. Further, because $T$ is an $A$-orthogonal projection, the symmetric V(1,1)-cycle error-propagation operator reduces to

$$(I - MA)\left(I - P\left(P^T AP\right)^{-1} P^T A\right)\left(I - M^T A\right) = GTG^{\dagger} = (GT)\left(TG^{\dagger}\right).$$

As a result, $\|GTG^{\dagger}\|_A = \|GT\|_A^2$. Thus, we focus only on bounds for the V(0,1)-cycle, $\|GT\|_A$, because bounds for V(1,0)-cycle and V(1,1)-cycle follow naturally.

## 2. ON SHARP THEORY

Early approaches to multigrid convergence theory were based on geometric properties of the continuum PDE to be solved, including knowledge of the eigenfunctions of the operator and information about the discretization [1, 13]. In [38], a new approach was proposed, specifically focusing on the algebraic nature of the matrix problem. In this section, we revisit the basics of this early theory and expose the loss of sharpness in bounding the multilevel convergence factor.

For a given multigrid hierarchy, we seek a bound on the resulting per-cycle convergence factor, estimating (exactly, if possible) the reduction in $\|\mathbf{e}\|_A$ from a two-level V(0,1)-cycle. Given the CGC operator, $T = I - P\left(P^T AP\right)^{-1} P^T A$, the $A$-orthogonal projection onto the range of interpolation, $\mathcal{R}(P)$, is given by $S = I - T = P\left(P^T AP\right)^{-1} P^T A$, showing that $\mathcal{R}(T)$ is the $A$-orthogonal complement of $\mathcal{R}(P)$. Thus, the energy norm is naturally decomposed as $\|\mathbf{e}\|_A^2 = \|(I-T)\mathbf{e}\|_A^2 + \|T\mathbf{e}\|_A^2$; this relation is useful in analyzing the multigrid error-propagation operator, because the role of CGC is explicitly represented.

The reduction in $\|\mathbf{e}\|_A$ from a two-level V(0,1)-cycle is naturally expressed by the form

$$\|GT\mathbf{e}\|_A^2 \leqslant \left(1 - \delta^*\right)\|\mathbf{e}\|_A^2.$$

Ideally, we target a *sharp* bound in the sense that

$$\|GT\|_A^2 := \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|GT\mathbf{e}\|_A^2}{\|\mathbf{e}\|_A^2} = 1 - \delta^*. \tag{3}$$

We note that the use of the supremum, rather than maximum, in (3) is unnecessary, because we are restricted to $\mathbb{R}^n$, yet, to be consistent with the literature, we use the supremum and infimum notation throughout the paper.



DOI: 10.1002/nla

In order to guarantee a parameter-independent bound on convergence, the key challenge is to identify sufficient conditions on the matrix operator, the relaxation method, and CGC (interpolation). One approach is to assume that relaxation is effective on the range of $T$—that is, there exists $\delta > 0$ such that

$$\|GT\mathbf{e}\|_A^2 \leqslant (1-\delta)\|T\mathbf{e}\|_A^2 \quad \text{for all } \mathbf{e}. \tag{4}$$

Then, because $T$ is an $A$-orthogonal projector,

$$\|GT\mathbf{e}\|_A^2 \leqslant (1-\delta)\|\mathbf{e}\|_A^2 \quad \text{for all } \mathbf{e},$$

also holds. Following [18, Lemma 2.3], we note that a similar inequality holds for $\mathbf{v} \perp \mathcal{R}(T)$,

$$\|G\mathbf{v}\|_A^2 \leqslant \|\mathbf{v}\|_A^2 \quad \text{for all } \mathbf{v} \perp \mathcal{R}(T),$$

simply by the norm-convergence of $G$. Combining these, we assume that there exists $\delta > 0$ such that

$$\|G\mathbf{v}\|_A^2 \leqslant \|\mathbf{v}\|_A^2 - \delta\|T\mathbf{v}\|_A^2 \quad \text{for all } \mathbf{v}.$$

This assumption, which asserts that relaxation effectively reduces errors remaining after CGC, establishes convergence of a two-grid V(0,1)-cycle as summarized in the following theorem.

*Theorem 2*
Under Assumption 1, if there is a $\delta > 0$ such that

$$\|G\mathbf{e}\|_A^2 \leqslant \|\mathbf{e}\|_A^2 - \delta\|T\mathbf{e}\|_A^2 \quad \text{for all } \mathbf{e}, \tag{5}$$

then

$$\|GT\|_A^2 \leqslant 1-\delta. \tag{6}$$

An equivalent form of the assumption in (5) is given in [7, 39]. This assumption yields a bound on convergence; however, the sharpness of this bound has not yet been established. To be sharp, the largest $\delta$ that satisfies (5) should equal $\delta^*$ in (3). Indeed, writing

$$\hat{\delta} = \inf_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e}\|_A^2 - \|G\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2},$$

we show that $\delta^* = \hat{\delta}$ through the following theorem, establishing the sharpness of (6).

*Theorem 3*
Under Assumption 1, if

$$\hat{\delta} = \inf_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e}\|_A^2 - \|G\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2}, \tag{7}$$

then $\hat{\delta}$ defines a sharp bound on the two-grid V(0,1)-cycle convergence factor; that is,

$$\|GT\|_A^2 = 1-\hat{\delta}. \tag{8}$$

*Proof*
Writing the $A$-orthogonal decomposition of $\mathbf{e}$ as

$$\mathbf{e} = T\mathbf{e} + (I-T)\mathbf{e},$$

gives $\|\mathbf{e}\|_A^2 = \|T\mathbf{e}\|_A^2 + \|(I-T)\mathbf{e}\|_A^2$. Noting that the supremum in the succeeding text is achieved over the constrained set with $T\mathbf{e} \neq \mathbf{0}$ (because $T\mathbf{e} = \mathbf{0}$ gives $\|GT\mathbf{e}\|_A = 0$), we have

$$\|GT\|_A^2 = \sup_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|GT\mathbf{e}\|_A^2}{\|\mathbf{e}\|_A^2} = \sup_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|GT\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2 + \|(I-T)\mathbf{e}\|_A^2}.$$



DOI: 10.1002/nla

Then, defining $\hat{\mathbf{e}} = \arg\sup_{\mathbf{e}:T\mathbf{e}\neq 0} \dfrac{\|GT\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2 + \|(I-T)\mathbf{e}\|_A^2}$, we see that $T\hat{\mathbf{e}}$ also achieves the supremum, because $T\hat{\mathbf{e}}$ also satisfies the constraint and

$$\frac{\|GTT\hat{\mathbf{e}}\|_A^2}{\|TT\hat{\mathbf{e}}\|_A^2 + \|(I-T)T\hat{\mathbf{e}}\|_A^2} = \frac{\|GT\hat{\mathbf{e}}\|_A^2}{\|T\hat{\mathbf{e}}\|_A^2} \geqslant \frac{\|GT\hat{\mathbf{e}}\|_A^2}{\|\hat{\mathbf{e}}\|_A^2}.$$

Thus, as expected, the supremum is achieved by error $\mathbf{e}$ that satisfies $(I-T)\mathbf{e} = 0$. With this,

$$\|GT\|_A^2 = \sup_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|GT\mathbf{e}\|_A^2}{\|\mathbf{e}\|_A^2} = \sup_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|G(T\mathbf{e}+(I-T)\mathbf{e})\|_A^2}{\|T\mathbf{e}\|_A^2} = \sup_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|G\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2},$$

which results in

$$1 - \|GT\|_A^2 = \inf_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|T\mathbf{e}\|_A^2 - \|G\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2} = \inf_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e}\|_A^2 - \|G\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2} = \hat{\delta}. \tag{9}$$

□

The utility of the convergence estimate in (8) is limited by its computability. Indeed, determining $\hat{\delta}$ from (7) is computationally prohibitive and motivates the development of more tractable estimates at a cost of sharpness in the bound. Correspondingly, early theory [16,40] in AMG initially considers a separated, less-sharp form of the convergence bound.

To see this, from (9), we define $\delta(\mathbf{e})$, $\alpha_g(\mathbf{e})$, and $\beta_g(\mathbf{e})$ for a given error $\mathbf{e}$ with

$$\delta(\mathbf{e}) = \underbrace{\frac{\|\mathbf{e}\|_A^2 - \|G\mathbf{e}\|_A^2}{g(\mathbf{e})}}_{\alpha_g(\mathbf{e})} \underbrace{\frac{g(\mathbf{e})}{\|T\mathbf{e}\|_A^2}}_{1/\beta_g(\mathbf{e})}, \tag{10}$$

for any non-negative function $g(\mathbf{e})$. Then, letting

$$\hat{\alpha}_g = \inf_{\mathbf{e}:g(\mathbf{e})\neq\mathbf{0}} \alpha_g(\mathbf{e}), \tag{11}$$

$$\hat{\beta}_g = \sup_{\mathbf{e}:g(\mathbf{e})\neq\mathbf{0}} \beta_g(\mathbf{e}), \tag{12}$$

we can naturally bound the convergence factor in terms of $\hat{\alpha}_g$ and $\hat{\beta}_g$ by noting that, for $\mathbf{e}$ such that $g(T\mathbf{e}) \neq 0$,

$$\begin{aligned} \|GT\mathbf{e}\|_A^2 &\leqslant \|T\mathbf{e}\|_A^2 - \hat{\alpha}_g g(T\mathbf{e}) \\ &\leqslant \|T\mathbf{e}\|_A^2 - \frac{\hat{\alpha}_g}{\hat{\beta}_g}\|T\mathbf{e}\|_A^2 = \left(1 - \frac{\hat{\alpha}_g}{\hat{\beta}_g}\right)\|T\mathbf{e}\|_A^2 \\ &\leqslant \left(1 - \frac{\hat{\alpha}_g}{\hat{\beta}_g}\right)\|\mathbf{e}\|_A^2. \end{aligned} \tag{13}$$

For $g(\mathbf{e})$ chosen such that $g(\mathbf{e}) \neq 0$ for all $\mathbf{e} \neq \mathbf{0}$, or so that $g(\mathbf{e}) \neq 0$ for all $\mathbf{e}$ such that $T\mathbf{e} \neq \mathbf{0}$, this leads to a bound on the two-grid convergence rate and, consequently, the sharp convergence bound, $\hat{\delta}$, according to

$$\|GT\|_A = \sqrt{1-\hat{\delta}} \leqslant \sqrt{1 - \frac{\hat{\alpha}_g}{\hat{\beta}_g}} \quad \text{and, equivalently,} \quad \hat{\delta} \geqslant \frac{\hat{\alpha}_g}{\hat{\beta}_g}.$$

Several natural choices of $g(\mathbf{e})$ exist. In [4, Eqn. (3.3a)], $g(\mathbf{e}) = \|\mathbf{e}\|_{AD^{-1}A}^2$ is used, although the weaker approximation property in [4, Eqn. (5.4)] takes $g(\mathbf{e}) = \|T\mathbf{e}\|_{AD^{-1}A}^2$. More recently,



DOI: 10.1002/nla

Aricó and Donatelli [41] noted that a more-general norm of $\boldsymbol{e}$ can be used as $g(\mathbf{e})$, in contrast to the choices made in [4]. They use this only in the case of multilevel circulant/Toeplitz type matrices, where the difference between $\|\mathbf{e}\|^2_{AD^{-1}A}$ and their choice of $\|\mathbf{e}\|^2_{A^2}$ is minimal, but the circulant/Toeplitz structure guides their smoothing and approximation properties. We note, however, that the freedom of choice of $g(\mathbf{e})$ clearly offers avenues for controlling the sharpness of the resulting bounds. Choosing $g(\mathbf{e}) = \|T\mathbf{e}\|^2_A$ naturally leads to $\hat{\beta}_g = 1$ and $\hat{\alpha}_g = \hat{\delta}$, which we know gives a sharp bound, with a similar result for $g(\mathbf{e}) = \|\mathbf{e}\|^2_A - \|G\mathbf{e}\|^2_A$ (discussed further in Section 3.1), while other choices are expected to lead to bounds of varying sharpness if the corresponding smoothing and approximation bounds can be shown.

Breaking down (13), we see that bounding $\hat{\alpha}_g$,

$$\bar{\alpha}_g \leqslant \hat{\alpha}_g = \inf_{\mathbf{e}: g(\mathbf{e}) \neq 0} \frac{\|\mathbf{e}\|^2_A - \|G\mathbf{e}\|^2_A}{g(\mathbf{e})}, \tag{14}$$

leads to the assumption that relaxation satisfies

$$\|G\mathbf{e}\|^2_A \leqslant \|\mathbf{e}\|^2_A - \bar{\alpha}_g g(\mathbf{e}) \quad \text{for all } \mathbf{e}, \tag{15}$$

for some $\bar{\alpha}_g$; this is the so-called smoothing assumption (terminology which first appeared in [42]). While practical, the smoothing assumption typically results in immediate loss of sharpness in the convergence estimate, even when using the optimal bound, because $\hat{\alpha}_g$ and $\hat{\beta}_g$ generally do not simultaneously achieve their extreme values.

Because the assumptions in (15) are only placed on relaxation, similar assumptions are needed on CGC in order to bound the convergence factor, $\|GT\|^2_A$. Indeed, if we assume the smoothing property in (15) with $\bar{\alpha}_g$ and consider the bounds in (12) and (13), then a bound analogous to (14) results in

$$\bar{\beta}_g \geqslant \hat{\beta}_g = \sup_{\mathbf{e}: g(\mathbf{e}) \neq 0} \frac{\|T\mathbf{e}\|^2_A}{g(\mathbf{e})}.$$

That is, there exists a $\bar{\beta}_g > 0$ such that

$$\|T\mathbf{e}\|^2_A \leqslant \bar{\beta}_g g(\mathbf{e}) \quad \text{for all } \mathbf{e}. \tag{16}$$

Because the two-level bound in (13) only depends on $g(T\mathbf{e})$, and not on $g(\mathbf{e})$ for general $\mathbf{e}$, a reasonable *weak approximation assumption* is that there exists $\bar{\beta}_g > 0$ such that

$$\|T\mathbf{e}\|^2_A \leqslant \bar{\beta}_g g(T\mathbf{e}) \quad \text{for all } \mathbf{e}. \tag{17}$$

With this, we obtain sufficient conditions for two-grid convergence, summarized by the following theorem.

*Theorem 4*
Under Assumption 1, if there exists $\bar{\alpha}_g > 0$ such that

$$\|G\mathbf{e}\|^2_A \leqslant \|\mathbf{e}\|^2_A - \bar{\alpha}_g g(\mathbf{e}) \quad \text{for all } \mathbf{e} \quad \text{(smoothing)},$$

and there exists $\bar{\beta}_g > 0$ such that

$$\|T\mathbf{e}\|^2_A \leqslant \bar{\beta}_g g(T\mathbf{e}) \quad \text{for all } \mathbf{e} \quad \text{(approximation)},$$

then $\|GT\|_A \leqslant \sqrt{1 - \bar{\alpha}_g \,/\, \bar{\beta}_g}$.

Equation (16), which establishes a stronger form of the approximation assumption in (17), is sufficient to establish convergence for multilevel V-cycles [4]. The *strong* approximation assumption in (16) is dependent on $g(\mathbf{e})$ directly, contrasting with (17), which is only defined over $\mathcal{R}(T)$. As shown in [7], the strong approximation assumption ensures $\ell_2$-boundedness of the CGC step.



DOI: 10.1002/nla

The choice in [4] of $g(\mathbf{e}) = \|\mathbf{e}\|^2_{AD^{-1}A}$ is appealing because it naturally leads to weaker but more intuitive forms of the weak and strong approximation properties. From the form of the CGC operator, $T$, as an $A$-orthogonal projection, we can rewrite $\|T\mathbf{e}\|_A = \inf_{\mathbf{e}_c} \|\mathbf{e} - P\mathbf{e}_c\|_A$. Thus, the main theoretical challenge in analyzing CGC is to define conditions on the interpolation operator that guarantee a bound of the form in (16) or (17). With this choice of $g(\mathbf{e})$ and denoting $\bar{\beta}_s = \bar{\beta}_g$, the strong approximation assumption in (16) is rewritten as

$$\inf_{\mathbf{e}_c} \|\mathbf{e} - P\mathbf{e}_c\|^2_A \leqslant \bar{\beta}_s \|\mathbf{e}\|^2_{AD^{-1}A} \quad \text{for all } \mathbf{e}. \tag{18}$$

For the weaker approximation assumption in (17), a similar statement arises in a weaker norm. Because, by definition, the range of $T$ is $A$-orthogonal to the range of $P$, for any $\mathbf{e}_c$, we have that

$$\begin{aligned} \|T\mathbf{e}\|^2_A &= \langle AT\mathbf{e}, T\mathbf{e}\rangle = \langle AT\mathbf{e}, T\mathbf{e} - P\mathbf{e}_c\rangle \\ &\leqslant \|T\mathbf{e}\|_{AD^{-1}A} \|T\mathbf{e} - P\mathbf{e}_c\|_D. \end{aligned} \tag{19}$$

Thus, if we are able to quantify the quality of interpolation with

$$\inf_{\mathbf{e}_c} \|\mathbf{e} - P\mathbf{e}_c\|^2_D \leqslant \bar{\beta}_w \|\mathbf{e}\|^2_A \quad \text{for all } \mathbf{e}, \tag{20}$$

for some $\bar{\beta}_w > 0$, then (19) shows that the approximation assumption in (17) holds. In (18) and in (20), we use subscripts to distinguish these strong and weak forms of the bound.

In this form, these approximation assumptions place natural demands on interpolation and, thus, were recognized early in the development of AMG as a critical component in establishing convergence. In [38], this restriction on interpolation was analyzed in terms of eigenvectors of $A$, which led to one of the first concise statements of the so-called Brandt–McCormick principle, namely, that $P$ should be chosen so that, for any $\lambda > 0$,

$$\sup_{\mathbf{e} \in V_{\lambda(A)}} \inf_{\mathbf{e}_c} \|\mathbf{e} - P\mathbf{e}_c\|^2_A \leqslant c\lambda^a h^s. \tag{21}$$

Here, $V_{\lambda(A)}$ is the set of all eigenvectors of $A$ with unit $A$-norm with eigenvalue no larger than $\lambda$, $a$ and $s$ are constants, and $h$ is a parameter reflecting the discretization mesh size [38, §3]. The strong approximation property in (18) naturally implies (21). To see this, let $\mathbf{e} \in V_\lambda(A)$ with $A$ denoting a standard finite-difference discretization of a differential operator. Then $\|\mathbf{e}\|^2_{AD^{-1}A} \leqslant \|D^{-1}\|\|A\mathbf{e}\|^2 \leqslant Ch^2\lambda$, where the $Ch^2$ factor is a result of the size, $1/(Ch^2)$, of the diagonal entries in $A$.

Such approximation assumptions first appeared in a similar form, also implying this choice of $g(\mathbf{e}) = \|\mathbf{e}\|^2_{AD^{-1}A}$. The original formulation of an approximation assumption was by Hackbusch as

$$\|A^{-1} - P\left(P^TAP\right)^{-1}P^T\| \leqslant ch^s; \tag{22}$$

see [42, Eqn. (3.2)], [40, Eqn. (2.7)], and [43, Eqn. (C1)]. Rewriting (18) as

$$\sup_{\mathbf{e}\neq\mathbf{0}} \inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|^2_A}{\|\mathbf{e}\|^2_{AD^{-1}A}} \leqslant \bar{\beta}_s.$$

and explicitly replacing the infimum with the $A$-orthogonal projection onto the complement of $\mathcal{R}(P)$ give

$$\sup_{\mathbf{e}\neq\mathbf{0}} \inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|^2_A}{\|\mathbf{e}\|^2_{AD^{-1}A}} = \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e} - P\left(P^TAP\right)^{-1}P^TA\mathbf{e}\|^2_A}{\|\mathbf{e}\|^2_{AD^{-1}A}}.$$

Next, we relate these norms to $\ell_2$-norms in the usual way, giving

$$\sup_{\mathbf{e}\neq\mathbf{0}} \inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|^2_A}{\|\mathbf{e}\|^2_{AD^{-1}A}} = \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|A^{1/2}\mathbf{e} - A^{1/2}P\left(P^TAP\right)^{-1}P^TA\mathbf{e}\|^2}{\|D^{-1/2}A\mathbf{e}\|^2}.$$



DOI: 10.1002/nla

Now, taking $\mathbf{u} = D^{-1/2}A\mathbf{e}$, we have

$$\begin{aligned}\sup_{\mathbf{e}\neq\mathbf{0}} \inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_A^2}{\|\mathbf{e}\|_{AD^{-1}A}^2} &= \sup_{\mathbf{u}\neq\mathbf{0}} \frac{\|A^{-1/2}D^{1/2}\mathbf{u} - A^{1/2}P\left(P^TAP\right)^{-1}P^TD^{1/2}\mathbf{u}\|^2}{\|\mathbf{u}\|^2} \\ &= \|A^{-1/2}D^{1/2} - A^{1/2}P\left(P^TAP\right)^{-1}P^TD^{1/2}\|^2 \\ &\leqslant \|A^{1/2}\|^2\|A^{-1} - P\left(P^TAP\right)^{-1}P^T\|^2\|D^{1/2}\|^2.\end{aligned}$$

Thus, Assumption (22) can imply the strong approximation assumption in (16), depending on $s$. Assumption (22) was also formulated in [16], although in a slightly more general context.

In general, the approximation assumptions as written in (17) and (16) directly relate the boundedness of $\hat{\beta}_g$ to $P$, when $g(\mathbf{e})$ is chosen appropriately. In addition, it is important to note that other forms of the approximation assumptions emerge with choices other than $g(\mathbf{e}) = \|\mathbf{e}\|_{AD^{-1}A}^2$. Indeed, the computation in (19) is replicated for any symmetric, positive-definite matrix $B$, leading to a *generalized* weak approximation assumption: this assumption is

$$\inf_{\mathbf{e}_c} \|\mathbf{e} - P\mathbf{e}_c\|_B^2 \leqslant \bar{\beta}_{w,B}\|\mathbf{e}\|_A^2 \quad \text{for all } \mathbf{e}, \tag{23}$$

compatible with the choice of $g(\mathbf{e}) = \|\mathbf{e}\|_{AB^{-1}A}^2$, with the corresponding *generalized* strong approximation assumption becoming

$$\inf_{\mathbf{e}_c} \|\mathbf{e} - P\mathbf{e}_c\|_A^2 \leqslant \bar{\beta}_{s,B}\|\mathbf{e}\|_{AB^{-1}A}^2 \quad \text{for all } \mathbf{e}.$$

These generalized assumptions, as well as their classical counterparts, have an intuitive explanation: each vector, $\mathbf{e}$, must be approximated by something in the range of interpolation, $\mathcal{R}(P)$, with accuracy proportional to $\|\mathbf{e}\|_A$. In (21), this is imposed on every eigenvector, where $\|\mathbf{e}\|_A^2$ is just the eigenvalue. In (20), this is imposed in the more general sense of [19, Thm. 4.1].

We note that, in the case where $B = A$, these generalized weak and strong approximation assumptions coincide and that the norms on both sides of the inequality are $A$-norms. Thus, in this case, $\bar{\beta} = 1$ naturally satisfies the inequality. While this may seem to be an optimal case for the theory, because the approximation assumption is trivially satisfied, the corresponding smoothing assumption from (15) is

$$\|G\mathbf{e}\|_A^2 \leqslant \|e\|_A^2 - \bar{\alpha}\|e\|_A^2 \quad \text{for all } \mathbf{e},$$

which cannot generally be satisfied for a suitable $\bar{\alpha}$ bounded away from 0. It seems natural, therefore, to consider $B$ to be an approximation to $A$ but not particularly close to $A$. We explore such a choice of $B$ in Section 3.1.

## 3. ON MEASURES

Implicit in the separated bounds—for example, (10)—is the additional assumption on smoothing that *only* error that is slow to reduce through relaxation—that is, *algebraically smooth error*—yields relatively small residuals. Thus, approximation assumption (18) ensures that these errors are adequately reduced by CGC. It is important to note that the extra assumption on smoothing does not hold for all relaxation schemes, even if they satisfy the regular smoothing assumption in (15). In particular, certain forms of distributed relaxation for Maxwell's equations are effective because they accurately address a large family of errors that yield small residuals [44]. Such relaxation techniques, however, rely upon much more specialized knowledge of the problem than we assume. For more standard AMG approaches, typical choices of relaxation include the Richardson, Jacobi, and Gauss–Seidel iterations, which often satisfy this assumption for the discrete PDE models for which AMG is effective.



DOI: 10.1002/nla

To formalize the role of relaxation in addressing algebraically smooth error, we use modified versions of the approximation assumptions discussed earlier. In particular, we define optimal weak and strong constants, $\hat{\beta}_w$ and $\hat{\beta}_s$, by using (20) and (18) to arrive at

$$\hat{\beta}_w = \sup_{\mathbf{e}\neq\mathbf{0}} \inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_D^2}{\|\mathbf{e}\|_A^2}, \tag{24}$$

and

$$\hat{\beta}_s = \sup_{\mathbf{e}\neq\mathbf{0}} \inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_A^2}{\|\mathbf{e}\|_{AD^{-1}A}^2}. \tag{25}$$

In [10], the element-based AMG (AMGe) framework is introduced. The goal of AMGe is to break up the global problem of building the AMG hierarchy based on properties of the system matrix into localized pieces that are directly optimized with low cost. At the center of this approach are the so-called AMGe measures,

$$M_1(Q,\mathbf{e}) = \frac{\|(I-Q)\mathbf{e}\|_D^2}{\|\mathbf{e}\|_A^2} \quad \text{and} \quad M_2(Q,\mathbf{e}) = \frac{\|(I-Q)\mathbf{e}\|_A^2}{\|\mathbf{e}\|_{AD^{-1}A}^2},$$

although we note that it is assumed in [10] that $A$ is pre-scaled to have unit diagonal, simplifying the $D$-norm to the standard $\ell_2$ norm and the $AD^{-1}A$ norm to the $A^2$ norm. In these measures, $Q$ is taken to be any projection onto the range of $P$, written $Q = PR$, for some $R$ such that $RP = I$, the identity on the coarse scale. By choosing $R = \left(P^T DP\right)^{-1} P^T D$ or $R = \left(P^T AP\right)^{-1} P^T A$, the projection-based bounds in (24) and (25) are respectively recovered. For more general choices of $R$, these measures are generalizations of those in (24), for measure $M_1$, and (25), for measure $M_2$, where the minimization over vectors $\mathbf{v}_c$ is replaced by the action of $R$ on $\mathbf{e}$ itself. That is, the variational principle in the strong and weak measures is replaced by an explicit projection onto the range of $P$.

Replacing the minimization with such a projection has the consequence that boundedness of the AMGe measures is a weaker condition than that of the corresponding AMG measure—that is, for any $\mathbf{e}$,

$$\inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_D^2}{\|\mathbf{e}\|_A^2} \leqslant M_1(Q,\mathbf{e}), \tag{26}$$

and

$$\inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_A^2}{\|\mathbf{e}\|_{AD^{-1}A}^2} \leqslant M_2(Q,\mathbf{e}). \tag{27}$$

As such, boundedness of $M_1$ for all $\mathbf{e}$ is sufficient to guarantee two-level convergence, and boundedness of $M_2$ for all $\mathbf{e}$ guarantees both two-level convergence (as above) and multilevel convergence (if the corresponding bound holds uniformly across all grids), simply by taking the supremum over $\mathbf{e} \neq \mathbf{0}$ on both sides of (26) and (27).

When the coarse set, $C$, is restricted to be a subset of the fine-grid degrees of freedom, further specialization is possible. In this setting, we implicitly partition the matrix $A$ as

$$A = \begin{bmatrix} A_{ff} & -A_{fc} \\ -A_{fc}^T & A_{cc} \end{bmatrix}, \tag{28}$$

with vectors partitioned to match; for example, $\mathbf{e} = \begin{pmatrix} \mathbf{e}_f \\ \mathbf{e}_c \end{pmatrix}$. In this setting, many interpolation schemes take the form $P = \begin{bmatrix} W_{fc} \\ I \end{bmatrix}$, with $W_{fc}$ denoting the coarse-to-fine interpolation weights.



DOI: 10.1002/nla

Under these conditions, a common special case of measure $M_1$ is when $R = \begin{bmatrix} 0 & I \end{bmatrix}$, giving $Q\mathbf{e} = P\mathbf{e}_c$. Thus, this choice of $R$ means that the value of

$$\hat{\tau}_w = \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_D^2}{\|\mathbf{e}\|_A^2}, \tag{29}$$

is always an upper bound on the weak AMG approximation constant, $\hat{\beta}_w$, in (24). For choices of interpolation $P$ that conforms to the classical multigrid expectation that $P = \begin{bmatrix} W_{fc} \\ I \end{bmatrix}$, the numerator of this bound then simplifies to a fine-grid only calculation. From this point of view, only upper bounds on $\hat{\tau}_w$ are relevant; small values of $\hat{\tau}_w$ guarantee reasonable AMG performance (as long as the corresponding smoothing assumption holds with a reasonable constant, $\bar{\alpha}_w$); however, large values for $\hat{\tau}_w$ do not lead to the opposite conclusion. Indeed, large values of $\hat{\tau}_w$ may result from a poorly performing multigrid algorithm but may also arise if the difference between the optimal choice of $\mathbf{v}_c$ is far from $\mathbf{e}_c$ (as is the case in the example in Section 5.1) or if there is a large gap in the constants given by the sharp theory and the bound due to the separated smoothing and approximation assumptions. Thus, the lower bound on $\hat{\tau}_w$ provided in [45] offers limited insight into AMG performance. Consequently, we do not consider these lower bounds in the following discussion.

A thorough analysis of the relationship between $\hat{\tau}_w$, AMG convergence, and other fundamental constants is given in [45]. The partitioning of $A$ in (28) highlights the block structure of the partitioned form, for which the block-diagonal dominance of $A_{ff}$ and $A_{cc}$ can be measured with the Cauchy–Bunyakowski–Schwarz (or CBS) constant [46],

$$\gamma_A = \sup_{\mathbf{v}_f\neq\mathbf{0},\mathbf{v}_c\neq\mathbf{0}} \frac{\mathbf{v}_f^T A_{fc}\mathbf{v}_c}{\left(\mathbf{v}_f^T A_{ff}\mathbf{v}_f\right)^{\frac{1}{2}}\left(\mathbf{v}_c^T A_{cc}\mathbf{v}_c\right)^{\frac{1}{2}}}. \tag{30}$$

Many characterizations of $\gamma_A$ exist, especially for matrices arising from finite-element discretizations of elliptic PDEs. One intuitive result for a symmetric and positive-definite matrix, $A$, is that the eigenvalues, $\lambda$, of the generalized eigenvalue problem,

$$\begin{bmatrix} A_{ff} & -A_{fc} \\ -A_{fc}^T & A_{cc} \end{bmatrix}\begin{pmatrix} \mathbf{x}_f \\ \mathbf{x}_c \end{pmatrix} = \lambda \begin{bmatrix} A_{ff} & 0 \\ 0 & A_{cc} \end{bmatrix}\begin{pmatrix} \mathbf{x}_f \\ \mathbf{x}_c \end{pmatrix},$$

are in the interval $[1 - \gamma_A, 1 + \gamma_A]$ with the endpoints obtained by the extremal eigenvalues [46, Corollary 9.4].

With this, the combined quality of the partitioning and choice of $P = \begin{bmatrix} W_{fc} \\ I \end{bmatrix}$ can be assessed by the CBS constant, $\gamma_{\hat{A}}$, of the transformed matrix,

$$\hat{A} = \begin{bmatrix} I & W_{fc} \\ 0 & I \end{bmatrix}^T \begin{bmatrix} A_{ff} & -A_{fc} \\ -A_{fc}^T & A_{cc} \end{bmatrix}\begin{bmatrix} I & W_{fc} \\ 0 & I \end{bmatrix}. \tag{31}$$

In this setting, $\hat{\tau}_w$ is bounded from above according to the next theorem.

*Theorem 5 (Theorem10 of* [45]*)*
Under Assumption 1, with $A$ partitioned as in (28) and with $D_{ff} = \text{diag}\left(A_{ff}\right)$, a diagonal matrix with entries matching the diagonal of $A_{ff}$, if $\gamma_{\hat{A}}$ is the CBS constant for $\hat{A}$ defined by (31) and $\hat{\tau}_w$ is defined as in (29), then

$$\hat{\tau}_w \leqslant \frac{1}{\lambda_{min}\left(D_{ff}^{-1}A_{ff}\right)}\frac{1}{1-\gamma_{\hat{A}}^2}. \tag{32}$$

The implications of this result are not immediately apparent for a typical $A$. In order to ensure a small value of $\hat{\tau}_w$, guaranteeing fast AMG convergence, two properties must hold. First, the smallest eigenvalue of $D_{ff}^{-1}A_{ff}$ must be bounded sufficiently away from zero, while, second, the CBS



DOI: 10.1002/nla

constant for $\hat{A}$ must be bounded away from one. The condition on the spectrum of $D_{ff}^{-1}A_{ff}$ depends only on the properties of the partition and is guaranteed to hold if (unweighted) Jacobi relaxation on the fine-grid submatrix, $A_{ff}$, is quick to converge. Incidentally, this is closely related to the principle of CR, discussed in Section 4. The condition on the CBS constant, $\gamma_{\hat{A}}$, is exposed by considering the block entries of $\hat{A}$,

$$\hat{A} = \begin{bmatrix} A_{ff} & A_{ff}W_{fc} - A_{fc} \\ W_{fc}^T A_{ff} - A_{fc}^T & A_{cc} - A_{fc}^T W_{fc} - W_{fc}^T A_{fc} + W_{fc}^T A_{ff} W_{fc} \end{bmatrix}. \tag{33}$$

For $\gamma_{\hat{A}}$ to be small, $\hat{A}$ must be spectrally close to its diagonal blocks. This requires that the off-diagonal $A_{ff}W_{fc} - A_{fc} \approx 0$, which suggests defining $W_{fc}$ as the 'ideal' interpolation of $A_{ff}^{-1}A_{fc}$.

Recent interest in the AMG literature has focused in many ways on the weak approximation property; however, a generalization of Theorem 5 (as originally stated in [45]) can also be derived for the strong analogue of $\hat{\tau}_w$,

$$\hat{\tau}_s = \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_A^2}{\|\mathbf{e}\|_{AD^{-1}A}^2},$$

that forms an upper bound for $\hat{\beta}_s$ as given in (25). The resulting bound is

$$\hat{\tau}_s \leqslant \frac{1}{\lambda_{\min}\left(D_{ff}^{-1}A_{ff} + A_{ff}^{-1}A_{fc}D_{cc}^{-1}A_{fc}^T\right)} \frac{1}{1 - \gamma^2_{\widehat{AD^{-1}A}}}, \tag{34}$$

where $\widehat{AD^{-1}A}$ is defined by applying the same transformation as is used in (31). There are two notable features of this bound. First, it depends on $\lambda_{\min}\left(D_{ff}^{-1}A_{ff} + A_{ff}^{-1}A_{fc}D_{cc}^{-1}A_{fc}^T\right)$ rather than $\lambda_{\min}\left(D_{ff}^{-1}A_{ff}\right)$, as in the weak bound in (32). Because $A_{ff}$ and $A_{fc}D_{cc}^{-1}A_{fc}^T$ are both symmetric positive semi-definite, we have that

$$\lambda_{\min}\left(D_{ff}^{-1}A_{ff} + A_{ff}^{-1}A_{fc}D_{cc}^{-1}A_{fc}^T\right) \geqslant \lambda_{\min}\left(D_{ff}^{-1}A_{ff}\right),$$

suggesting that the bound on $\hat{\tau}_s$ may be smaller than that on $\hat{\tau}_w$. Second, the typical 'ideal' interpolation in AMG results from minimizing $\gamma_{\hat{A}}$ as follows in the weak bound. The strong bound, however, suggests an analogous definition of an ideal *strong* interpolation operator, given by choosing $W_{fc}$ to have $\gamma_{\widehat{AD^{-1}A}} = 0$, yielding

$$W_{fc} = \left(A_{ff}D_{ff}^{-1}A_{ff} + A_{fc}D_{cc}^{-1}A_{fc}^T\right)^{-1}\left(A_{ff}D_{ff}^{-1}A_{fc} + A_{fc}D_{ff}^{-1}A_{cc}\right).$$

While the aforementioned bounds on these measures are in many ways more computable than directly bounding the approximation property or the smoothing assumption, the cost is in the form of sharpness of the bound. That is, two-level (for $M_1$) and multilevel (for $M_2$) convergence estimates based on bounding these measures are no better than those obtained by directly bounding the approximation properties from which they are derived. While the bounds on these measures may be easily expressed in terms of fine-scale generalized eigenvalue problems, these are only given as upper bounds to optimal bounds of the approximation properties. As a result, if no acceptable bound is found on these measures, the true performance of the AMG algorithm cannot be assessed.

### *3.1. Return to sharp theory*

In [11], the two-level theory for measure $M_1$ is extended to general relaxation techniques with error-propagation operators of the form $G = I - MA$. The generalized measure of [11] weights the norm



DOI: 10.1002/nla

in the numerator of $M_1$ by the matrix $X = \left(M + M^T - M^T AM\right)^{-1}$. This weight arises naturally from the symmetric smoother,

$$G^\dagger G = \left(I - M^T A\right)(I - MA) = I - \left(M + M^T - M^T AM\right)A = I - X^{-1}A.$$

Norm-convergence of $G = I - MA$ guarantees symmetry and positive definiteness of $X^{-1}$. Thus, weighting matrix $X$ represents a natural norm in which to measure the performance of relaxation, because $X^{-1}$ is the approximation to $A^{-1}$ used in the symmetrized relaxation scheme. Furthermore, the measure,

$$\mu(Q, \mathbf{e}) = \frac{\left\langle \left(M + M^T - M^T AM\right)^{-1}(I - Q)\mathbf{e}, (I - Q)\mathbf{e} \right\rangle}{\|\mathbf{e}\|_A^2} = \frac{\langle X(I - Q)\mathbf{e}, (I - Q)\mathbf{e}\rangle}{\|\mathbf{e}\|_A^2}, \quad (35)$$

for $Q = PR$, where $P$ is the AMG interpolation operator and $R$ satisfies $RP = I$, is a measure of the performance of relaxation on those components orthogonal to the range of $P$. Moreover, $\mu(Q, \mathbf{e})$ is bounded below by a generalized weak approximation assumption of the type considered in (23), namely

$$\mu(Q, \mathbf{e}) \geqslant \inf_{\mathbf{v}_c} \frac{\|\mathbf{e} - P\mathbf{v}_c\|_X^2}{\|\mathbf{e}\|_A^2}.$$

The corresponding smoothing assumption that

$$\|(I - MA)\mathbf{e}\|_A^2 \leqslant \|\mathbf{e}\|_A^2 - \alpha\|\mathbf{e}\|_{AX^{-1}A}^2,$$

holds with constant $\alpha = 1$, by the definition of the $AX^{-1}A$-norm. Thus, boundedness of $\mu(Q, \mathbf{e})$ for all $\mathbf{e} \neq \mathbf{0}$,

$$\sup_{\mathbf{e}\neq\mathbf{0}} \inf_{\mathbf{e}_c} \frac{\|\mathbf{e} - P\mathbf{e}_c\|_X^2}{\|\mathbf{e}\|_A^2} \leqslant \sup_{\mathbf{e}\neq\mathbf{0}} \mu(Q, \mathbf{e}) \leqslant K,$$

for any choice of $R$ with $Q = PR$, is sufficient to guarantee convergence of the two-level V(0, 1)-cycle whose error-propagation operator is $GT = (I - MA)T$. Consequently, Theorem 2.2 of [11] may be stated as

*Theorem 6*
Under Assumption 1, let $Q$ be any projection onto the range of $P$. If $\mu(Q, \mathbf{e}) \leqslant K$ for all $\mathbf{e} \neq \mathbf{0}$, then

$$\|GT\|_A \leqslant \left(1 - \frac{1}{K}\right)^{\frac{1}{2}}.$$

We note that the earlier argument establishing Theorem 6 essentially follows the proof of Theorem 2.2 in [11]; however, we state it in the more general setting of generalized weak approximation properties. Theorem 6.9 of [7] gives conditions under which a rigorous lower bound is also possible.

If we again consider the special case where $P = \begin{bmatrix} W_{fc} \\ I \end{bmatrix}$, with $Q = PR$ for $R = \begin{bmatrix} 0 & I \end{bmatrix}$, then the numerator of the bound in (35) simplifies to include the fine-grid submatrix, $X_{ff}$, of $X$. As noted in [45], in this case, boundedness of the measures $\mu(Q, \mathbf{e})$ and $M_1(Q, \mathbf{e})$ is equivalent in the sense that

$$\lambda_{\min}\left(D_{ff}^{-1}X_{ff}\right) \leqslant \frac{\langle X(I - Q)\mathbf{e}, (I - Q)\mathbf{e}\rangle}{\langle D(I - Q)\mathbf{e}, (I - Q)\mathbf{e}\rangle} \leqslant \lambda_{\max}\left(D_{ff}^{-1}X_{ff}\right), \quad (36)$$

where $D_{ff}$ is the diagonal of $A_{ff}$. Thus, defining $\hat{\mu} = \sup_{\mathbf{e}\neq\mathbf{0}} \mu(Q, \mathbf{e})$ leads to the inequality [45],

$$\lambda_{\min}\left(D_{ff}^{-1}X_{ff}\right)\hat{\tau}_w \leqslant \hat{\mu} \leqslant \lambda_{\max}\left(D_{ff}^{-1}X_{ff}\right)\hat{\tau}_w.$$



DOI: 10.1002/nla

Two conclusions follow directly from this observation. First, when $Q = PR$ for $R = \begin{bmatrix} 0 & I \end{bmatrix}$, boundedness of $\mu(Q, \mathbf{e})$ for all $\mathbf{e} \neq \mathbf{0}$ is again sufficient to guarantee two-level multigrid convergence, giving another point of view on the reasoning discussed earlier. Second, this bound again loses sharpness, because of the lack of sharpness in (36).

While this choice of the projection, $Q$, onto the range of $P$ leads to a bound that is several steps away from being sharp, it is interesting to note that the measure, $\mu(Q, \mathbf{e})$, can be made sharp by choosing the optimal projection, $Q_X = P(P^T XP)^{-1} P^T X$. Theorem 4.3 from [12] shows this and can be rewritten as follows.

*Theorem 7*
Under Assumption 1, let $X = (M + M^T - M^T AM)^{-1}$ and $Q_X = P(P^T XP)^{-1} P^T X$. Then, taking $K = \sup_{\mathbf{e} \neq \mathbf{0}} \mu(Q_X, \mathbf{e})$, we have

$$\|GT\|_A^2 = 1 - \frac{1}{K}. \tag{37}$$

*Remark 1*
Here, we present an alternate proof to that given in [12]. While the proof in [12] is based on a more general bound that also provides insight into the related hierarchical basis preconditioner, the proof we give here is more direct, following from the generalized split theory discussed in Section 2.

*Proof*
From (7), we have that

$$1 - \|GT\|_A^2 = \inf_{\mathbf{e}: T\mathbf{e} \neq 0} \frac{\|\mathbf{e}\|_A^2 - \|G\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2} = \inf_{\mathbf{e}: T\mathbf{e} \neq 0} \frac{\|\mathbf{e}\|_{AX^{-1}A}^2}{\|T\mathbf{e}\|_A^2}.$$

Equivalently, we can characterize convergence by

$$K = \frac{1}{1 - \|GT\|_A^2} = \sup_{\mathbf{e} \neq 0} \frac{\|T\mathbf{e}\|_A^2}{\|\mathbf{e}\|_{AX^{-1}A}^2}.$$

This immediately leads to a generalized strong approximation assumption: if

$$\|T\mathbf{e}\|_A^2 \leqslant K \|\mathbf{e}\|_{AX^{-1}A}^2 \quad \text{for all } \mathbf{e},$$

then $\|GT\|_A^2 \leqslant 1 - 1/K$. The bound is sharp if $K$ is the smallest constant for which the generalized strong approximation assumption holds.

Alternately, $\|T\mathbf{e}\|_A^2$ is bounded by using a generalized weak approximation assumption by writing

$$\begin{aligned} \|T\mathbf{e}\|_A^2 &= \left\langle X^{-1/2} AT\mathbf{e}, X^{1/2}(T\mathbf{e} - P\mathbf{e}_c) \right\rangle \\ &\leqslant \|T\mathbf{e}\|_{AX^{-1}A} \|T\mathbf{e} - P\mathbf{e}_c\|_X, \end{aligned} \tag{38}$$

for any $\mathbf{e}_c$, where the Cauchy–Schwarz inequality is sharp if

$$X^{-1/2} AT\mathbf{e} = \lambda X^{1/2}(T\mathbf{e} - P\mathbf{e}_c),$$

for some scalar $\lambda$. This is equivalent to

$$AT\mathbf{e} = \lambda X(T\mathbf{e} - P\mathbf{e}_c).$$

As a result, the generalized weak approximation assumption,

$$\inf_{\mathbf{e}_c} \|T\mathbf{e} - P\mathbf{e}_c\|_X^2 \leqslant K \|T\mathbf{e}\|_A^2 \quad \text{for all } \mathbf{e}, \tag{39}$$

guarantees that

$$\|T\mathbf{e}\|_A^2 \leqslant K \|T\mathbf{e}\|_{AX^{-1}A}^2 \quad \text{for all } \mathbf{e},$$



DOI: 10.1002/nla

from (38). This bound is sharp if the bound in (39) is sharp for some $\mathbf{e}$, and the Cauchy–Schwarz inequality in (38) is sharp for that $\mathbf{e}$. Then, writing $\mathbf{e}_c$ as a projection yields

$$\begin{aligned}\inf_{\mathbf{e}_c}\|T\mathbf{e}-P\mathbf{e}_c\|_X^2 &= \left\|\left(I-P\left(P^TXP\right)^{-1}P^TX\right)T\mathbf{e}\right\|_X^2\\ &= \left\|\left(I-P\left(P^TXP\right)^{-1}P^TX\right)\mathbf{e}\right\|_X^2\\ &= \inf_{\mathbf{e}_c}\|\mathbf{e}-P\mathbf{e}_c\|_X^2.\end{aligned}$$

As a result, choosing

$$K=\sup_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}}\inf_{\mathbf{e}_c}\frac{\|\mathbf{e}-P\mathbf{e}_c\|_X^2}{\|T\mathbf{e}\|_A^2},$$

yields a sharp bound on the generalized weak approximation assumption in (39). The maximizer of the supremum satisfies the generalized eigenvalue problem

$$T^TAT\mathbf{e}=\lambda\left(I-P\left(P^TXP\right)^{-1}P^TX\right)^TX\left(I-P\left(P^TXP\right)^{-1}P^TX\right)\mathbf{e},$$

which, using properties of projections gives

$$AT\mathbf{e}=\lambda X\left(I-P\left(P^TXP\right)^{-1}P^TX\right)\mathbf{e},$$

thus ensuring the sharpness of the Cauchy-Schwarz inequality in (38) for the $\mathbf{e}$ that gives the sharp definition of $K$.

Finally, because $\|GT\|_A^2 \leqslant 1-1/K$ from Theorem 6, and we have found an $\mathbf{e}$ such that $\|GT\mathbf{e}\|_A^2=(1-1/K)\|\mathbf{e}\|_A^2$, the equality in (37) holds. □

Theorem 7 shows that the generalized weak approximation assumption in (23) yields a sharp two-level bound (choosing $B=X$ in the notation of (23)). Because both this bound and the bound in (7) are sharp, the constant $1/K$ must be equal to $\hat{\delta}$ as defined in Theorem 3. However, a direct proof of the equality between these constants requires more work. To show this equality, we first rewrite $\hat{\delta}$ in terms of $X$ and, then, review the necessary details of [12] to directly demonstrate the equivalence.

*Lemma 1*
Under Assumption 1, constant $\hat{\delta}$ from (7) satisfies

$$\hat{\delta}=\inf_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}}\frac{\|\mathbf{e}\|_A^2-\|G\mathbf{e}\|_A^2}{\|T\mathbf{e}\|_A^2}=\inf_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}}\frac{\left\langle X^{-1}AT\mathbf{e},AT\mathbf{e}\right\rangle}{\|T\mathbf{e}\|_A^2},\tag{40}$$

where $X=\left(M+M^T-M^TAM\right)^{-1}$.

*Proof*
Direct calculation yields

$$\|GT\mathbf{e}\|_A^2=\langle AGT\mathbf{e},GT\mathbf{e}\rangle=\langle AT\mathbf{e},T\mathbf{e}\rangle-\left\langle X^{-1}AT\mathbf{e},AT\mathbf{e}\right\rangle,$$

which leads to (40). □

*Theorem 8*
Under Assumption 1, constants $\hat{\delta}$, defined in (7), and $K=\sup_{\mathbf{e}\neq\mathbf{0}}\mu(Q_X,\mathbf{e})$ for $X=\left(M+M^T-M^TAM\right)^{-1}$ and $Q_X=P\left(P^TXP\right)^{-1}P^TX$ satisfy

$$\hat{\delta}=\frac{1}{K}.$$



DOI: 10.1002/nla

*Proof*
We use (40) to prove equality with $K$ given as

$$\hat{\delta} = \inf_{\mathbf{e}:T\mathbf{e}\neq 0} \frac{\langle X^{-1}AT\mathbf{e}, AT\mathbf{e}\rangle}{\langle AT\mathbf{e}, T\mathbf{e}\rangle},$$

$$K = \sup_{\mathbf{e}\neq 0} \frac{\langle X(I-Q_X)\mathbf{e}, (I-Q_X)\mathbf{e}\rangle}{\langle A\mathbf{e}, \mathbf{e}\rangle}. \tag{41}$$

The goal of the proof is to show that the $X^{-1}$ in the numerator of the infimum-based expression for $\hat{\delta}$ can be moved to the denominator of a similar expression for $1/K$, appearing as $X$ in the numerator of (41). To do this, we first follow the proof of Theorem 4.3 in [12] to derive an equivalent expression for $K$, then make use of the properties of orthogonal projections to show that this is equal to $1/\hat{\delta}$.

Writing $\mathbf{e} = P\mathbf{v} + T\mathbf{w}$ in (41), it is apparent that the supremum is achieved when $\mathbf{v} = 0$. As a result, an equivalent expression for $K$ is

$$K = \sup_{\mathbf{e}:T\mathbf{e}\neq 0} \frac{\langle X(I-Q_X)\mathbf{e}, (I-Q_X)\mathbf{e}\rangle}{\langle AT\mathbf{e}, T\mathbf{e}\rangle}. \tag{42}$$

Now, writing $\mathbf{e} = T\mathbf{w}$, we argue that

$$K = \sup_{\mathbf{e}:\mathbf{e}=T\mathbf{w}\neq 0} \inf_{\mathbf{w}} \frac{\langle X((I-T)\mathbf{w} + T\mathbf{e}), ((I-T)\mathbf{w} + T\mathbf{e})\rangle}{\langle A\mathbf{e}, \mathbf{e}\rangle},$$

by noting that the choice of $\mathbf{w}$ only affects the value of the numerator. Thus, in order to minimize the numerator over all choices of $\mathbf{w}$, we first write $(I-T)\mathbf{w} + T\mathbf{e} = \mathbf{e} + (I-T)(\mathbf{w}-\mathbf{e})$ and then note that $(I-T)(\mathbf{w}-\mathbf{e}) \in \text{Range}(P)$. Thus, the infimum is attained when $(I-T)(\mathbf{w}-\mathbf{e}) = -Q_X\mathbf{e}$, the $X$-orthogonal projection of $\mathbf{e}$ onto the range of $P$.

Now, because $\mathbf{e} = T\mathbf{w}$, we write $\mathbf{w} = T\mathbf{e} + (I-T)\mathbf{w}$, simplifying the numerator to give

$$K = \sup_{\mathbf{e}:\mathbf{e}=T\mathbf{w}\neq 0} \inf_{\mathbf{w}} \frac{\langle X\mathbf{w}, \mathbf{w}\rangle}{\langle A\mathbf{e}, \mathbf{e}\rangle}.$$

Here, the outer supremum in this expression is over all non-zero vectors, $\mathbf{e} \in \text{Range}(T)$, while, for each such $\mathbf{e}$, the inner infimum is over all $\mathbf{w}$ such that $\mathbf{e} = T\mathbf{w}$.

This expression for $K$ is shown to imply that

$$K = \sup_{\mathbf{e}\neq 0} \frac{\langle B\mathbf{e}, \mathbf{e}\rangle}{\langle A\mathbf{e}, \mathbf{e}\rangle},$$

by Theorem 4.1 in [12], where $B$ is defined by

$$E = I - B^{-1}A = (I-MA)T\left(I - M^TA\right).$$

Because $\lambda_{\max}(E) = 1 - 1/K$, we also have $\lambda_{\max}\left(A^{\frac{1}{2}}EA^{-\frac{1}{2}}\right) = 1 - 1/K$. Defining the $\ell_2$-orthogonal projection onto the range of $A^{\frac{1}{2}}P$ as $\overline{Q}_A = A^{\frac{1}{2}}P\left(P^TAP\right)^{-1}P^TA^{\frac{1}{2}}$, we write

$$A^{\frac{1}{2}}EA^{-\frac{1}{2}} = \left(I - A^{\frac{1}{2}}MA^{\frac{1}{2}}\right)\left(I - \overline{Q}_A\right)\left(I - A^{\frac{1}{2}}M^TA^{\frac{1}{2}}\right),$$



DOI: 10.1002/nla

yielding $A^{\frac{1}{2}}EA^{-\frac{1}{2}} = YY^T$ for $Y = \left(I - A^{\frac{1}{2}}MA^{\frac{1}{2}}\right)\left(I - \overline{Q}_A\right)$. Thus,

$$\begin{aligned}1 - \frac{1}{K} &= \lambda_{\max}\left(Y^T Y\right) = \lambda_{\max}\left(\left(I - \overline{Q}_A\right)\left(I - A^{\frac{1}{2}}M^T A^{\frac{1}{2}}\right)\left(I - A^{\frac{1}{2}}MA^{\frac{1}{2}}\right)\left(I - \overline{Q}_A\right)\right)\\ &= \sup_{\mathbf{v}\neq 0} \frac{\mathbf{v}^T\left(I - \overline{Q}_A\right)\left(I - A^{\frac{1}{2}}\left(M + M^T - M^T AM\right)A^{\frac{1}{2}}\right)\left(I - \overline{Q}_A\right)\mathbf{v}}{\mathbf{v}^T\mathbf{v}}\\ &= \sup_{\mathbf{v}\neq 0} \frac{\mathbf{v}^T\left(I - \overline{Q}_A\right)\left(I - \left(I - \overline{Q}_A\right)A^{\frac{1}{2}}\left(M + M^T - M^T AM\right)A^{\frac{1}{2}}\left(I - \overline{Q}_A\right)\right)\left(I - \overline{Q}_A\right)\mathbf{v}}{\mathbf{v}^T\mathbf{v}}.\end{aligned}$$

Noting that the numerator is independent of $\overline{Q}_A\mathbf{v}$, we conclude that $\mathbf{v} \in \text{Range}\left(I - \overline{Q}_A\right)$, implying that

$$1 - \frac{1}{K} = 1 - \inf_{\mathbf{v}:\left(I-\overline{Q}_A\right)\mathbf{v}\neq 0} \frac{\left\langle\left(M + M^T - M^T AM\right)A^{\frac{1}{2}}\left(I - \overline{Q}_A\right)\mathbf{v}, A^{\frac{1}{2}}\left(I - \overline{Q}_A\right)\mathbf{v}\right\rangle}{\left\langle\left(I - \overline{Q}_A\right)\mathbf{v}, \left(I - \overline{Q}_A\right)\mathbf{v}\right\rangle}.$$

Finally, taking $\mathbf{e} = A^{-\frac{1}{2}}\mathbf{v}$ gives

$$\frac{1}{K} = \inf_{\mathbf{e}:T\mathbf{e}\neq 0} \frac{\langle X^{-1}AT\mathbf{e}, AT\mathbf{e}\rangle}{\langle AT\mathbf{e}, T\mathbf{e}\rangle} = \hat{\delta}.$$

□

## 4. THEORETICALLY MOTIVATED ALGORITHMS

Several algorithmic advances in AMG methods that are independent of the theoretical development have also been proposed recently. First discussed in [33], the method of CR attempts to measure the quality of a partition in (28) based on the potential multigrid convergence. While the CR principle only gives a measure of quality, it was quickly adapted to guide the coarse-grid selection process within AMG [34, 35]. More direct control of the CR principle was proposed in the reduction-based AMG (AMGr) framework [47], with a corresponding coarse-grid selection algorithm that guarantees a two-level convergence bound [36].

A similar framework for interpolation, adaptive (or bootstrap) multigrid, was developed simultaneously [29–31, 37, 48]. In adaptive multigrid methods, the interpolation operator is treated as a 'work in progress' and adapted to fit a collection of algebraically smooth errors. While these approaches were developed largely independently of the theory discussed earlier, we show that they can, in fact, be directly related to the bounds presented in the preceding sections. A recently introduced third family of methods based on M-matrix properties includes strong bounds on the multigrid approximation but less certain bounds on the coarsening rates [49, 50].

### *4.1. Theoretically motivated partitioning approaches*

In [33], CR is defined as 'a modified relaxation scheme which keeps the coarse-level variables invariant'. While there are several ways to enforce this (cf. [35]), the simplest form of CR is through $F$-relaxation, where relaxation is applied only to the fine-grid subproblem, $A_{ff}\mathbf{v}_f = \mathbf{r}_f$. CR was introduced with the motivation that 'a general measure for the quality of the set of coarse variables is the convergence rate of the compatible relaxation', and it is claimed that the convergence factor of CR must be bounded away from 1 in order to yield an efficient multigrid algorithm [33]. While this assertion was not substantiated in [33], later analysis (as discussed in the succeeding text) shows that if CR is fast to converge, then there is an interpolation operator that leads to fast convergence in the two-grid sense.

From the theoretical point of view, the question is whether a quickly converging CR method is sufficient to guarantee a bound on $M_1(Q, \mathbf{e})$ or $\mu(Q, \mathbf{e})$ for some choice of $P$. The bound in



DOI: 10.1002/nla

(32) indicates a connection in this respect, where we see that if $\lambda_{\min}\left(D_{ff}^{-1}A_{ff}\right)$ is small, then the bound on $\hat{\tau}_w$ can be large and, conversely, that if $\lambda_{\min}\left(D_{ff}^{-1}A_{ff}\right)$ is large, then $\hat{\tau}_w$ is bounded by a small factor times $\frac{1}{1-\gamma_{\hat{A}}^2}$. Furthermore, the connection between the spectrum of $D_{ff}^{-1}A_{ff}$ and the rate of convergence of a compatible, weighted Jacobi relaxation scheme is straightforward. If $\lambda_{\min}\left(D_{ff}^{-1}A_{ff}\right)$ is small, then the spectral radius of the weighted Jacobi error-propagation matrix, $I_{ff}-\omega D_{ff}^{-1}A_{ff}$, must be close to one. If, on the other hand, a weighted-Jacobi relaxation is quick to converge, then either $\lambda_{\min}\left(D_{ff}^{-1}A_{ff}\right)$ must be bounded away from zero or the eigenvalues of $D_{ff}^{-1}A_{ff}$ must be tightly clustered (even if they are all small).

To formalize this, consider the minimization of $\mu(Q,\mathbf{e})$ over all possible choices of $P$ for a given coarse-grid, defining

$$\mu^\star = \inf_P \sup_{\mathbf{e}\neq\mathbf{0}} \mu(PR,\mathbf{e}),$$

for fixed $R$ such that $RP = I$ for all $P$ admissible in the minimization. Then Theorem 5.1 of [11] establishes a bound on $\mu^\star$ only in terms of properties of CR based on a given relaxation error-propagation operator for the full-grid system, $G = I - MA$.

*Theorem 9 (Theorem 5.1 of* [11]*)*
Under Assumption 1, let $\sigma(M) = \frac{1}{2}\left(M^{-1}+M^{-T}\right)$ and take $\omega = \lambda_{\max}\left(\sigma(M)^{-1}A\right)$ and $\Delta \geqslant 1$ to be such that $\langle M^{-1}\mathbf{v},\mathbf{w}\rangle \leqslant \Delta\langle\sigma(M)\mathbf{v},\mathbf{v}\rangle^{\frac{1}{2}}\langle\sigma(M)\mathbf{w},\mathbf{w}\rangle^{\frac{1}{2}}$. Let $M_{ff}$ be the block of $M$ associated with $F$, as in (28), and $\rho_f = \|I - M_{ff}A_{ff}\|_{A_{ff}}$. Then

$$\mu^\star \leqslant \frac{\Delta^2}{2-\omega}\cdot\frac{1}{1-\rho_f}. \tag{43}$$

The bound on $\mu^\star$ in (43) depends on three parameters: $\Delta$, $\omega$, and $\rho_f$. Parameters $\Delta$ and $\omega$ depend only on the fine-grid relaxation process. In particular, bounding $\omega$ away from 2 to limit the size of $1/(2-\omega)$ is equivalent to assuming that the slow-to-converge modes of the symmetric smoothing operator, $I-\sigma(M)^{-1}A$, are those associated with small eigenvalues of $\sigma(M)^{-1}A$, noting that this is reliant on the approximation of $\sigma(M)$ to $A$, rather than that of $\frac{1}{2}\left(M+M^T\right)$ to $A^{-1}$. In many ways, multigrid relies on this relationship for efficiency, and as a result, many common relaxation schemes satisfy this property. Ensuring that $\Delta$ is small is again a property of $M$; for example, if $M$ is symmetric and positive definite, then $\Delta = 1$. To ensure that the term $1/(1-\rho_f)$ is not large requires that $\rho_f$ be bounded away from 1, which implies quick convergence of CR method with operator $I - M_{ff}A_{ff}$, measured in the $A_{ff}$-norm. Thus, Theorem 9 establishes the existence of an interpolation operator that leads to fast multigrid convergence if CR is quick to converge (and $\Delta$ and $\omega$ are appropriately bounded).

In addition, a stronger result [47] gives the form for an interpolation operator that guarantees good multigrid performance.

*Theorem 10 (Theorem 1 of* [47]*)*
Under Assumption 1, let $M_{ff}$ be symmetric and positive definite and satisfy $\mathbf{v}_f^T M_{ff}^{-1}\mathbf{v}_f \leqslant \mathbf{v}_f^T A_{ff}\mathbf{v}_f \leqslant \lambda_{\max}\mathbf{v}_f^T M_{ff}^{-1}\mathbf{v}_f$ for all $\mathbf{v}_f$, and assume that $\begin{bmatrix} M_{ff}^{-1} & -A_{fc} \\ -A_{fc}^T & A_{cc}\end{bmatrix}$ is positive semidefinite. Partition $A$ as in (28) and define relaxation by $G = I - \frac{2}{1+\lambda_{\max}}\begin{bmatrix} M_{ff} & 0 \\ 0 & 0\end{bmatrix}A$ and interpolation by $P = \begin{bmatrix} M_{ff}A_{fc} \\ I\end{bmatrix}$. Then,



DOI: 10.1002/nla

$$\|GT\|_A^2 \leqslant 1 - \frac{4}{(\lambda_{\max} + 1)^2}.$$

Here, the bound on the multigrid convergence factor depends only on the spectral-equivalence bound between $A_{ff}$ and $M_{ff}^{-1}$. This, in turn, determines the convergence factor of the weighted CR process, with error-propagation operator $I - \frac{2}{\lambda_{\max}+1} M_{ff} A_{ff}$. As a result, if the weighted compatible-relaxation process converges quickly, then the prescribed two-grid multigrid cycle also converges quickly.

Turning the principle of CR into practical algorithms has, in many ways, been less successful than the use of CR as a theoretical tool. Two similar approaches are proposed in [34, 35], where the fine-grid relaxation is defined as relaxation on $A_{ff}\mathbf{v}_f = \mathbf{r}_f$ with error-propagation operator $G_{ff} = I - M_{ff} A_{ff}$. A coarse grid is selected by first running relaxation on the homogeneous problem with a non-zero initial guess over the entire grid, followed by marking points where relaxation made a small pointwise change to the error as candidate points for the coarse grid. A subset of these points, typically selected as a maximal independent subset, is chosen as a candidate set of coarse-grid points, and the relaxation and selection phase is repeated on the remaining points until relaxation quickly converges.

A more direct approach for compatible weighted-Jacobi relaxation is proposed in [36], where a diagonal-dominance criterion is proposed to select a set of fine-grid points where $A_{ff}$ is guaranteed to be sufficiently diagonally dominant that weighted-Jacobi can be proven to be quick to converge on the fine-grid subproblem. In combination with the AMGr algorithm described in Theorem 10, this approach leads to efficient and scalable solvers for several PDE-based problems.

### *4.2. Adaptive and bootstrap algebraic multigrid*

The principle of 'self-correcting' (later called 'bootstrap' or 'adaptive') multigrid methods was first suggested in the original AMG papers [2, 3], where it was realized that while simple scalings of the matrix can lead to serious problems with AMG heuristics, acceptable performance can be recovered by adjusting interpolation to account for the scaling. This idea has been greatly expanded, with three independent implementations of this principle: adaptive smoothed aggregation [30, 31], adaptive AMG [29], and bootstrap AMG [37, 51, 52]. All of these methods share a common goal of building a more general AMG setup stage, thereby allowing treatment of more general problems within the AMG family of algorithms. The common approach is to develop a single vector, or set of vectors, that represents 'prototypical' errors that are slow to be reduced by relaxation; these vectors are then used in the definition of the AMG interpolation operators. The various methods differ in how they generate the vectors (sequentially or concurrently) and how the vectors are used in the AMG interpolation process (directly in [30, 31], indirectly in [29], and in a least-squares process in [37, 51, 52]).

In [29], a connection between the adaptive/bootstrap methods and classical AMG theory is studied, in terms of the use of a single vector to characterize the slow-to-converge modes of relaxation. There, it is shown that the bound on the weak approximation property in (20) is not made unduly small by the prototypical error vector used in determining interpolation. If this vector is representative of the slow-to-converge modes, then this indicates, but does not guarantee, good AMG convergence. In the bootstrap setting [37], in contrast, the use of a least-squares minimization in the (row-wise) definition of interpolation again indirectly controls the constant in (20), with optimal constants achieved over the set of prototype vectors indicating, but not guaranteeing, a good overall approximation property.

### *4.3. Graph Laplacians and M-matrices*

In the past few years, a new direction has been opened in AMG theory, on the basis of (unsmoothed) aggregation multigrid principles applied to graph Laplacians and other M-matrices [49, 50, 53]. These approaches build on two-level theory for unsmoothed aggregation, an approach that does not



DOI: 10.1002/nla

normally extend to optimal V-cycle algorithms, in combination with aggregation-based AMLI/K-cycle coarsening algorithms, where each coarse-grid equation is solved, in turn, using either stationary iteration (based on Chebyshev polynomials) or Krylov acceleration [9].

In contrast to the theory presented in Sections 2 and 3, these approaches focus on restricted classes of coarsening and interpolation operators, along with relaxation schemes that guarantee fast convergence of the underlying two-level methods. In [53], general aggregation is considered, with the focus on developing an a posteriori estimate of aggregate quality that can be used in the two-level bound; for certain PDE-based problems, sharp bounds are observed. In contrast, [50] fixes a two-level scheme based on pairwise aggregation that guarantees a good two-level scheme when considering a graph Laplacian matrix. In [49, 50], these two-level schemes are extended to multilevel approaches through the AMLI (K-cycle) methodology, employing an iterative solution scheme on each level of the multigrid hierarchy to ensure sufficient error reduction for optimal performance from the cycle as a whole.

## 5. NUMERICAL EXAMPLES

### *5.1. An illustrative example*

To summarize and give insight into the various bounds discussed earlier, we consider a model problem, the graph Laplacian of a dense graph with $n$ nodes, where the computations of these bounds are made explicit. Let $A = (n+1)I - \mathbf{1}\mathbf{1}^T$, where $\mathbf{1}$ is an $n \times 1$ vector with all entries equal to unity and $I$ is the $n \times n$ identity matrix. That is,

$$A = \begin{bmatrix} n & -1 & -1 & \cdots & -1 \\ -1 & n & -1 & \cdots & -1 \\ -1 & -1 & n & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -1 \\ -1 & -1 & \cdots & -1 & n \end{bmatrix}.$$

By construction, $A$ has eigenvector $\mathbf{1}$ with eigenvalue 1, and eigenspace $\{\mathbf{v} \perp \mathbf{1}\}$ of dimension $n-1$ with eigenvalue $n+1$. Moreover, $A$ is an M-matrix, thus allowing application of the classical AMG theory, although the classical AMG coarsening algorithms are challenged by this problem because $A$ is dense. Using the Sherman–Morrison formula, the inverse is readily expressed as $A^{-1} = \frac{1}{n+1}\left(I + \mathbf{1}\mathbf{1}^T\right)$.

A simple, yet effective multigrid algorithm for this matrix is defined using Richardson relaxation and a standard variational CGC aimed at the vector $\mathbf{1}$. Using Gerschgorin Discs to bound the spectral radius, $\rho(A)$, by $2n$, the error-propagation operator for the convergent Richardson relaxation is $G = I - \frac{1}{2n}A = \frac{n-1}{2n}I + \frac{1}{2n}\mathbf{1}\mathbf{1}^T$; in the context of writing the error-propagation operator for $A$ as $I - MA$, this corresponds to choosing $M = \frac{1}{2n}I$. Then, taking $P = \mathbf{1}$, we define the usual variational CGC step by its error-propagation operator, $T = I - P\left(P^T A P\right)^{-1} P^T A = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$.

With this method, $G$ has eigenvector $\mathbf{1}$ with eigenvalue $1 - \frac{1}{2n}$ and eigenspace $\{\mathbf{v} \perp \mathbf{1}\}$ of dimension $n-1$ with eigenvalue $\frac{n-1}{2n}$, while, by construction, $T$ has eigenvector $\mathbf{1}$ with eigenvalue 0 and eigenspace $\{\mathbf{v} \perp \mathbf{1}\}$ of dimension $n-1$ with eigenvalue 1. So, $GT$ has eigenvector $\mathbf{1}$ with eigenvalue 0 and eigenspace $\{\mathbf{v} \perp \mathbf{1}\}$ with eigenvalue $\frac{n-1}{2n}$; because $A$ and $GT$ are both symmetric and are diagonalized by the same eigenspaces,

$$\|GT\|_A = \frac{n-1}{2n} < \frac{1}{2}.$$

The sharp bound of (3), $\|GT\|_A^2 = 1 - \delta^*$, gives $\delta^* = 1 - \left(\frac{n-1}{2n}\right)^2$. Then, as stated in Theorem 3, an equivalent formulation of an exact bound is $\|GT\|_A^2 = 1 - \hat{\delta}$ as given in (7). Writing $\mathbf{e} = \alpha\mathbf{1} + \beta\mathbf{v}$



DOI: 10.1002/nla

for $\mathbf{v} \perp \mathbf{1}$ with $\|\mathbf{v}\| = 1$ results in $T\mathbf{e} = \beta\mathbf{v}$. Thus, the constraint $T\mathbf{e} \neq \mathbf{0}$ implies that $\beta \neq 0$. Computing the terms in (7) yields

$$\begin{aligned}\hat{\delta} &= \inf_{\alpha,\beta:\beta\neq 0} \frac{\alpha^2 n + \beta^2(n+1) - \left(1-\frac{1}{2n}\right)^2 \alpha^2 n - \frac{(n-1)^2(n+1)}{(2n)^2}\beta^2}{(n+1)\beta^2} \\ &= \inf_{\alpha,\beta:\beta\neq 0} \left( \frac{\left(1-\left(1-\frac{1}{2n}\right)^2\right)\alpha^2 n}{(n+1)\beta^2} + 1 - \left(\frac{n-1}{2n}\right)^2 \right).\end{aligned}$$

Because $\left(1-\frac{1}{2n}\right)^2 < 1$, the minimum is clearly attained with $\alpha = 0$, giving $\hat{\delta} = 1 - \left(\frac{n-1}{2n}\right)^2$, as expected.

A multitude of bounds are possible within the framework of split bounds, starting with the splitting in (10) and optimal constants (11) and (12). Two particularly interesting bounds occur with the optimal strong, $\hat{\beta}_s$, and weak, $\hat{\beta}_w$, approximation properties, analogous to the bounds in (18) and (20), taking $g(\mathbf{e}) = \|\mathbf{e}\|^2_{AD^{-1}A}$ for the strong bound (cf.(16)) and $g(T\mathbf{e}) = \|T\mathbf{e}\|^2_{AD^{-1}A}$ for the weak bound (cf.(17)). For this matrix, the weak bound is immediately a sharp bound, namely,

$$\hat{\beta}_w = \sup_{\mathbf{e}:T\mathbf{e}\neq\mathbf{0}} \frac{\|T\mathbf{e}\|^2_A}{\|T\mathbf{e}\|^2_{AD^{-1}A}} = \frac{n}{n+1},$$

and more generally, $\frac{\|T\mathbf{e}\|^2_A}{\|T\mathbf{e}\|^2_{AD^{-1}A}} = \frac{n}{n+1}$ for all $\mathbf{e}$ such that $T\mathbf{e} \neq \mathbf{0}$. Thus, $\hat{\alpha}_w = \frac{n}{n+1}\hat{\delta}$, and the weak bound is sharp. For the strong bound, bounding $\hat{\beta}_s$ is again straightforward:

$$\hat{\beta}_s = \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|T\mathbf{e}\|^2_A}{\|\mathbf{e}\|^2_{AD^{-1}A}} = \sup_{\alpha,\beta:\alpha^2+\beta^2\neq 0} \left( \frac{(n+1)\beta^2}{\alpha^2 + \frac{(n+1)^2}{n}\beta^2} \right) = \frac{n}{n+1},$$

because the supremum is clearly achieved when $\alpha = 0$, $\beta \neq 0$. Bounding $\hat{\alpha}_s$ is slightly more complex, because only of the need for more algebraic manipulation, with

$$\begin{aligned}\hat{\alpha}_s = \inf_{\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e}\|^2_A - \|G\mathbf{e}\|^2_A}{\|\mathbf{e}\|^2_{AD^{-1}A}} &= \inf_{\alpha,\beta:\alpha^2+\beta^2\neq 0} \left( \frac{\alpha^2 n\left(1-\left(1-\frac{2}{n}\right)^2\right) + \beta^2(n+1)\left(1-\frac{(n-1)^2}{(2n)^2}\right)}{\alpha^2 + \frac{(n+1)^2}{n}\beta^2} \right) \\ &= \min\left( n\left(1-\left(1-\frac{2}{n}\right)^2\right), \frac{n}{n+1}\left(1-\frac{(n-1)^2}{(2n)^2}\right)\right) \\ &= \frac{n}{n+1}\left(1-\frac{(n-1)^2}{(2n)^2}\right).\end{aligned}$$

Thus, $\hat{\alpha}_s/\hat{\beta}_s = \hat{\delta}$, and the strong bound is also sharp; as noted earlier, this occurs because the two constants, $\hat{\alpha}_s$ and $\hat{\beta}_s$, are extremized by the same vector. Direct calculation shows that the aforementioned values for $\hat{\beta}_w$ and $\hat{\beta}_s$ match those given by the more intuitive calculations in (24) and (25).

Moving from these bounds, which are based on explicitly accounting for the variational CGC process, to one for which the CGC is fixed a priori has a serious loss of sharpness as a consequence. The optimal weak approximation constant, $\hat{\beta}_w$, is bounded by $\hat{\tau}_w$, given in (29). In this setting, $\mathbf{e}_c$ is a single variable; without loss of generality, we take it to be the last entry in $\mathbf{e}$, $e_n$. With $\mathbf{e} = \alpha\mathbf{1} + \beta\mathbf{v}$ for unit vector $\mathbf{v}$ such that $\mathbf{v} \perp \mathbf{1}$, this gives $e_n = \alpha + \beta v_n$, with $v_n^2 \leqslant (n-1)/n$. Thus,

$$\begin{aligned}\hat{\tau}_w = \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e} - P\mathbf{e}_c\|^2_D}{\|\mathbf{e}\|^2_A} &= \sup_{\alpha,\beta:\alpha^2+\beta^2\neq 0} \frac{n\left((\alpha-e_n)^2 n + \beta^2\right)}{\alpha^2 n + \beta^2(n+1)} \\ &= \sup_{\beta\neq 0} \frac{n\left((\beta v_n)^2 n + \beta^2\right)}{\beta^2(n+1)} = \sup_{v_n} \frac{n\left(v_n^2 n + 1\right)}{n+1} = \frac{n^2}{n+1}.\end{aligned}$$



DOI: 10.1002/nla

We note that the value of $\hat{\tau}_w$ is larger than that of $\hat{\beta}_w$ by a factor of $n$, yielding a significant loss of sharpness.

The final bound of $\hat{\tau}_w$ $\left(\text{and, consequently, of } \hat{\beta}_w\right)$ is given in Theorem 5. In the current setting, the matrix $A$ partitions easily with $A_{ff} = (n+1)I - \mathbf{1}_{n-1}\mathbf{1}_{n-1}^T$, where $\mathbf{1}_{n-1}$ is the vector of all ones of length $n-1$, $A_{fc} = \mathbf{1}_{n-1}$, and $A_{cc} = n$, a scalar value. Using $P = \mathbf{1}$ gives $W_{fc} = \mathbf{1}_{n-1}$, while $D_{ff} = nI$. Thus, $\lambda_{\min}\left(D_{ff}^{-1}A_{ff}\right) = 2/n$ (the eigenvalue associated with eigenvector $\mathbf{1}_{n-1}$). To compute $\gamma_{\hat{A}}$, we use the definition of the CBS constant in (30) and that of $\hat{A}$ in (33); writing $\mathbf{v}_f = \alpha\mathbf{1}_{n-1} + \beta\mathbf{v}_{n-1}$ for $\mathbf{v}_{n-1} \perp \mathbf{1}_{n-1}$, $\|\mathbf{v}_{n-1}\| = 1$, and $\mathbf{v}_c = \omega$ (a scalar), this gives

$$\begin{aligned}\gamma_{\hat{A}} &= \sup_{\mathbf{v}_f\neq\mathbf{0},\mathbf{v}_c\neq\mathbf{0}} \frac{\mathbf{v}_f^T\hat{A}_{fc}\mathbf{v}_c}{\left(\mathbf{v}_f\hat{A}_{ff}\mathbf{v}_f\right)^{\frac{1}{2}}\left(\mathbf{v}_c\hat{A}_{cc}\mathbf{v}_c\right)^{\frac{1}{2}}} \\ &= \sup_{\alpha,\beta:\alpha^2+\beta^2\neq 0} \frac{(n-1)\alpha}{\sqrt{n}(2\alpha^2(n-1)+\beta^2(n+1))^{1/2}} = \sqrt{\frac{n-1}{2n}}.\end{aligned}$$

Thus, we arrive at the bound

$$\hat{\tau}_w \leqslant \frac{n}{2}\frac{1}{1-\frac{n-1}{2n}} = \frac{n^2}{n+1}.$$

We note that this is, in fact, a sharp bound for the true value of $\hat{\tau}_w$ found earlier. In this example, the loss of sharpness in the weak bound occurs from the transition from $\hat{\beta}_w$ to $\hat{\tau}_w$; further approximation of the bound using the CBS constant does not make the value worse.

Similar calculations can be performed for the bounds of the strong approximation constant, $\hat{\beta}_s$. Thus, we first find the non-variational bound, $\hat{\tau}_s$, as

$$\begin{aligned}\hat{\tau}_s = \sup_{\mathbf{e}\neq\mathbf{0}} \frac{\|\mathbf{e}-P\mathbf{e}_c\|_A^2}{\|\mathbf{e}\|_{AD^{-1}A}^2} &= \sup_{\alpha,\beta:\alpha^2+\beta^2\neq 0} \frac{(\alpha-e_n)^2 n+\beta^2(n+1)}{\alpha^2+\beta^2\frac{(n+1)^2}{n}} \\ &= \sup_{v_n} \frac{v_n^2 n+(n+1)}{\frac{(n+1)^2}{n}} = \frac{2n^2}{(n+1)^2}.\end{aligned}$$

Thus, we see that $\hat{\tau}_s$ is a relatively sharp bound for $\hat{\beta}_s$, even though $\hat{\tau}_w$ is a poor bound for $\hat{\beta}_w$. Extending this to the CBS-based bound in (34), we find that

$$\lambda_{\min}\left(D_{ff}^{-1}A_{ff} + A_{ff}^{-1}A_{fc}D_{cc}^{-1}A_{fc}^T\right) = \frac{n+3}{2n},$$

so that this term is bounded independently of $n$, in contrast to the weak case mentioned earlier, and that

$$\gamma_{\widehat{AD^{-1}A}} = \sqrt{\frac{n-1}{n(n+3)}}.$$

These give the bound

$$\hat{\tau}_s \leqslant \frac{2n}{n+3}\frac{n(n+3)}{(n+1)^2} = \frac{2n^2}{(n+1)^2}.$$

We note that this gives a sharp bound for the true value of $\hat{\tau}_s$, just as was the case with the weak bounds. Again, the only loss of sharpness in this bound is that between $\hat{\tau}_s$ and $\hat{\beta}_s$.



DOI: 10.1002/nla

### 5.2. *Anisotropy and high-order approximations*

Two well-known challenges to basic AMG methods and, consequently, interesting test cases for the bounds discussed earlier are anisotropic diffusion problems and the use of high-order finite-element methods [54]. Therefore, we investigate the theoretical convergence bounds in these settings in order to further highlight the differences in sharpness in the bounds.

Consider the model diffusion problem

$$-\nabla \cdot \kappa \nabla u = f \quad \text{in } \Omega = [0,1]^2,$$

with homogeneous Dirichlet boundary conditions. Here, we take $\kappa = \Theta K \Theta^T$, where $K = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix}$ and $\Theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, thus describing rotated, anisotropic diffusion.

We consider a finite-element discretization of the problem using triangular, $p$-order Lagrange elements, and we focus on two numerical experiments:

Varying $p$: an unstructured triangulation with $\varepsilon = 1$ (isotropic) is constructed. The finite element order is varied from $p = 1, \ldots, 4$.

Varying $\varepsilon$: a structured triangulation with $p = 1$ (linear elements) is constructed. The anisotropy is rotated by $\theta = \pi/4$, and the strength is varied from $\varepsilon = 1.0, \ldots, 0.01$.

We study the *weak* convergence bounds in the previous sections (and, so, omit the subscripted $w$ in the constants that follow). In particular, we consider the sharp estimate on convergence, $1 - \delta^*$, from (3), the estimate $1 - \frac{\hat{\alpha}}{\hat{\beta}}$ based on the separated assumptions from (10), the AMGe-based bound, $1 - \hat{\alpha}/\hat{\tau}$, from (29), and the CBS bound, $1 - \hat{\alpha}/\bar{\tau}$, where $\bar{\tau}$ is the bound on $\hat{\tau}_w$ in (32). We compare these theoretical convergence bounds on $\|GT\|_A^2$ with the observed convergence of

$$\rho^2 = \lim_{k \to \infty} \frac{\|\mathbf{e}_{k+1}\|_A^2}{\|\mathbf{e}_k\|_A^2}.$$

In practice, we measure $\|\mathbf{e}_{k+1}\|_A^2 / \|\mathbf{e}_k\|_A^2$ for some finite $k$ at which asymptotic convergence is observed. In the tests below, $k = 100$ is used, although asymptotic behavior is observed much earlier. Moreover, because convergence bounds are often deceiving as the bound approaches 1.0, we also report the expected number of iterations to reduce the initial error by 10 orders of magnitude, namely,

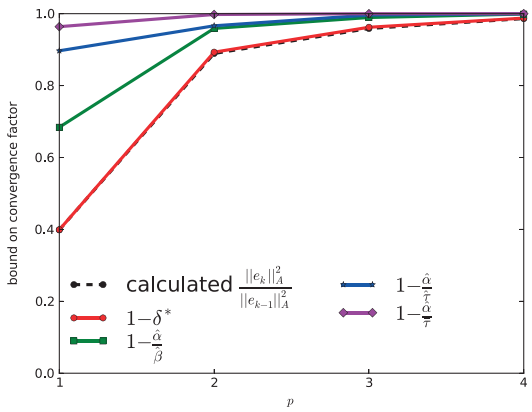
$$\text{iterations} \approx \frac{10}{\log_{10} \rho_{\text{bound}}},$$

where $\rho_{\text{bound}}$ represents the convergence bound on $\|GT\|_A$. In addition, we use the classical strength measure with a threshold of 0.0, along with the standard splitting and interpolation defined in [4].

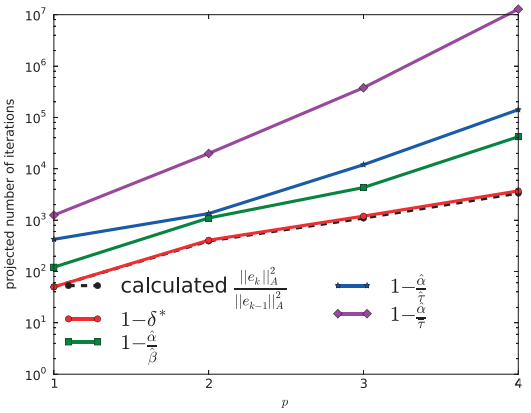
From Figures 1 and 2, the sharpness of the convergence bounds is exposed. The sharp bound, $1 - \delta^*$, matches the observed convergence closely (the separation is only due to the non-symmetry of $GT$ and, hence, the difference between the operator norm and spectral radius). Likewise, by separating the convergence assumptions into the weak approximation assumption and the weak smoothing assumption, the resulting bound, $1 - \hat{\alpha}/\hat{\beta}$, is less accurate. This is expected, because both the approximation assumption and the smoothing assumption in the case of high-order elements and for anisotropy are weak using the standard smoother (weighted Jacobi) and interpolation (Ruge-Stüben). Moreover, as $\varepsilon \to 0$ and $p$ increases, thus losing the natural elliptic character on which the basic interpolation and coarsening decisions are based, the gap between these bounds increases. That is, the sharpness of the separated bounds deteriorates as the problem becomes less elliptic, or as $A$ becomes less like an M-matrix. Finally, we see that the bounds based on $\hat{\tau}$ are highly inaccurate in each case, and again, as the problem increases in complexity, the sharpness also diminishes. Likewise, in the bounds based on $\bar{\tau}$ in (32), the $\lambda_{\min}\left(D_{ff}^{-1} A_{ff}\right)$ term is bounded far away from 0, while the CBS constant plays a more dominant role in the bound, which is magnified as the problem increases in complexity.



DOI: 10.1002/nla

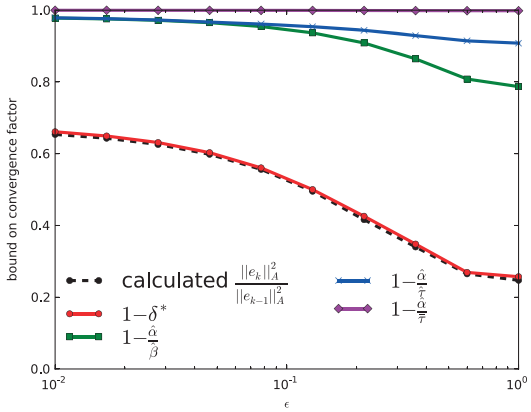



(a) Varying $p$. Convergence.

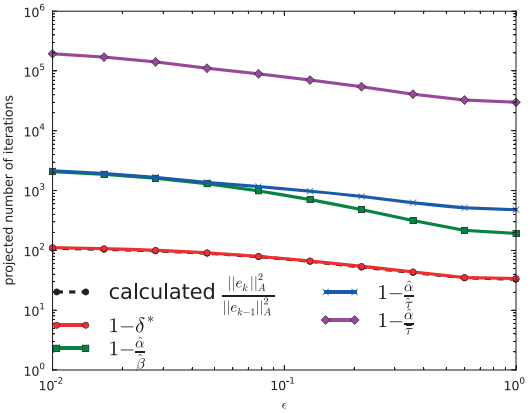



(b) Varying $p$. Iterations.

Figure 1. Convergence and effective iterations for various convergence bounds.




(a) Varying $\varepsilon$. Convergence.




(b) Varying $\varepsilon$. Iterations.

Figure 2. Convergence and effective iterations for various convergence bounds.

## 6. CONCLUSIONS

Convergence bounds play an important role in AMG methods. Bounds are not only useful in practice as a diagnostic tool, but they also serve as a critical element in the design of new methods. This paper presented a unified approach to understanding theoretical convergence bounds, from the early theory as geometric methods transformed toward algebraic, to more recent developments in areas such as element-based AMG (AMGe) and CR. Central to the theoretical analysis and method development over the years, and as a consequence to our discussion in this paper, is the dichotomy between sharpness of the bound and computability. That is, we observed the transition of sharp convergence estimates that yield very little practical value in the design of new methods to bounds that play a key role in defining a multigrid hierarchy yet result in poor predictors of actual performance.

Consequently, through our theoretical presentation, a larger issue emerges: soft upper bounds on convergence do not always adequately guide the design of new methods. Indeed, many AMG variants and many approaches to convergence theory exist. However, most theoretical analyses take the form of upper bounds on convergence, which, as we have observed, may not adequately predict performance.



DOI: 10.1002/nla

ACKNOWLEDGEMENTS

The work of Olson was partially supported by the NSF under award DMS 07–46676. The work of MacLachlan was partially supported by the NSF under award DMS 0811022.

REFERENCES

1. Brandt A. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation* 1977; **31**(138):333–390.
2. Brandt A, McCormick SF, Ruge JW. Algebraic multigrid (AMG) for automatic multigrid solution with application to geodetic computations. *Technical Report*, Institute for Computational Studies, Colorado State University, 1982.
3. Brandt A, McCormick SF, Ruge JW. Algebraic multigrid (AMG) for sparse matrix equations. In *Sparsity and Its Applications*, Evans DJ (ed.). Cambridge University Press: Cambridge, 1984; 257–284.
4. Ruge JW, Stüben K. Algebraic multigrid. In *Multigrid Methods*, Vol. 3, Frontiers in Applied Mathematics. SIAM: Philadelphia, PA, 1987; 73–130.
5. Napov A, Notay Y. Comparison of bounds for V-cycle multigrid. *Applied Numerical Mathematics. An IMACS Journal* 2010; **60**(3):176–192.
6. Napov A, Notay Y. When does two-grid optimality carry over to the V-cycle? *Numerical Linear Algebra with Applications* 2010; **17**(2-3):273–290.
7. Vassilevski PS. *Multilevel Block Factorization Preconditioners*. Springer: New York, 2008.
8. Xu J. Iterative methods by space decomposition and subspace correction. *SIAM Review. A Publication of the Society for Industrial and Applied Mathematics* 1992; **34**(4):581–613.
9. Notay Y, Vassilevski PS. Recursive Krylov-based multigrid cycles. *Numerical Linear Algebra with Applications* 2008; **15**(5):473–487.
10. Brezina M, Cleary AJ, Falgout RD, Henson VE, Jones JE, Manteuffel TA, McCormick SF, Ruge JW. Algebraic multigrid based on element interpolation (AMGe). *SIAM Journal on Scientific Computing* 2000; **22**(5):1570–1592.
11. Falgout RD, Vassilevski PS. On generalizing the algebraic multigrid framework. *SIAM Journal on Numerical Analysis* 2004; **42**(4):1669–1693.
12. Falgout RD, Vassilevski PS, Zikatanov LT. On two-grid convergence estimates. *Numerical Linear Algebra with Applications* 2005; **12**(5-6):471–494.
13. Stüben K, Trottenberg U. Multigrid methods: fundamental algorithms, model problem analysis and applications. In *Multigrid Methods*, Vol. 960, Hackbusch W, Trottenberg U (eds), Lecture Notes in Mathematics. Springer–Verlag: Berlin, 1982; 1–176.
14. Maitre JF, Musy F. Algebraic formalisation of the multigrid method in the symmetric and positive definite case—a convergence estimation for the $V$-cycle. In *Multigrid Methods for Integral and Differential Equations (Bristol, 1983)*, Vol. 3, Institute of Mathematics and Its Applications Conference Series New Series. Oxford University Press: New York, 1985; 213–223.
15. Bank RE, Douglas CC. Sharp estimates for multigrid rates of convergence with general smoothing and acceleration. *SIAM Journal on Numerical Analysis* 1985; **22**(4):617–633.
16. McCormick SF. An algebraic interpretation of multigrid methods. *SIAM Journal on Numerical Analysis* 1982; **19**(3):548–560.
17. McCormick SF. Multigrid methods for variational problems: further results. *SIAM Journal on Numerical Analysis* 1984; **21**(2):255–263.
18. McCormick SF. Multigrid methods for variational problems: general theory for the $V$-cycle. *SIAM Journal on Numerical Analysis* 1985; **22**(4):634–643.
19. Brandt A. Algebraic multigrid theory: the symmetric case. *Applied Mathematics and Computation* July 1986; **19**:23–56.
20. Bramble JH, Pasciak JE, Wang J, Xu J. Convergence estimates for product iterative methods with applications to domain decomposition. *Mathematics of Computation* 1991; **57**:1–21.
21. Xu J, Zikatanov L. The method of alternating projections and the method of subspace corrections in Hilbert space. *Journal of the American Mathematical Society* 2002; **15**(3):573–597.
22. Mandel J. Algebraic study of multigrid methods for symmetric, definite problems. *Applied Mathematics and Computation* 1988; **25**(1, part I):39–56.
23. Jones JE, McCormick SF. Parallel multigrid methods. In *Parallel Numerical Algorithms (Hampton, VA, 1994)*, Vol. 4, ICASE LaRC Interdisciplinary Series in Science and Engineering. Kluwer Acad. Publ.: Dordrecht, 1997; 203–224.
24. Henson VE, Yang UM. BoomerAMG: a parallel algebraic multigrid solver and preconditioner. *Applied Numerical Mathematics. An IMACS Journal* 2002; **41**(1):155–177. Developments and trends in iterative methods for large systems of equations—in memoriam Rüdiger Weiss (Lausanne, 2000).
25. Kolev TV, Vassilevski PS. Auxiliary space AMG for H(curl) problems. In *Domain Decomposition Methods in Science and Engineering XVII*, Vol. 60, Langer U, Discacciati M, Keyes DE, Widlund OB, Zulehner W, Barth TJ, Griebel M, Keyes DE, Nieminen RM, Roose D, Schlick T (eds), Lecture Notes in Computational Science and Engineering. Springer Berlin: Heidelberg, 2008; 147–154.
26. Olson LN, Schroder JB. Smoothed aggregation for Helmholtz problems. *Numerical Linear Algebra with Applications* 2010; **17**(2–3):361–386.



DOI: 10.1002/nla

27. Reitzinger S, Schöberl J. An algebraic multigrid method for finite element discretizations with edge elements. *Numerical Linear Algebra with Applications* 2002; **9**:223–238.
28. Chartier T, Falgout RD, Henson VE, Jones J, Manteuffel T, McCormick S, Ruge J, Vassilevski PS. Spectral AMGe ($\rho$ AMGe). *SIAM Journal on Scientific Computing* 2003; **25**(1):1–26.
29. Brezina M, Falgout R, MacLachlan S, Manteuffel T, McCormick S, Ruge J. Adaptive algebraic multigrid. *SIAM Journal on Scientific Computing* 2006; **27**(4):1261–1286.
30. Brezina M, Falgout R, MacLachlan S, Manteuffel T, McCormick S, Ruge J. Adaptive smoothed aggregation $(\alpha\mathrm{SA})$ multigrid. *SIAM Review* 2005; **47**(2):317–346.
31. Brezina M, Falgout R, MacLachlan S, Manteuffel T, McCormick S, Ruge J. Adaptive smoothed aggregation $(\alpha SA)$. *SIAM Journal on Scientific Computing* 2004; **25**(6):1896–1920.
32. Vaněk P, Mandel J, Brezina M. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing. Archives for Scientific Computing* 1996; **56**(3):179–196.
33. Brandt A. General highly accurate algebraic coarsening. *Electronic Transactions on Numerical Analysis* 2000; **10**:1–20.
34. Brannick JJ, Falgout RD. Compatible relaxation and coarsening in algebraic multigrid. *SIAM Journal on Scientific Computing* 2010; **32**(3):1393–1416.
35. Livne O. Coarsening by compatible relaxation. *Numerical Linear Algebra with Applications* 2004; **11**(2):205–227.
36. MacLachlan S, Saad Y. A greedy strategy for coarse-grid selection. *SIAM Journal on Scientific Computing* 2007; **29**(5):1826–1853.
37. Brandt A, Brannick J, Kahl K, Livshits I. Bootstrap AMG. *SIAM Journal on Scientific Computing* 2011; **33**(2):612–632.
38. McCormick SF, Ruge JW. Multigrid methods for variational problems. *SIAM Journal on Numerical Analysis* 1982; **19**(5):924–929.
39. Shaĭdurov VV. *Multigrid Methods for Finite Elements*, Mathematics and its Applications, Vol. 318. Kluwer Academic Publishers Group: Dordrecht, 1995. Translated from the 1989 Russian original by N. B. Urusova and revised by the author.
40. Hackbusch W. Convergence of multi-grid iterations applied to difference equations. *Mathematics of Computation* April 1980; **34**(150):425–440.
41. Aricò A, Donatelli M. A $V$-cycle multigrid for multilevel matrix algebras: proof of optimality. *Numerische Mathematik* 2007; **105**(4):511–547.
42. Hackbusch W. On the convergence of multi-grid iteration. *Technical Report 79-4*, Mathematisches Institut, Universität zu Köln, 1979.
43. Hackbusch W. Survey of convergence proofs for multigrid iteration. In *Special Topics in Applied Mathematics*, Frehse J, Pallasche D, Trottenberg U (eds). North Holland Publishing Company: Amsterdam, New York, 1980; 151–164.
44. Hiptmair R. Multigrid method for $H(\mathrm{div})$ in three dimensions. *Electronic Transactions on Numerical Analysis* 1997; **6**(Dec.):133–152.
45. Notay Y. Algebraic multigrid and algebraic multilevel methods: a theoretical comparison. *Numerical Linear Algebra with Applications* 2005; **12**(5-6):419–451.
46. Axelsson O. *Iterative Solution Methods*. Cambridge University Press: Cambridge, 1994.
47. MacLachlan S, Manteuffel T, McCormick S. Adaptive reduction-based AMG. *Numerical Linear Algebra with Applications* 2006; **13**:599–620.
48. Brandt A, Ron D. Multigrid solvers and multilevel optimization strategies. In *Multilevel Optimization in VLSI-CAD*, Vol. 14, Cong J, Shinnerl JR (eds), Combinatorial Optimization. Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; 1–69.
49. Napov A, Notay Y. An algebraic multigrid method with guaranteed convergence rate. *SIAM Journal on Scientific Computing* 2012; **34**(2):A1079–A1109.
50. Brannick J, Chen Y, Kraus J, Zikatanov L. Algebraic multilevel preconditioners for the graph Laplacian based on matching in graphs. *SIAM Journal on Numerical Analysis* 2013; **51**(3):1805–1827.
51. Manteuffel T, McCormick S, Park M, Ruge J. Operator-based interpolation for bootstrap algebraic multigrid. *Numerical Linear Algebra with Applications* 2010; **17**(2-3):519–537.
52. Brezina M, Ketelsen C, Manteuffel T, McCormick S, Park M, Ruge J. Relaxation-corrected bootstrap algebraic multigrid (rBAMG). *Numerical Linear Algebra with Applications* 2012; **19**(2):178–193.
53. Napov A, Notay Y. Algebraic analysis of aggregation-based multigrid. *Numerical Linear Algebra with Applications* 2011; **18**(3):539–564.
54. Heys JJ, Manteuffel TA, McCormick SF, Olson LN. Algebraic multigrid for higher-order finite elements. *Journal of Computational Physics* 2005; **204**(2):520–532.



DOI: 10.1002/nla