A new perspective on strength measures in algebraic multigrid

Luke N. Olson[1], Jacob Schroder[1, *, †] and Raymond S. Tuminaro[2]

[1]*Siebel Center for Computer Science, University of Illinois at Urbana-Champaign, 201 N. Goodwin Avenue, Urbana, IL 61801, U.S.A.*
[2]*Sandia National Laboratories, PO Box 969, MS 9159, Livermore, CA 94551, U.S.A.*

SUMMARY

Algebraic-based multilevel solution methods (e.g. classical Ruge–Stüben and smoothed aggregation style algebraic multigrid) attempt to solve or precondition sparse linear systems without knowledge of an underlying geometric grid. The automatic construction of a multigrid hierarchy relies on *strength-of-connection* information to coarsen the matrix graph and to determine sparsity patterns for the inter-grid transfer operators. Strength-of-connection as a general concept is not well understood and the first task of this paper is therefore on understanding existing strength-of-connection measures and their limitations. In particular, we present a framework to interpret and clarify existing measures through differential equations. This framework leads to a new procedure for making pointwise strength-of-connection decisions that combines knowledge of local algebraically smooth error and of the local behavior of interpolation. The new procedure effectively addresses a variety of challenges associated with strength-of-connection and when incorporated within an algebraic multigrid procedure gives rise to a robust and efficient solver. Copyright © 2009 John Wiley & Sons, Ltd.

1. INTRODUCTION

Algebraic multigrid (AMG) [1, 2] is an attractive black box solver or preconditioner. Coarse problems are automatically constructed by defining coarse sets of degrees-of-freedom and associated inter-grid transfer operators without explicit knowledge of the underlying problem.

The focus of this paper is on strength-of-connection in the matrix graph between two degrees-of-freedom. Strength-of-connection information is used in coarse grid selection and in construction of the inter-grid transfer operators. Specifically, strength-of-connection information is used to

construct a directed graph, G, whose vertices are the degrees-of-freedom present in the operator, A, and where vertex i is connected by an edge to vertex j only if i is strongly connected to j. The coarse grid is constructed through a graph coarsening of G. Strength information is often used to determine the sparsity pattern of an inter-grid transfer operator by enforcing that degree-of-freedom j interpolates to degree-of-freedom i only if i is strongly connected to j.

The ideal strength-of-connection measure for AMG is neither well understood nor well defined. Classic strength-of-connection decisions are often based directly on the matrix stencil. AMG was first proposed in [1] and further developed in [2], where the 'classical' strength-of-connection measure is proposed. Degree-of-freedom i is strongly connected to j if

$$-A_{ij} \geqslant \theta \max_{m \neq i} -A_{im} \tag{1a}$$

for some user supplied $\theta \in [0, 1]$. Similarly, smoothed aggregation (SA) [3, 4] sets i to be strongly connected to j if

$$|A_{ij}| \geqslant \theta \sqrt{A_{ii} A_{jj}}. \tag{1b}$$

Both measures are motivated by M-matrix assumptions on A, which yield heuristics that relate matrix entries to algebraically smooth error and the complementary interpolation operators. Consequently while (1a) and (1b) are reliable and efficient measures in many situations, their utility is limited for more general non-M-matrices. One example is the grid-aligned anisotropic diffusion equation,

$$-u_{xx} - \epsilon u_{yy} = f, \tag{2}$$

on a uniform mesh using bilinear Q1 finite elements. Equation (3a) gives the corresponding generic matrix stencil. The 3×3 array notation displays the coefficients in the matrix from a degree-of-freedom (center point) to its eight nearest neighbors in the mesh

$$[A]_\epsilon = \frac{1}{3} \begin{pmatrix} -\dfrac{1+\epsilon}{2} & 1-2\epsilon & -\dfrac{1+\epsilon}{2} \\ -2+\epsilon & 4+4\epsilon & -2+\epsilon \\ -\dfrac{1+\epsilon}{2} & 1-2\epsilon & -\dfrac{1+\epsilon}{2} \end{pmatrix}. \tag{3a}$$

Setting $\epsilon = 1.0$ and $\epsilon = 0.0$ yields

$$[A]_{1.0} = \frac{1}{3} \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad \text{and} \quad [A]_{0.0} = \frac{1}{3} \begin{pmatrix} -\dfrac{1}{2} & 1 & -\dfrac{1}{2} \\ -2 & 4 & -2 \\ -\dfrac{1}{2} & 1 & -\dfrac{1}{2} \end{pmatrix}. \tag{3b}$$

When ϵ is small, the y-direction is weakly coupled and standard pointwise smoothing, such as Gauss–Seidel, is ineffective at reducing error that is oscillatory in the y-direction. In geometric multigrid, one solution is to coarsen only in the x-direction—i.e. the direction in which the relaxed error is smooth. However, to algebraically coarsen only in the x-direction the strength-of-connection

measure must determine that coupling in the vertical and diagonal directions is weak compared to the horizontal direction. Unfortunately, the classic measures only partially indicate a strong coupling in the horizontal direction. As a result, direct use of the matrix coefficients alone is not sufficient to reliably identify a connection strength.

The matrix inverse, A^{-1}, has been used to motivate ideal strength-of-connection information. We outline here, however, a simple example for which the matrix inverse is not appropriate. Consider (2) on the unit square with $\epsilon = 0.001$. Let the problem be discretized with linear elements on a uniform simplicial mesh and apply Dirichlet boundary conditions at $x = 0.0, 1.0$ and pure Neumann boundary conditions at $y = 0.0$ and 1.0. Equation (4) compares the stencils of the matrix and the matrix inverse, respectively, at a point in the center of the domain. The inverse is generally dense and so only the stencil for the local neighborhood of the center point is given.

For this problem, A is an M-matrix and the matrix stencil itself clearly indicates weak connections in the y-direction. The inverse, however, indicates nearly isotropic behavior, yielding incorrect strength-of-connection values. Additional examples using mixed boundary conditions also illustrate the misleading global nature of the matrix inverse, but the misleading character of the inverse can also be exposed through the PDE. One such example is the rotated anisotropic diffusion equation with Q1 elements and Dirichlet boundary conditions

$$[A] = \begin{pmatrix} & -0.001 & \\ -1.00 & 2.002 & -1.00 \\ & -0.001 & \end{pmatrix}, \quad [A^{-1}] = \begin{pmatrix} 225.0 & 225.0 & 225.0 \\ 250.7 & 251.6 & 250.7 \\ 225.0 & 225.0 & 225.0 \end{pmatrix}. \quad (4)$$

The main deficiencies in using the matrix inverse itself is its global nature and the fact that it contains both high- and low-energy information (the energy of a mode, v, is defined as $\|v\|_A / \|v\|$ for a symmetric positive-definite matrix, A). The low-energy modes correspond to algebraically smooth error and must be captured by coarse level correction, while high-energy modes are reduced by relaxation and need not be considered when constructing a coarse level operator.

Interestingly, local approximations to the matrix inverse yield effective strength information since the measure is no longer global in nature. Local inverse approximations are pursued in [5, 6], however, locality is pursued to reduce the cost of the approximation rather than as a method to a better strength-of-connection measure. It is recognized in [5] that approximations to the inverse may provide better information if they are generated by relaxation, since high-energy modes are filtered when using a moderate number of steps. Yet even the local relaxation-based approximations are not ideal. The measure in [5] can lead to less valuable strength information even after a modest number of relaxation steps as the matrix inverse is approached. We illustrate this effect in Section 4.1.

A promising approach related to the matrix inverse is associated with compatible relaxation [7, 8] and follows more naturally from classical Ruge–Stüben AMG, as opposed to SA which is pursued in this paper. In a classical AMG framework, one splits the unknowns into two sets, F-points and C-points, where the C-points effectively define the next coarser level grid. Instead of looking at A^{-1}, one can now look at the inverse of the submatrix associated with only the F-points, i.e. A_{ff}^{-1}. This matrix is less global than A^{-1} and can provide better strength information [9].

Given the weaknesses of the matrix stencil and the matrix inverse, we are motivated to pursue a more suitable way to determine strength-of-connection information. In this paper, we develop a new strength measure that also uses a relaxation-based process, but does not approximate the matrix

inverse. The new strength measure incorporates strength-of-connection information, the local nature of algebraically smooth error, and the local behavior of interpolation. To infer local information about algebraically smooth error, we consider the connection between weighted-Jacobi-relaxation and the time integration of ordinary differential equations (ODEs) for the specific case of evolving δ-functions. To account for the local behavior of interpolation, the new strength measure considers the user-provided near null-space modes. In a SA setting [3], the prolongation and restriction operators are constructed directly from these modes. Hence, it is appropriate to view them as indicative of interpolation. We consider this integration of local knowledge about algebraically smooth error and interpolation to be one of the main achievements of the paper.

In Section 2.1, a novel ODE perspective is presented that examines the time evolution of δ-functions that is equivalent to Jacobi-relaxation. This new perspective further exposes the difficulties associated with existing strength-of-connection measures and provides a useful way to generate local algebraically smooth functions. In Section 3.1, the Approximate Inverse Measure from [5] is discussed. In Section 3.2, the new strength-of-connection measure is proposed and analyzed. A proof of scale invariance is given, followed by a discussion of the complexity. Lastly in Section 4, numerical evidence is provided in support of the new measure. Convergence results are encouraging and indicate that the new strength measure is more parameter independent than the classical measures—i.e. performance is less sensitive to the tuning of the drop-tolerance parameter.

2. ODE PERSPECTIVE ON STRENGTH

2.1. ODE perspective

Consider the transient solution of (2) with $\epsilon = 0$ and $f = 0$ given by

$$u_t = -u_{xx} \tag{5}$$

and take $u(t=0)$ to be a Dirac δ-function centered at (x^*, y^*). The initial evolution of this function indicates directions from (x^*, y^*) that are most strongly influenced by the solution at (x^*, y^*) and these directions are used in identifying strength-of-connection . The physics of (5) suggests no coupling in the y-direction, so the δ-function diffuses only in the x-direction as time evolves—i.e. $u(x, y) = 0$ for $y \neq y^*$ and $t \geqslant 0$.

Discretizing (5) only in space yields the ODE,

$$u_t = -Au \quad \text{with} \ u(0) = \delta_i \tag{6}$$

where δ_i is a Dirac δ-function centered at the ith grid point and A is a symmetric positive-definite matrix that is diagonally scaled to a unit diagonal. The exact solution to (6) is

$$u = \mathrm{e}^{-At}\delta_i. \tag{7}$$

When A corresponds to a Laplacian, (6) models the time diffusion of an initial point distribution, δ_i, with steady-state $u = 0$. In this case, the evolution of δ_i in (6) clearly describes the connection between point i and a neighboring point j.

At $t = 0$, we have

$$u_t = -A\delta_i. \tag{8}$$

$A\delta_i$ is simply the matrix coefficients in the ith row or column, as A is symmetric. The growth of u at j for $t=0$ is given by the size of the coefficient A_{ij}. This is identical to the classical strength-of-connection measures, (1a) and (1b), which are based on the matrix stencil. From this perspective, the matrix stencil is valuable because it describes how u changes in time based on the information at point i. However as indicated above, this direct use of matrix entries is misleading and motivates a discussion of the relationship with the underlying discretization.

The transient solution of (5) has no coupling in the y-direction. The use of Q1 finite elements, however, yields a matrix with the stencil given by $[A]_{0.0}$ in (3b), which has nontrivial couplings in the y-direction. This discrepancy is due to the omission of a mass matrix, M, on the u_t term, which properly reflects a finite element representation of the ODE.

If Q1 finite elements are used, then M has the stencil

$$[M] = \frac{1}{36} \begin{pmatrix} 1 & 4 & 1 \\ 4 & 16 & 4 \\ 1 & 4 & 1 \end{pmatrix}. \tag{9}$$

The generic stencil for $M^{-1}A$ on a regular grid with Dirichlet boundary conditions is $[M^{-1}A]$ in (10). $M^{-1}A$ more accurately reflects the ODE (6), in that it recovers the absence of any coupling in the y-direction

$$[M^{-1}A] = \begin{pmatrix} 0 & 0 & 0 & -0.0028 & 0 & 0 & 0 \\ -0.20 & 0.75 & -2.78 & 4.40 & -2.78 & 0.75 & -0.20 \\ 0 & 0 & 0 & -0.0028 & 0 & 0 & 0 \end{pmatrix}. \tag{10}$$

Therefore, the problem from a strength-of-connection perspective is that smearing is reflected in the matrix stencil for A and is significantly reduced by considering a mass matrix. Using Jacobi iterations to approximate $M^{-1}A$ is not feasible since AMG methods do not have access to a mass matrix (on fine or projected coarse levels). Further, (3b) may represent a valid finite difference stencil of the operator u_{xx}, thus necessitating strength-of-connection techniques that properly address (3b)-type stencils without requiring the use of a mass matrix.

Consider the smearing reflected in the matrix stencil (3b) via a Taylor series expansion of the truncation error. The y-direction terms in (3b) give rise to cross derivatives that are $O(h^2)$. These terms are insignificant for smooth functions and small h. However, δ-functions are not smooth, yielding large error terms for initial time steps. After the initial time steps, the function smooths and the troublesome cross coupling terms are less significant.

Given the difficulties associated with (8), it is natural to consider integration of (6) to an intermediate time t_f and to determine strength-of-connection based on the evolved δ-function. The steady-state solution to (6) corresponds to $u=0$, which provides no strength-of-connection information. This is similar to approximating the matrix inverse in that steady-state information is global and often has little accurate local strength-of-connection information. Thus, a balance is sought in choosing t_f. We smooth the δ-function sufficiently, but avoid global effects by limiting t_f. The specific choice of t_f is further discussed in Section 3.2.2.

Numerically, we solve (6) by the forward Euler method, resulting in an iteration of the form

$$u^{(k)} = (I - \Delta t A)^k \delta_i. \tag{11}$$

of the spectrum range where the Chebyshev polynomial is focused. A possible fix might consider reducing the spectrum range in the polynomial definition each time the order is increased, though this would require rethinking Chebyshev recurrences. Another possible fix could consider some type of weighted Chebyshev approximations.

Other polynomial alternatives could be based on weighted-Jacobi, or as considered in [5], local Gauss–Seidel methods, which are effectively a combination of Jacobi and Gauss–Seidel. Using Gauss–Seidel within strength measures is attractive because Gauss–Seidel is commonly used as a multigrid smoother. The complementary relationship between smoother and coarse level correction argues that coarse level operators (which are based in part on strength decisions) should be defined using information from the multigrid smoother. Unfortunately, a single Gauss–Seidel iteration can propagate information across the entire domain. This makes Gauss–Seidel very costly within a strength measure. Local Gauss–Seidel can be used to mitigate this issue, though this is somewhat awkward. Additionally, one has to either cap the Gauss–Seidel iterations or tune the iterations in some way to focus only on high-energy modes.

One final alternative should be mentioned when strength is considered in the context of classical Ruge–Stüben-like AMG methods, as opposed to the SA setting here. As previously mentioned, points are split into two sets, F- and C-points. Using compatible relaxation arguments, one could assume that A_{ff} is well-conditioned [11]. Further, an ideal interpolation operator of $A_{ff}^{-1}A_{fc}$ can be justified based on a Schur complement perspective, where A_{ff} is the submatrix corresponding to F-points and A_{fc} is the submatrix associated with F-point rows and C-point columns. As A_{ff} is well-conditioned and its inverse is less global than the true inverse, a polynomial approximation to the entire spectrum of A_{ff}^{-1} could be used to generate prolongator sparsity patterns or strength information [9].

The ODE-based approach deviates somewhat from a traditional polynomial perspective. However, this approach provides an effective means and logic for defining strength-of-connection and so indirectly provides a mechanism for defining a polynomial. It exhibits good results in practice and provides a natural connection between weighted-Jacobi and ODE time-stepping. Lastly, this approach also provides a link to the original physical problem.

3. STRENGTH MEASURES

The numerically integrated ODEs of the previous section trace the evolution of a δ-function in time. This evolution gives a basic indication of the local nature of algebraically smooth error. In this section, we describe two techniques for transforming the evolved δ-function to concrete strength-of-connection measures. For generality, we no longer assume that matrix A is diagonally scaled—i.e. we consider the ODEs (6) and (12) with $D^{-1}A$ as opposed to A.

3.1. Approximate inverse measure

First, we consider a technique from [5], which is based on the columns of the inverse of A. The Approximate Inverse Measure is derived by applying a local relaxation method to obtain approximations of the inverse. For the case of weighted-Jacobi, this is equivalent to numerically

solving the ODE (12). Denoting S as a matrix of strength-of-connection entries and letting z be the ith approximate column of the inverse,

$$S_{ij} = \frac{\|z^{(j)}\|_A - \|z\|_A}{\|z\|_A},$$ (14)

where $z_k^{(j)} = z_k$ for $k \neq j$ and $z_j^{(j)} = 0$. Measure (14) corresponds to a normalized change in energy for z. Assuming z is an algebraically smooth vector with local support, z is a suitable interpolatory basis function around i. A large change in energy when comparing $z^{(j)}$ to z indicates that the jth value is significant in reducing the energy of the ith interpolation basis function. Thus, j strongly influences i. Through this use of A-norms, the measure indirectly accounts for how algebraically smooth error local to i should be interpolated.

In Section 4.1, we illustrate several advantages of the Approximate Inverse Measure and show how it performs on model problems. However, the measure is motivated through the matrix inverse, which as previously noted, may not provide sufficient strength information. In particular, (14) reduces to

$$S_{ij} = \sqrt{1 + \frac{A_{jj}z_j^2}{z_i}} - 1 = \sqrt{1 + \frac{z_j^2}{z_i}} - 1$$ (15)

in the limit of relaxation where z is the ith column of A^{-1} and A is assumed to have unit diagonal. When z_j^2/z_i is small (as occurs for a 3D Laplace operator with a standard 7-pt stencil) an approximation is

$$S_{ij} \approx \frac{z_j^2}{2z_i}.$$ (16)

When comparing two elements within the ith row to identify the stronger connection, the key factor is given by z_j^2 (as z_i is the same for both elements in row i). Thus, this measure effectively squares z_j which accentuates differences between different values of z_j. For strongly anisotropic phenomena this is helpful, but may offer misleading information for mildly anisotropic problems.

3.2. Evolution measure

To define strength-of-connection for a degree-of-freedom i, knowledge is first obtained about local algebraically smooth error by letting z be an evolved δ-function centered at i. We choose the time-stepping formula (11) for the ODE (6), in order to avoid any of the misleading effects of the inverse. To make strength-of-connection decisions, we measure how well z will be pointwise interpolated by the eventual multigrid cycle. This is problematic because strength-of-connection decisions are necessary in order to construct the interpolation operators. However because of the SA framework, the construction process of the tentative prolongator is known and naturally indicates how interpolation is expected to behave. That is, strength-of-connection for i is defined in terms of the ability of the tentative prolongator to interpolate z in the neighborhood of i.

The tentative prolongator is constructed during the SA setup phase. Here, SA coarsens a matrix graph by first aggregating nodes. An *ideal* aggregate consists of a root node and all strongly connected neighbors. SA is provided a set of vectors, $B = \{B_j\}$, corresponding to a basis of the

near null-space—e.g. null-space functions of the underlying PDE. For example, the near null-space for the standard scalar elliptic PDE is the constant vector, while for elasticity in 3D, the six rigid body modes are effective. A tentative prolongator is constructed such that each block column corresponds to one aggregate with the columns per block given by the number of near null-space vectors. Each block column is nonzero only at the associated degrees-of-freedom for the corresponding aggregate, where the near null-space vectors are injected. The tentative prolongator represents the near null-space components, which are algebraically smooth and slow to relax on the fine mesh. Effective SA prolongators are constructed by improving the tentative prolongator, while maintaining the representation of B [12].

If the tentative prolongator effectively interpolates z, then so will the multigrid prolongator, since it also maintains the representation of B. In addition, we assume that z is representative of algebraically smooth error local to i. If these assumptions hold for the z corresponding to each degree-of-freedom, then the range of the multigrid prolongator will accurately represent algebraically smooth error and be effective within a multigrid cycle.

As a result, our goal is to ensure that tentative prolongation accurately represents the z generated for each degree-of-freedom. This will yield aggregates that correspond to strongly connected degrees-of-freedom. To do this, consider a mock aggregate centered at i that consists of Ω_i, the union of i with i's strongly connected degrees-of-freedom. That is,

$$\Omega_i = \{j \mid A_{ij} \neq 0 \text{ and } i \text{ is strongly connected to } j\}.$$

Initially, we assume that all neighboring connections are strong. Denote by $\bar{B} = B|_{\Omega_i}$ the restriction of the near null-space basis to the mock aggregate, which corresponds to the mock aggregate's block column in a tentative prolongator. In general, we use an overbar symbol to denote restrictions of quantities to Ω_i (with a corresponding local numbering within the mock aggregate). For a vector, this restriction corresponds to a subset of elements from the original vector. For matrices, the restriction corresponds to a subset of rows and columns. For example, z_j is the jth element of z and $\bar{z}_{\bar{j}}$ refers to the same element.

For each z, we define a strength-of-connection measure that allows for the largest mock aggregate that yields an accurate tentative prolongator. To do this, the strength-of-connection measure determines how well each element of $\bar{z}$ is represented by a best approximation of $\bar{z}$ in the span of $\bar{B}$. For $B \in \mathbb{R}^{n,m}$, consider the constrained minimization problem

$$\min_{x \in \mathbb{R}^m} \|\bar{B}x - \bar{z}\|_Q^2 \tag{17}$$

subject to the linear constraint

$$\langle \bar{B}x - \bar{z}, \bar{e} \rangle_Q = 0, \tag{18}$$

where Q is either I or $\bar{D}_A$, $\bar{D}_A = \text{diag}(A)|_{\Omega_i}$ and $\bar{e}$ is the ith unit vector restricted to Ω_i. The linear constraint centers strength decisions around degree-of-freedom i.

Using a Lagrange multiplier, λ, this constrained minimization problem is equivalent to the solution of the linear system,

$$\begin{bmatrix} 2\bar{B}^{\mathrm{T}}Q\bar{B} & q\bar{B}^{\mathrm{T}}\bar{e} \\ \bar{e}^{\mathrm{T}}\bar{B} & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 2\bar{B}^{\mathrm{T}}Q\bar{z} \\ \bar{z}_{\bar{i}} \end{bmatrix}, \tag{19}$$

where $q = Q_{\bar{i}\bar{i}}$. Let $\widetilde{z} = \bar{B}x$, such that x solves (19). Then $\widetilde{z}$ is the best approximation to $\bar{z}$ in the span of $\bar{B}$ subject to the constraint (18). We define the strength-of-connection between degrees-of-freedom i and j to be the pointwise relative approximation error for $\widetilde{z}$ with respect to $\bar{z}$ at degree-of-freedom, j:

$$S_{ij} = \left| \frac{\bar{z}_{\bar{j}} - \widetilde{z}_{\bar{j}}}{\bar{z}_{\bar{j}}} \right|. \tag{20}$$

Since S_{ij} represents relative approximation error, smaller values imply stronger connections.

Next, we show that for $B \in \mathbb{R}^n$, measure (20) reduces to a simple ratio that does not require directly solving (19) and is independent of the choice of Q.

Theorem 1
For $B \in \mathbb{R}^n$ and for $Q = I$ or $Q = \bar{D}_A$, measure (20) reduces to

$$S_{ij} = \left| 1 - \frac{\bar{B}_{\bar{j}}\bar{z}_{\bar{i}}}{\bar{B}_{\bar{i}}\bar{z}_{\bar{j}}} \right|. \tag{21}$$

Proof
Since $B \in \mathbb{R}^n$, then $x \in \mathbb{R}$ and the solution to the constrained minimization problem (17)–(18) is solved only by enforcing (18), which yields

$$x\bar{B}^{\mathrm{T}}Q\bar{e} = \bar{z}^{\mathrm{T}}Q\bar{e}$$

$$Q_{\bar{i}\bar{i}}\bar{B}_{\bar{i}}x = Q_{\bar{i}\bar{i}}\bar{z}_{\bar{i}}$$

$$x = \frac{\bar{z}_{\bar{i}}}{\bar{B}_{\bar{i}}}.$$

By definition, $\widetilde{z} = \bar{B}x$, and $\widetilde{z}_{\bar{j}} = (\bar{B}x)_{\bar{j}} = \bar{B}_{\bar{j}}x = (\bar{B}_{\bar{j}}\bar{z}_{\bar{i}})/\bar{B}_{\bar{i}}$. Substituting into (20), we arrive at (21). □

3.2.1. Scale invariance. As shown in [5], an important property of the Approximate Inverse Measure is invariance to symmetric diagonal scaling of A. This is an important property for any strength-of-connection measure since it ensures that scaling the underlying equations does not distort the strength measure. The measure (20) also exhibits this property when $Q = \bar{D}_A$.

Lemma 1
Let A be a nonsingular matrix with nonzero diagonal in $\mathbb{R}^{n \times n}$ and $\hat{A} = DAD$, where D is an arbitrary nonsingular diagonal matrix. Let

$$\omega = \frac{1}{\rho(D_{\hat{A}}^{-1}\hat{A})} = \frac{1}{\rho(D_A^{-1}A)}$$

be the weighted-Jacobi damping parameter (i.e. Δt) for both A and $\hat{A}$ and let $D_A = \mathrm{diag}(A)$ and $D_{\hat{A}} = \mathrm{diag}(\hat{A})$. If $z = (I - \omega D_A^{-1}A)^k \delta_i$ and $\hat{z} = (I - \omega D_{\hat{A}}^{-1}\hat{A})^k \delta_i$, then $\hat{z} = D_{ii}D^{-1}z$, for $k \in \mathbb{Z}^+$.

Proof

δ_i is first smoothed k times with $\hat{A}$ in order to derive a relationship to smoothing with A. We also use the identity, $D_{\hat{A}} = D^2 D_A$.

$$
\begin{aligned}
\hat{z} = (I - \omega D_{\hat{A}}^{-1} \hat{A})^k \delta_i &= (I - \omega D_{\hat{A}}^{-1} \hat{A}) D^{-1} D (I - \omega D_{\hat{A}}^{-1} \hat{A}) D^{-1} D \\
&\quad \dots (I - \omega D_{\hat{A}}^{-1} \hat{A}) D^{-1} D \delta_i \\
&= (D^{-1} - \omega D_A^{-1} D^{-1} A) D (D^{-1} - \omega D_A^{-1} D^{-1} A) D \\
&\quad \dots (D^{-1} - \omega D_A^{-1} D^{-1} A) D \delta_i \\
&= D_{ii} D^{-1} (I - \omega D_A^{-1} A)(I - \omega D_A^{-1} A) \dots (I - \omega D_A^{-1} A) \delta_i \\
&= D_{ii} D^{-1} z.
\end{aligned}
$$

$\square$

Lemma 2

Let A and $\hat{A}$ be defined as in Lemma 1. For degree-of-freedom i, $[x, \lambda]^{\mathrm{T}}$ solves the constrained minimization problem (17)–(18) for A with $Q = \bar{D}_A$, if and only if, $[D_{ii} x, \lambda]^{\mathrm{T}}$ solves the constrained minimization problem (17)–(18) for $\hat{A}$ with $Q = \bar{D}_{\hat{A}}$.

Proof

Assume $[D_{ii} x, \lambda]^{\mathrm{T}}$ solves system (19) for degree-of-freedom i of $\hat{A}$ with $Q = \bar{D}_{\hat{A}}$. Then (19) becomes

$$
\begin{bmatrix} 2 \bar{\hat{B}}^{\mathrm{T}} \bar{D}_{\hat{A}} \bar{\hat{B}} & \hat{A}_{ii} \bar{\hat{B}}^{\mathrm{T}} \bar{e} \\ \bar{e}^{\mathrm{T}} \bar{\hat{B}} & 0 \end{bmatrix} \begin{bmatrix} D_{ii} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 2 \bar{\hat{B}}^{\mathrm{T}} \bar{D}_{\hat{A}} \bar{\hat{z}} \\ \bar{\hat{z}}_{\bar{i}} \end{bmatrix}. \tag{22}
$$

Using Lemma 1, the identity $D_{\hat{A}} = D^2 D_A$ and the near null-space vectors, $\hat{B} = D^{-1} B$, of $\hat{A}$, we write the system in terms of the original operator, A.

$$
\begin{bmatrix} 2 \bar{B}^{\mathrm{T}} \bar{D}_A \bar{B} & D_{ii} A_{ii} \bar{B}^{\mathrm{T}} \bar{e} \\ D_{ii}^{-1} \bar{e}^{\mathrm{T}} \bar{B} & 0 \end{bmatrix} \begin{bmatrix} D_{ii} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 2 D_{ii} \bar{B}^{\mathrm{T}} \bar{D}_A \bar{z} \\ \bar{z}_{\bar{i}} \end{bmatrix}. \tag{23}
$$

Then canceling D_{ii} results in

$$
\begin{bmatrix} 2 \bar{B}^{\mathrm{T}} \bar{D}_A \bar{B} & A_{ii} \bar{B}^{\mathrm{T}} \bar{e} \\ \bar{e}^{\mathrm{T}} \bar{B} & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 2 \bar{B}^{\mathrm{T}} \bar{D}_A \bar{z} \\ \bar{z}_{\bar{i}} \end{bmatrix}, \tag{24}
$$

which yields the solution $[x, \lambda]^{\mathrm{T}}$ to the constrained minimization problem for A with $Q = \bar{D}_A$. The steps in the proof are reversible, thus proving the lemma. $\square$

Theorem 2

Measure (20) is invariant to symmetric diagonal scaling if $Q = \bar{D}_A$.

Proof

Let A and $\hat{A}$ be defined as in Lemma 1 and z and $\hat{z}$ be the time-evolved δ-functions, respectively. Also let S and $\hat{S}$ be the associated strength-of-connection matrices defined by (20) with $Q = \bar{D}_A$ and $Q = \bar{D}_{\hat{A}}$, respectively. We show that S and $\hat{S}$ are equivalent.

By Lemmas 1 and 2, we have $\bar{\bar{z}} = D_{ii}\bar{D}^{-1}\bar{z}$ and $\tilde{\tilde{z}} = \hat{\bar{B}}(D_{ii}x)$. Using $\hat{B} = D^{-1}B$, we then obtain $\tilde{\tilde{z}} = D_{ii}\bar{D}^{-1}\bar{B}x = D_{ii}\bar{D}^{-1}\tilde{z}$.

Using the above substitutions for $\tilde{\tilde{z}}$ and $\bar{\bar{z}}$ in the definition (20) of $\hat{S}$,

$$\hat{S}_{ij} = \left| \frac{\bar{\bar{z}}_{\bar{j}} - \tilde{\tilde{z}}_{\bar{j}}}{\bar{\bar{z}}_{\bar{j}}} \right| = \left| \frac{D_{ii}D_{jj}^{-1}\bar{z}_{\bar{j}} - D_{ii}D_{jj}^{-1}\tilde{z}_{\bar{j}}}{D_{ii}D_{jj}^{-1}\bar{z}_{\bar{j}}} \right| = \left| \frac{\bar{z}_{\bar{j}} - \tilde{z}_{\bar{j}}}{\bar{z}_{\bar{j}}} \right| = S_{ij}. \tag{25}$$

$\square$

3.2.2. Algorithm. The algorithm for the Evolution Measurecontains two approaches, Evolution Measure 1, which uses (20) and Evolution Measure 2, an energy-based approach related to [5]. Evolution Measure 1 uses a two-stage process. The first stage examines the sign of the ratio of $\tilde{z}_{\bar{j}}/\bar{z}_{\bar{j}}$. If the ratio is negative, then clearly $span(\bar{B})$ cannot approximate $\bar{z}$ at i and j well—i.e. the connection is weak. If the ratio is positive, then the second stage calculates measure (20). In line 7, we invert Equation (14) to maintain consistency with the concept of smaller values representing stronger connections.

Algorithm 1 relies on the parameters $k, \Delta t, \theta$ and B. The default, elliptic case uses B as a vector of ones. For Δt, we take a time-step that does not violate stability considerations and is commensurate with relaxation used in a multigrid cycle—i.e. does not accentuate oscillatory modes. Specifically, we take $\Delta t = 1/\rho(D^{-1}A)$, which works well in practice.

Algorithm 1. Evolution Measure(A, k, t_f, B, θ).

```
1  for i = 1,…,n
2      z = (I − (t_f/k)D^{-1}A)^k δ_i
3      for j ∋ A_ij ≠ 0
4          if (energy)                          /* Evolution Measure 2 */
5              z^(j) = z
6              z^(j)_j = 0
7              S_ij = | ‖z‖_A / (‖z^(j)‖_A − ‖z‖_A) |
8          else                                 /* Evolution Measure 1 */
9              x ← solution of (19)
10             z̃ = B̄x
11             if (z̃_j̄/z̄_j̄ < 0)
12                 S_ij = weak
13             else
14                 S_ij = | (z̄_j̄ − z̃_j̄)/z̄_j̄ |
15     for j ∋ A_ij ≠ 0                          /* apply drop−tolerance */
16         if S_ij > θ min({S_{i,m} : m ≠ i, A_{i,m} ≠ 0})
17             S_ij = weak
18 return (S)
```

Table I. Typical parameters for Algorithm 1.

k	$\max(\lfloor\rho(D^{-1}A)\rfloor, 1)$	Number of time steps
Δt	$\frac{1}{\rho(D^{-1}A)}$	Time step size
θ	4.0	Drop tolerance for weak connections
B	$\mathbf{1}$	Near null-space vector(s)

The parameter k is defined by choosing a final time, t_f, where $t_f = k\Delta t$. We seek t_f such that a significant portion of the δ-function is transferred to neighboring points, which indicates that the discretization error is damped. Experiments indicate that a precise t_f value is not critical, but we nonetheless offer heuristics for guidance. We stipulate that the maximum value of the δ-function be reduced by a factor of two from its original maximum. Formally, we want to find the smallest t_f such that $\|u(0)\|_\infty \geqslant 2\|u(t_f)\|_\infty$. Denote by $u_{t,i}$ the value of u_t at point i and assume that $u_{t,i}$ can be modeled in a linear fashion. Define $\alpha = u_{t,i}(t_f)$, and recall that $u_{t,i}(0) = -1$, since we are working with the ODE, $u_t = -D^{-1}Au$. Taking this linear model of $u_{t,i}$ and integrating it in time gives

$$u_{t,i}(t) = -1 + \frac{(1+\alpha)t}{t_f} \quad \text{and} \quad u_i(t) = -t + \frac{(1+\alpha)t^2}{2t_f} + C. \tag{26}$$

Using $u_i(0) = 1$, we arrive at $C = 1$. Setting $u_i(t_f) = \frac{1}{2}$ and solving for t_f results in

$$\frac{1}{2} = 1 - t_f + \frac{(1+\alpha)t_f^2}{2t_f} \quad \Rightarrow \quad t_f = \frac{1}{1-\alpha}. \tag{27}$$

We now determine an appropriate α. Consider $\alpha = -1$ implying that $u_{t,i}(t_f) = -1$. Normally, $u_{t,i}$ is largest at $t = 0$, so we view $u_{t,i}(t) = -1$ as an extreme case where the rate of decay is constant. This yields $t_f = \frac{1}{2}$. Consider another extreme where $\alpha = 0$, which represents no further decay in the δ-function at t_f. In this case, $t_f = 1$. We expect an accurate linear model of $u_{t,i}$ to behave between these two extremes and therefore consider

$$t_f \in \left[\frac{1}{2}, 1\right]. \tag{28}$$

Based on experiments, we favor $t_f = 1$. This leads to $k = \max(\lfloor\rho(D^{-1}A)\rfloor, 1)$, which is used in this paper. One reasonable alternative in parallel would be to dynamically iterate in time until $\|u\|_\infty$ drops by a factor of 2 from its original maximum.

Selecting t_f based on (28) results in sensible conclusions when comparing M-matrices to non-M-matrices. For an M-matrix arising from an elliptic PDE, $\rho(D^{-1}A) \approx 2$, which suggests 2 steps if $t_f = 1$. However, if the matrix is far from an M-matrix with a large $\rho(D^{-1}A)$, then more than 2 time steps may be required. Experimentally, we did not encounter such a case and notice the most benefit in moving from one to two time steps.

Table I shows typical parameters for Algorithm 1. The strength-of-connection parameter, θ, or drop-tolerance, is selected so that smaller values reflect stronger connections. Hence, we use $\theta > 1$ and the row-wise dropping strategy in line 16 of Algorithm 1. While tuning θ for Evolution Measure 1 is potentially beneficial, this approach is less sensitive to small changes in θ than the classic strength measure. We offer empirical evidence in Section 4.4.

There are four important implementation considerations with Algorithm 1. Since $\bar{B}^{\mathrm{T}} Q \bar{B}$ may be singular, a robust factorization of the pseudo-inverse is necessary—e.g. SVD. It is also beneficial to restrict Ω_i to those degrees-of-freedom belonging to the same PDE as i. Linearized elasticity problems in 3D involve three coupled PDEs at each spatial location and this approach reduces by a factor of 3 the number of elements in $\bar{z}$. For the elasticity examples examined in this paper, an unrestricted Ω_i often results in a $\bar{z}$ with many elements—e.g. 70. The reduction in the size of $\bar{z}$ significantly removes error introduced by approximating $\bar{z}$ with only the 6 near null-space vectors in $\bar{B}$.

Restricting Ω_i in this manner and the pointwise approach to strength-of-connection in Algorithm 1 is compatible with the SA code, ML [13], in which we have implemented Algorithm 1. In particular, our approach is compatible with ML's implementation of supernodes [4]. A supernode is a common approach in the SA framework to vector problems such as elasticity. The coarsening of the finest grid is broken up into 2 stages. In the first stage, supernodes are formed by aggregating degrees-of-freedom at the same spatial location together. The second stage then aggregates supernodes using strength-of-connection information. ML defines the strong neighbors of a supernode, i, as any supernode with a degree-of-freedom that is strongly connected to a degree-of-freedom in i.

The complex analogue of line 11 calculates the angle formed by $\widetilde{z}_{\bar{j}}$ and $\bar{z}_{\bar{j}}$. If

$$\operatorname{real}(\widetilde{z}_{\bar{j}})\operatorname{real}(\bar{z}_{\bar{j}})+\operatorname{imag}(\widetilde{z}_{\bar{j}})\operatorname{imag}(\bar{z}_{\bar{j}})<0, \tag{29}$$

then the angle is greater than $\pi/2$ and the connection is weak.

Lastly, using relative approximation error (20) with respect to $\bar{z}_{\bar{j}}$ is a design choice. If

$$\left|\frac{\bar{z}_{\bar{j}}-\widetilde{z}_{\bar{j}}}{\widetilde{z}_{\bar{j}}}\right| \tag{30}$$

is large, the connection is also likely weak. While we did not encounter cases where (20) failed and (30) did not, the most robust algorithm would include both measures.

3.2.3. Analysis of setup complexity. The complexity analysis is simplified by assuming a constant, η, number of nonzeros per row, a symmetric matrix, A, and the recommended two time steps. To derive complexities, we count the number of multiplies for dominant operations.

We first examine the complexity of Algorithm 1 for Evolution Measure 1. Here, only $\bar{z}$ is needed—i.e. z at the nearest neighbors of i in the matrix graph. The complexity of $z \leftarrow (I - \frac{1}{\rho(D^{-1}A)}D^{-1}A)^2\delta_i$ at the nearest neighbors is between $O(\eta)$ and $O(\eta^2)$ per degree-of-freedom, where the precise complexity depends on the structure of the matrix. For $B \in \mathbb{R}^{n,1}$, this is the dominant operation, as (21) is $O(\eta)$ per degree-of-freedom.

For $B \in \mathbb{R}^{n,m}$, we also account for the complexity of finding the minimum to (17)–(18). It is reasonable to assume that the number of near null-space vectors, m, is much smaller than η. With this, the linear solve for finding the minimum of (17)–(18), which is $O((m+1)^3)$, does not dominate the process. The dominant operation for finding the minimum is $\bar{B}^{\mathrm{T}} Q \bar{B}$, where Q is diagonal. $\bar{B}$ is a dense matrix of at most size $\eta \times m$, giving $\bar{B}^{\mathrm{T}} Q \bar{B}$ a complexity of $O(m^2\eta)$. It is likely that $m^2 \leqslant \eta$; thus, the complexity per degree-of-freedom for Algorithm 1 using Evolution Measure 1 is at most $O(\eta^2)$ (for two time steps).

Evolution Measure 2 and the Approximate Inverse Measureare of a higher complexity. Both measures require the $O(\eta^2)$ complete calculation of z in order to compute the A-norm's in

either (14) or line 7 of Algorithm 1. These A-norm's further add to the complexity and are the dominant operations. Each A-norm is $O(\mathrm{nnz}(z)\eta)$, where $\mathrm{nnz}(z)$ is the number of nonzeros in z and $\eta \leqslant \mathrm{nnz}(z) \leqslant \eta^2$. Per degree-of-freedom, these two measures are $O(\eta^3)$ (for two time steps).

As an example, consider the 'Iron Bar' problem from Section 4.3, where linear elasticity in three dimensions is discretized on a regular mesh by parallelpiped elements. For a generic interior degree-of-freedom and two time steps, $m=6$, $\eta=51$, and $\mathrm{nnz}(z)=375$. For Evolution Measure 1, finding the minimum of (17) is $O(m^2\eta/3)$, where $m^2 * \eta/3 = 36 * 17 = 612$. Since there are 3 coupled PDEs at each point, $\eta/3$ is used. Calculating $\bar{z}$ in this case requires 399 multiplies. On the other hand, Evolution Measure 2 and the Approximate Inverse Measure are of higher complexity, $O(\mathrm{nnz}(z)\eta)$, where $\mathrm{nnz}(z) * \eta = 51 * 375 = 19\,125$.

4. RESULTS

4.1. Scalar anisotropic diffusion strength stencils

Generic strength-of-connection stencils are presented using both the Approximate Inverse Measureand the two Evolution Measures for 2D diffusion problems discretized with Q1 elements. The stencils are arranged as in Equation (4), but with the diagonal entry represented as '***.' Connections labeled '*neg*' refer to a negative ratio in line 11 of Algorithm 1. Tables labeled 'Evolution Measure 1' and 'Evolution Measure 2' refer to the options in Algorithm 1. The parameter values in Table I are used, unless otherwise mentioned.

For both 'Evolution Measure 2' and 'Evolution Measure 1' stencils, smaller values reflect stronger connections. For comparison, we also present stencils for the Approximate Inverse Measure, which is similar to Evolution Measure 2, but with the reciprocal of line 7 in Algorithm 1 and a slightly different underlying ODE. Hence, larger values imply a stronger connection. For convenience, all stencils have the strongest connection normalized to one.

Tables III–V examine stencils corresponding to isotropic, vertically aligned anisotropic and rotated anisotropic diffusion. The rotated case is given by,

$$-(c^2+\epsilon s^2)u_{xx} - 2(1-\epsilon)csu_{xy} - (\epsilon c^2 + s^2)u_{yy} = f \tag{31}$$

where $\epsilon = 0.001$, $c = \cos(\pi/4)$, $s = \sin(\pi/4)$ and $\pi/4$ is the angle of rotation.

Table II gives the original matrix stencils for comparison. As expected, the strength measures are isotropic in the isotropic case. In the anisotropic cases, all measures are an improvement over the matrix stencil, as they accurately interpret positive off-diagonal entries in the matrix stencil and the separation between strong and weak connections is clear. In Table IV, strong connections are in the vertical direction. In Table V, strong connections are along the diagonal from the lower-left to the upper-right. For these problems $t_f \leqslant 2\Delta t$ is sufficient for Evolution Measure 1—i.e. only two time steps. Strength information does continue to improve in Table IV for larger values of t_f, though $t_f = 2\Delta t$ clearly indicates appropriate strong connections. The case of grid-aligned

Table II. Original matrix stencils.

Vertical anisotropy			Anisotropy Rot. by $\frac{\pi}{4}$		
−0.167	−0.666	−0.167	0.083	−0.167	−0.417
0.333	1.33	0.333	−0.167	1.33	−0.167
−0.167	−0.666	−0.167	−0.417	−0.167	0.083

Table III. Isotropic—strength stencils.

Measure	$t_f = \Delta t$			$t_f = 2\Delta t$		
Evolution 1	1.00	1.00	1.00	1.41	1.00	1.41
	1.00	***	1.00	1.00	***	1.00
	1.00	1.00	1.00	1.41	1.00	1.41
Evolution 2	2.00	1.00	2.00	1.68	1.00	1.68
	1.00	***	1.00	1.00	***	1.00
	2.00	1.00	2.00	1.68	1.00	1.68

Table IV. Vertical anisotropy—strength stencils.

Measure	$t_f = \Delta t$			$t_f = 2\Delta t$			$t_f = 4\Delta t$		
Evolution 1	5.00	1.00	5.00	11.9	1.00	11.9	155	1.00	155
	neg	***	*neg*	*neg*	***	*neg*	*neg*	***	*neg*
	5.00	1.00	5.00	11.9	1.00	11.9	155	1.00	155
Evolution 2	36.6	1.00	36.6	33.4	1.00	33.4	284	1.00	284
	9.71	***	9.71	25.6	***	25.6	35.7	***	35.7
	36.6	1.00	36.6	33.4	1.00	33.4	284	1.00	284
Approx. Inv.	0.037	1.00	0.037	0.017	1.00	0.017	0.003	1.00	0.003
	0.209	***	0.209	0.165	***	0.165	0.120	***	0.120
	0.037	1.00	0.037	0.017	1.00	0.017	0.003	1.00	0.003

Table V. Anisotropy rotated by $\frac{\pi}{4}$—strength stencils.

Measure	$t_f = \Delta t$			$t_f = 2\Delta t$			$t_f = 4\Delta t$		
Evolution 1	*neg*	3.47	1.00	*neg*	3.48	1.00	*neg*	3.46	1.00
	3.47	***	3.47	3.48	***	3.48	3.46	***	3.46
	1.00	3.47	*neg*	1.00	3.48	*neg*	1.00	3.46	*neg*
Evolution 2	45.0	3.00	1.00	39.6	2.07	1.00	143	1.82	1.00
	3.00	***	3.00	2.07	***	2.07	1.82	***	1.82
	1.00	3.00	45.0	1.00	2.07	39.6	1.00	1.82	143
Approx. Inv.	0.029	0.193	1.00	0.019	0.238	1.00	0.008	0.302	1.00
	0.193	***	0.193	0.238	***	0.238	0.302	***	0.302
	1.00	0.193	0.029	1.00	0.238	0.019	1.00	0.302	0.008

anisotropy is somewhat particular due to large discretization errors in weak directions compared to the analytic ODE in weak directions. When this is not the case as in Table V for nonaligned anisotropy, there is no benefit to large t_f.

An inherent problem of the Approximate Inverse Measure appears in Table V, where the strength-of-connection information begins to degrade for higher degrees. This also happens in the case of vertically aligned anisotropy, but at higher degrees. This phenomenon is due to the fact that as the measure more closely approximates the inverse, the strength information can degrade. It is important that this phenomenon is not observed for Evolution Measure 1 in these test cases. The separation between strong and weak connections for the Approximate Inverse Measure is roughly

the same as those obtained via Evolution Measure 1, though the cost for the Approximate Inverse Measure is higher.

4.2. Scalar convergence study

To study convergence, the following setup is used. Preconditioned conjugate gradient (PCG) is used with a relative residual tolerance of 10^{-8} and preconditioned with a V(1,1) cycle using symmetric Gauss–Seidel for relaxation. To highlight the advantages of the new strength-of-connection measure, we compare classic SA with classic SA enhanced by the energy minimization algorithm for the prolongator from [12]. We test the enhanced classic SA using both the classic strength measure (1b) (denoted by 'Enhanced Prol') and Evolution Measure 1 (denoted by 'Enhanced Prol Evolution Measure 1'). For classic SA, we use the default implementation in [13], which uses standard SA [3, 4]. Each of the three variants of SA takes advantage of the user-provided near null-space vectors, B, when calculating the prolongator. 'Grid l' refers to a level of uniform refinement starting with $l = 1$ as the coarsest, and p refers to the polynomial order of the finite elements.

For Evolution Measure 1, we use the parameter values in Table I. For efficiency, we capped the number of time steps at 3; however, rarely is this encountered. It is notable that for many problems, 1, 2 and 3 steps were all taken, but at different levels in the hierarchy. The drop-tolerance, θ, refers to the classic measure, (1b), if Evolution Measure 1 is not used. In this case, $\theta \in [0, 1]$.

In Table VI, we show PCG iterations for two types of problems. The first type is anisotropic diffusion discretized with Q1 elements for the same problems examined in Section 4.1. The second type is isotropic diffusion, corresponding to $\epsilon = 1$ in (2), discretized on an airfoil enclosed in a circular domain meshed with triangles. The mesh is heavily refined near the surface of the airfoil, as shown in Figure 1(a). The finest '$p = 1$' airfoil problem has 74 000 degrees-of-freedom and the finest '$p = 2$' problem has 296 992 degrees-of-freedom Operator complexities are generally comparable across methods. Operator complexity is the sum of the number of nonzeros in all the operators (not including prolongation or restriction) in a multilevel hierarchy divided by the number of nonzeros in the finest level operator.

'Enhanced Prol Evolution Measure 1' produces the most robust solver. The classic strength measure, (1b), with $\theta = \frac{1}{4}$ used with 'Classic SA' and 'Enhanced Prol' is effective for some problems, but exhibits unacceptable convergence in other situations.

4.3. Multi-dimensional near-nullspace tests

We next examine three isotropic linearized elasticity problems where $B \in \mathbb{R}^{n,6}$ and is equal to the six rigid body modes. Linear elasticity is defined by

$$-\mathrm{div}(\lambda \mathrm{tr}((\nabla \mathbf{u} + \nabla \mathbf{u}^\mathrm{T})/2)I + \mu(\nabla \mathbf{u} + \nabla \mathbf{u}^\mathrm{T})) = f \tag{32}$$

where λ and μ are the Lamé parameters, I is the identity matrix and $\mathrm{tr}(\cdot)$ is the trace function. For all examples, the GetFem++ package [14] is used to discretize the problem. One example is an iron bar attached to the left wall with a downward force applied to the top of the bar. The iron bar is defined on the region $[0, 4] \times [0, 1] \times [0, 1]$ and discretized with parallelpiped elements. The other two examples are for a steel tripod corresponding to the mesh in Figure 1(b). The tripod is discretized using either 4-point or 10-point (quadratic edge shape) tetrahedrons. A downward external force is applied to the top of the tripod.

Table VI. PCG iterations.

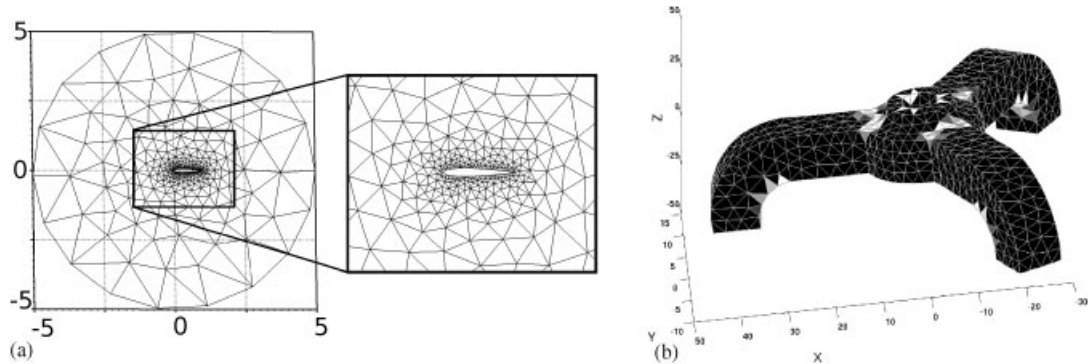|	Classic SA		Enhanced Prol		Enhanced Prol Evolution Measure 1
	$\theta=0$	$\theta=\frac{1}{4}$	$\theta=0$	$\theta=\frac{1}{4}$	$\theta=4.00$
Vertical Ani.					
16×16	23	7	24	21	8
32×32	33	13	33	33	7
64×64	54	18	54	42	10
128×128	72	23	71	49	8
Rot. by $\frac{\pi}{4}$ Ani.					
16×16	10	6	10	7	10
32×32	11	9	11	9	11
64×64	16	13	16	11	12
128×128	21	17	20	14	13
Rot. by $\frac{\pi}{8}$ Ani.					
16×16	13	9	13	10	9
32×32	19	12	16	15	12
64×64	28	15	21	21	15
128×128	38	18	28	28	18
Airfoil Diffusion					
$p=1$ Grid 1	7	8	7	6	7
Grid 2	8	12	9	7	8
Grid 3	9	17	9	8	9
Grid 4	11	24	10	9	10
Grid 5	12	30	11	9	11
$p=2$ Grid 1	13	13	12	8	10
Grid 2	14	19	13	8	10
Grid 3	16	33	13	10	11
Grid 4	20	49	16	13	12
Grid 5	22	88	21	15	14

Figure 1. Example meshes: (a) coarse airfoil mesh and (b) coarse tripod mesh.

In Table VII, 'l_2-Proj.' refers to $Q=I$ and 'D_A-Proj.' refers to $Q=\bar{D}_A$, because of the relationship between the minimization problem (17) and an orthogonal projection. For the iron bar cases, the finest '$p=1$' problem has 52 920 degrees-of-freedom while the finest '$p=2$' problem has 27 744 degrees-of-freedom. For both the *4-pt* and *10-pt* tripod cases, the finest '$p=1$' problem

Table VII. PCG iterations.

|	Classic SA		Enhanced Prol		Enhanced Prol	Evolution Measure 1
					D_A- Proj.	l_2-Proj.
	$\theta=0$	$\theta=\frac{1}{4}$	$\theta=0$	$\theta=\frac{1}{4}$	$\theta=4.00$	$\theta=4.00$
4-pt Tripod						
$p=1$ Grid 1	18	48	15	27	15	13
Grid 2	17	74	14	62	14	13
Grid 3	26	130	20	92	17	17
$p=2$ Grid 1	41	85	24	46	18	17
10-pt Tripod						
$p=1$ Grid 1	21	44	26	33	13	16
Grid 2	45	88	42	67	13	14
Grid 3	63	148	66	129	18	17
$p=2$ Grid 1	49	86	52	61	16	17
Iron Bar						
$p=1$ Grid 1	11	9	9	7	14	11
Grid 2	17	31	11	11	15	17
Grid 3	13	44	10	9	15	14
Grid 4	14	53	10	14	13	12
Grid 5	15	52	12	11	17	18
$p=2$ Grid 1	17	37	12	12	19	23
Grid 2	28	63	16	13	15	16

has 109 551 degrees-of-freedom and the finest '$p=2$' problem has 16 341 degrees-of-freedom Operator complexities were comparable across methods.

The challenging strength-of-connection questions are posed by the matrices from the two tripod test problems, in particular the case of the *10-pt* tetrahedron. For these two test problems, 'Enhanced Prol, Evolution Measure 1' performs well, while no other method produced a scalable solver. 'Enhanced Prol Evolution Measure 1' is comparatively weakest for the iron bar test case; however, this is a problem for which classic SA is designed, and the new method still outperforms 'Classic SA' for $p=2$.

4.4. Drop-tolerance sensitivity

This section compares the sensitivity of the classic strength measure (1b) and Evolution Measure 1 to their respective drop-tolerances. The PyAMG package [15], which implements standard SA [3, 4], was used for this test. In Figure 2, we plot work per digit of accuracy for the stand-alone SA method versus the drop-tolerance, θ. Work per digit of accuracy for the simple V-cycles employed here is the operator complexity divided by $|\log_{10}(\rho)|$, where ρ is the average residual reduction ratio. Denoting the drop-tolerance for the classic strength measure (1b) and Evolution Measure 1 as Classic-θ and Evolution-θ, respectively, we let Classic-$\theta=x$ and Evolution-$\theta=1/x$. In this way, $x=0$ indicates that all connections are strong and $x=1$ indicates that all off-diagonal connections are weak.

The SA methods compared are those used in Tables VI and VII and are applied to the 'Rot. by $\pi/8$ Ani., 128×128 Grid' case (using Table VI nomenclature). The data points end for 'Classic SA' and 'Enhanced Prol' when the drop-tolerance creates an initial coarse grid of all singleton aggregates.
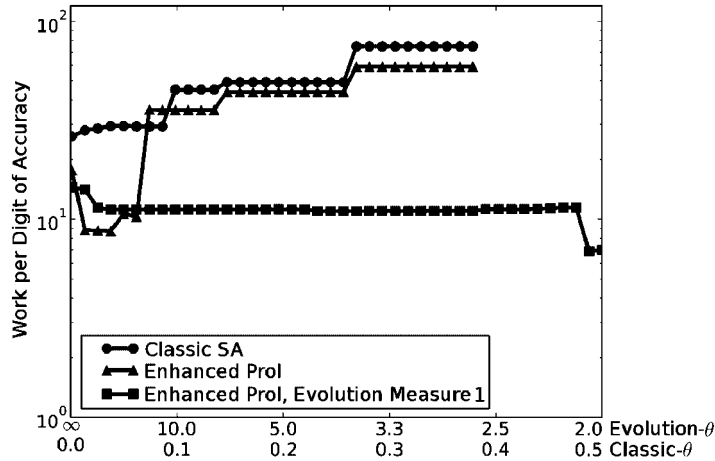
Figure 2. Anisotropic diffusion rotated by $\frac{\pi}{8}$.

As indicated in Figure 2, there is generally a narrow band of θ-values for which the classic strength measure produces a suitable method, while there is a wide range of acceptable θ-values for Evolution Measure 1. This indicates that the new measure is less sensitive and is hence a more parameter-free approach by requiring less tuning of the drop-tolerance.

5. CONCLUSIONS

We presented a novel strength-of-connection measure that explicitly combines the local nature of algebraically smooth error and the local behavior of interpolation. In particular, the use of near null-space modes in order to account for interpolation has, to the knowledge of the authors, not been previously attempted. The new measure is based on an ODE perspective. This perspective provides insight into subtle issues associated with determining strength information and how these issues relate to both classical measures and some newer measures associated with approximating matrix inverses. As discussed, measures based on matrix inverses can degrade as the inverse approximation is improved, due to misleading strength information that matrix inverses can give for even simple problems.

More significantly, the new strength measure is a robust and effective approach to strength-of-connection . For the case of one near null-space mode, Evolution Measure 1 performs as well as the Approximate Inverse Measure. Yet, the cost of Evolution Measure 1 is significantly less than that of the Approximate Inverse Measure, even for multiple near null-space modes. For the SA-based methods compared, Evolution Measure 1 yields the most robust and scalable performance in the numerical experiments. Evolution Measure 1 also appears to be a more parameter-free approach than the classic measure because our experiments indicate less sensitivity to the drop-tolerance.

REFERENCES

1. 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.
2. Ruge JW, Stüben K. Algebraic multigrid (AMG). In *Multigrid Methods*, McCormick SF (ed.). Frontiers in Applied Mathematics. SIAM: Philadelphia, 1987; 73–130.
3. Vaněk P, Mandel J, Brezina M. Algebraic multigrid based on smoothed aggregation for second and fourth order problems. *Computing* 1996; **56**:179–196.
4. Vanek P, Brezina M, Mandel J. Convergence of algebraic multigrid based on smoothed aggregation. *Numerische Mathematik* 2001; **88**(3):559–579.
5. Brannick J, Brezina M, MacLachlan S, Manteuffel T, McCormick S. An energy-based AMG coarsening strategy. *Numerical Linear Algebra with Applications* 2006; **13**:133–148.
6. Bröker O. Parallel multigrid methods using sparse approximate inverses. *Ph.D. Thesis*, Department of Computer Science, ETH Zürich, May 2003.
7. Brandt A. General highly accurate algebraic coarsening. *Electronic Transactions on Numerical Analysis* 2000; **10**:1–20.
8. Livne OE. Coarsening by compatible relaxation. *Numerical Linear Algebra with Applications* 2004; **11**(2–3): 205–227.
9. Brannick J, Zikatanov L. Algebraic multigrid methods based on compatible relaxation and energy minimization. In *Domain Decomposition Methods in Science and Engineering XVI*, Widlund O, Keyes DE (eds). Lecture Notes in Computational Science and Engineering, vol. 55. Springer: Berlin, 2007; 15–26.
10. Adams M, Brezina M, Hu J, Tuminaro R. Parallel multigrid smoothing: polynomial versus Gauss–Seidel. *Journal of Computational Physics* 2003; **188**:593–610.
11. Falgout RD, Vassilevski PS. On generalizing the algebraic multigrid framework. *SIAM Journal on Numerical Analysis* 2004; **42**(4):1669–1693.
12. Mandel J, Brezina M, Vanek P. Energy optimization of algebraic multigrid bases. *Computing* 1999; **62**:205–228.
13. Gee M, Siefert C, Hu J, Tuminaro R, Sala M. ML 5.0 smoothed aggregation user's guide. *Technical Report SAND2006-2649*, Sandia National Laboratories, 2006.
14. Renard Y. The GetFem++ Project. URL: http://download.gna.org/getfem/doc/getfem_project.pdf.
15. Bell WN, Olson LN, Schroder J. Pyamg: algebraic multigrid solvers in python 2008. http://www.pyamg.org, release 1.0 Beta.