

## GENERALIZED STRENGTH OF CONNECTION IN ALGEBRAIC MULTIGRID

JACOB SCHRODER<sup>‡</sup>, RAYMOND S. TUMINARO<sup>§</sup>, AND LUKE OLSON<sup>¶</sup>

**Abstract.** Algebraic multigrid (AMG) solves sparse linear systems without knowledge of any underlying geometric grid. The automatic construction of a multigrid hierarchy requires strength of connection information in order to coarsen the matrix graph and determine sparsity patterns for each intergrid transfer operator. This paper focuses on accessing strength of connection information, i.e. determining which degrees of freedom are strongly related to each other when algebraically smooth error is transferred between grids. Unfortunately, classic strength measures based on matrix stencils can be ineffective due to discretization errors and matrix inverses can be too global. We present an ODE framework for interpreting previous measures and propose a new strength of connection criteria. Some numerical results for the new criteria are also given.

**1. Introduction.** Algebraic multigrid (AMG) solves sparse linear systems without knowledge of any underlying geometric grid. The automatic construction of a multigrid hierarchy normally centers on three distinct tasks: coarse grid selection, determination of the sparsity pattern for each intergrid transfer, and the specification of the actual coefficients within intergrid transfer matrices. This paper focuses on the first two tasks which in turn rely on accessing strength of connection information, i.e. determining which degrees of freedom are strongly related to each other when algebraically smooth error is transferred between grids. Specifically, strength of connection information is used to construct a graph,  $G$ , whose vertices are the degrees of freedom present in the operator,  $A$ , and where  $i$  is connected by an edge to  $j$  only if  $i$  is strongly connected to  $j$ . The coarse grid is then constructed by applying some graph algorithm that coarsens  $G$ . Strength information is also used to construct the intergrid transfer operator, where degree of freedom,  $i$ , is used to interpolate to degree of freedom,  $j$ , only if  $i$  is strong connected to  $j$ .

The current state of strength of connection in AMG is primarily based on the seminal work of the 1980's that developed the classic strength of connection measure. The classic strength of connection measure uses the matrix stencil to determine the strength of connection between two degrees of freedom,  $i$  and  $j$ . For instance in [4],  $i$  is strongly connected to  $j$  with respect to a matrix  $A$  only if

$$-A(i, j) \geq \theta \max_{l \neq i} \{-A(i, l)\}, \quad (1.1a)$$

for some drop tolerance,  $0 < \theta \leq 1.0$ . Similarly, smoothed aggregation [5] sets degree of freedom  $i$  to be strongly connected to degree of freedom  $j$  only if

$$|A(i, j)| \geq \theta \sqrt{A(i, i)A(j, j)}. \quad (1.1b)$$

Unfortunately, the classic measure is most applicable to only M or near-M matrices.

A simple and common example of this measure's limitations can be seen by considering the use of bi-linear finite elements for

$$-u_{xx} + -\epsilon u_{yy} = f \quad (1.2)$$

on a uniform mesh. The corresponding matrix stencil at an interior point is

$$\frac{1}{3} \begin{pmatrix} -\frac{1+\epsilon}{2} & 1-2\epsilon & -\frac{1+\epsilon}{2} \\ -2+\epsilon & 4+4\epsilon & -2+\epsilon \\ -\frac{1+\epsilon}{2} & 1-2\epsilon & -\frac{1+\epsilon}{2} \end{pmatrix}. \quad (1.3a)$$

<sup>‡</sup>Department of Computer Science, University of Illinois at Urbana-Champaign, jschrod3@uiuc.edu

<sup>§</sup>Sandia National Laboratories, rstumin@sandia.gov

<sup>¶</sup>Department of Computer Science, University of Illinois at Urbana-Champaign, lukeo@uiuc.edu

$\epsilon = 1.0$  and  $\epsilon = 0.0$  yield respectively,

$$\frac{1}{3} \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad \text{and} \quad \frac{1}{3} \begin{pmatrix} -\frac{1}{2} & 1 & -\frac{1}{2} \\ -2 & 4 & -2 \\ -\frac{1}{2} & 1 & -\frac{1}{2} \end{pmatrix}. \quad (1.3b)$$

When  $\epsilon$  is small, the coupling in the  $y$ -direction is weak. This means that a standard point smoothing algorithm such as Gauss-Seidel will be ineffective at reducing errors which are smooth in the  $x$ -direction but oscillatory in the  $y$ -direction. This is not a problem if the multigrid coarse mesh is obtained by semi-coarsening, which coarsens only in the directions where the error after relaxation is smooth. Here, this implies coarsening only in the  $x$ -direction. To semi-coarsen, however, the strength of connection measure should determine that the coupling in the vertical direction is weak compared to the coupling in the horizontal direction. Unfortunately, the classic strength of connection measure only indicates modestly stronger coupling in the horizontal direction. Depending on the drop tolerance, the multigrid algorithm may or may not make the proper classification. The simple use of matrix coefficients is not sufficient to reliably reveal connection strength.

Another motivating concept for determining strength of connection has been the matrix inverse. The inverse can at first seem an attractive target for calculating strength of connection because the inverse relates the residual to the error,

$$A^{-1}r = e. \quad (1.4)$$

This relationship can appear useful for determining strength of connection in multigrid, because multigrid solves residual equations on coarse grids. However, the inverse does not necessarily give useful local strength of connection information. The information in the inverse is too global and includes information about both low and high energy modes.

For example, consider a standard 1-D finite difference approximation of

$$-\epsilon(x)u_{xx} = f \quad (1.5)$$

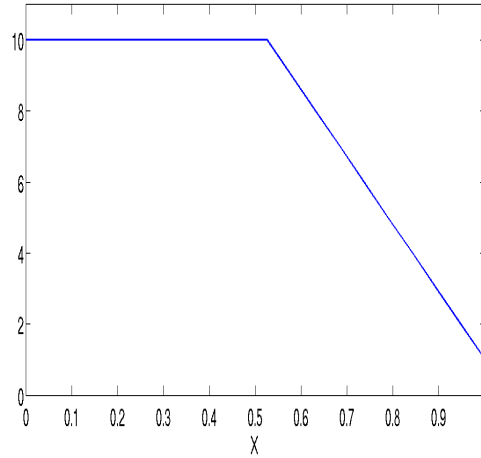
with 20 points on  $[0, 1]$  and  $h = 1/19$ . Define a Neumann boundary condition at  $x = 0.0$  and a Dirichlet boundary condition at  $x = 1.0$ . Let  $\epsilon(x) = 0.001$  if  $x \leq 0.5$  and  $\epsilon(x) = 1.0$  otherwise. The 11th row of the matrix contains the stencil,

$$x = \begin{bmatrix} -0.001 & 1.001 & -1 \\ 10h & 11h & 12h \end{bmatrix}. \quad (1.6)$$

However, the stencil of the matrix inverse for the two nearest neighbors of point 11 is

$$x = \begin{bmatrix} 10.0 & 10.0 & 9.0 \\ 10h & 11h & 12h \end{bmatrix}. \quad (1.7)$$

This is incorrect from a strength standpoint. The strong connection for point 11, should be to the right, in the direction of the large PDE coefficient. Instead, the strength information is inconclusive, and even hints at a slightly stronger connection in the direction of the small PDE coefficient. The reason for this can be explained by considering the Green's function. Since the finite difference stencil for the Neumann boundary condition sums to 0, the Green's function must be constant from point 11 to the Neumann condition on the left. On the other hand, the finite difference stencil for the Dirichlet boundary condition forces the Green's function to be zero at the Dirichlet boundary on the right. The Green's function corresponding to the 11th point is shown in Figure 1.1. It essentially corresponds to two linear functions,

FIG. 1.1. Column 11 of  $A^{-1}$ 

one to the left of point 11 and the other to the right of point 11, whose slopes are chosen to satisfy the boundary conditions, but contain no information about  $\epsilon$ .

If we continue examining this example, we can find even more problems with the inverse. Columns of the inverse corresponding to points to the left and to the right of the interface also give incorrect strength information. Even simpler examples can yield inverses with misleading strength information. Standard finite differencing applied to 1-D isotropic diffusion with both a Neumann and a Dirichlet boundary condition yields an inverse with misleading strength of connection information near the boundary conditions.

One recent idea that has attracted attention is Compatible Relaxation (CR), which is used to identify a subset of the original  $n$  degrees of freedom which will define a coarse mesh. CR iteratively carries out a mock-AMG cycle on  $Ae = 0$ , where  $e$  is an initial random guess. First,  $e$  is smoothed with the multigrid smoother but each point in the tentative set of coarse points is made to be invariant and held to 0. The basic idea is that this models a perfect coarse level operator which reduces the error at the coarse points to zero. Second, CR augments the tentative set of coarse points with a maximal independent set of points that were not sufficiently reduced in  $e$  by the mock-AMG cycle. The maximal independent set is chosen using the graph of the matrix. This process is repeated for  $e$  until convergence is satisfactory. Multiple initial random guesses are tried and the algorithm stops once a good balance has been struck between the coarsening ratio and the convergence rate of the CR smoothing step. While CR does not explicitly make strength of connection decisions, it does make strength related decisions when choosing a coarse grid. However, choosing a coarse grid is essentially an easier problem than calculating strength.

Another strength of connection avenue that has been explored is based on local approximations to the matrix inverse [1, 2]. These methods follow the reasoning that strength of connections within the matrix inverse are the most relevant when determining intergrid transfers. While our examples illustrate that this is incorrect, local approximations to the inverse can be much better than the actual inverse. This is because the local approximation may not suffer from being too global. In this paper, the strength of connection measure in [1] is examined and referred to as the  $\delta$ -function inverse measure. This method uses relaxation to calculate approximations to the matrix inverse with a 0 initial guess. An energy-based post-processing step is then applied to each column of the approximate inverse to determine

strength of connection.

Our central premise is that strength of connection information suitable for a multigrid algorithm is best determined by examining the evolution of an initial Dirac  $\delta$ -function during the standard multigrid relaxation process. Based on a relationship between weighted-Jacobi relaxation and the time marching of ordinary differential equations (ODEs), an ODE perspective is presented for understanding the  $\delta$ -function inverse measure and the evolution of delta functions during relaxation. The ODE perspective is used to shed new light on limitations associated with classical strength of connection measures as well as limitations associated with matrix inverses. In particular, classical strength of connection measures can be viewed as the initial evolution of a  $\delta$ -function within an ODE framework while matrix inverses are more closely tied to steady-state behavior. It is shown that the initial evolution of a  $\delta$ -function may be inaccurate due to high energy modes associated with discretization errors while the steady-state solution can be too global in nature to properly mimic the behavior of relaxation. A closely related modified measure to the one in [1] is then proposed based on the time evolution of a  $\delta$ -function until an intermediate time. A key issue in this proposed method is the determination of an appropriate intermediate time from which to base strength of connection.

In Section 2, an ODE perspective related to CR is presented that mirrors a Jacobi-relaxation type iteration. The ODE perspective provides an additional view on the limitations of classic strength measures and leads into the topic of matrix inverses and the  $\delta$ -function inverse measure. In Section 3, the  $\delta$ -function inverse measure is discussed from the ODE perspective. In Section 4, a new method related to the  $\delta$ -function inverse measure is proposed and analyzed from the ODE perspective. Experimental results are also then given along with a scale invariance proof.

**2. ODE Perspective On Strength.** An ordinary differential equation (ODE) perspective is presented. The perspective provides an additional view on the limitations of classical strength measures and leads into the topic of local approximations to matrix inverses and the delta-function inverse measure.

Consider the following ODE,

$$u_t = -Au, \text{ with } u(0) = \delta_i \quad (2.1)$$

where  $\delta_i$  is a Dirac delta function centered at the location of the  $i$ -th grid point and  $A$  is a symmetric positive definite matrix that is assumed to be diagonally scaled so that its diagonal entries are one. The exact solution to this system is

$$u = e^{-At} \delta_i. \quad (2.2)$$

When  $A$  corresponds to a Laplacian, this ODE models the diffusion in time of an initial point distribution,  $\delta_i$ . Obviously, the steady state solution is just  $u = 0$ .

Numerically, equation (2.1) can be solved by the forward Euler method resulting in an iteration of the form

$$\begin{aligned} u^{(0)} &= \delta_i, \\ u^{(k+1)} &= u^{(k)} - \Delta t A u^{(k)}. \end{aligned} \quad (2.3a)$$

or

$$u^{(k)} = (I - \Delta t A)^k \delta_j. \quad (2.3b)$$

Strength of connection corresponds to how much a point  $i$  influences a point  $j$ . In the context of (2.1), how does the  $\delta$ -function at point  $i$  spread to point  $j$ . At  $t = 0$ , we have

$$u_t = -A\delta_i. \quad (2.4)$$

$A\delta_i$  is simply the matrix coefficients in the  $i$ -th row. That is, the growth of  $u$  at  $j$  for  $t = 0$  is just given by the size of the coefficient  $A(i, j)$ . This is in principle identical to standard strength of connection measures. There is indeed a strong link between the matrix stencil and the time evolution of  $\delta$ -functions at  $t = 0$ .

As mentioned the steady state solution of equation (2.1) is just 0 and yields no useful information. Instead, suppose we consider the following ODE:

$$u_t = -Au + \delta_i, \text{ with } u(0) = 0. \quad (2.5)$$

When  $A$  corresponds to a Laplacian, this ODE models diffusion in time with a constant source applied at  $i$ . This too might provide insight into how  $i$  influences a point  $j$  via the operator  $A$ . At  $t = 0$ , we have  $u_t = \delta_i$ . After one step of a forward Euler scheme, we have  $u_t(\Delta t) = -\Delta t A\delta_i + \delta_i$ . Thus, once again we have the influence of  $i$  at the point  $j$  governed primarily by the size of the matrix coefficients. The steady state solution is now given by  $A^{-1}\delta_i$ , i.e. a column of the inverse of  $A$ . However as already discussed, the matrix inverse can be misleading as it is too global.

Now, consider the transient solution of (1.2) with  $\epsilon = 0$  given by,

$$u_t = -u_{xx} + \delta(x^*, y^*), \quad (2.6)$$

where  $f$  is taken as a  $\delta$ -function centered at some spatial location given by  $(x^*, y^*)$ . The solution of the associated PDE (2.6) properly indicates that there is no coupling in the  $y$ -direction. The  $\delta$ -function only spreads in the  $x$ -direction as time increases. More precisely,  $u(x, y) = 0$  for  $x \neq x^*$  and  $t > 0$ . The associated PDE should give accurate strength of connection information at time  $t = 0$ . A problem can occur, however, when the spatial term is discretized. For example while discretizing the spatial term with Q1 basis functions on a uniform grid, one can intuitively see that the basis functions still interact in the weak direction and will hence yield some discretization error in the weak direction where there is no PDE coupling. This is reflected in the right-most expression of (1.3b) by the top three and bottom three stencil coefficients. We spare the reader the details, but the overall stencil yields  $O(h^2)$  cross derivative truncation error terms, in the case of a sufficiently smooth function  $u$ . Normally, this error contribution is small for smooth functions. However, the  $\delta$ -functions are not smooth and so these error terms are significant during the initial time steps. Since there is no PDE coupling in the  $y$ -direction, any error terms involving derivatives taken in that direction are significant.

Thus in summary, solutions to equations (2.1) and (2.5) can be considered to generate strength of connection information. The solutions at  $t = 0$  for (2.1) and at  $t = \Delta t$  for equation (2.5) are similar to standard AMG strength of connection measures because they reduce to using the matrix stencil to make strength decisions. We know, however, that the use of the matrix stencil is quite limited due to discretization errors in directions of weak coupling within the PDE. Time evolution of the  $\delta$ -function can serve to damp the high frequency discretization errors and result in more accurate strength of connection information.

On the other hand, the  $t \rightarrow \infty$  solutions give us either  $A^{-1}$  or just 0. These solutions are often too global and do not accurately represent the local influence of  $i$  on  $j$ . Instead, we will try and consider an intermediate time where local discretization errors in the weak direction have decayed sufficiently, but where the time is not too large to render the solution global. This will be discussed in Section 3.

**3. Matrix Inverses And The  $\delta$ -Function Inverse Measure.** As discussed, the matrix inverse is often too global to adequately capture strength of connection information. However, there has been some success in using approximate matrix inverses as a means of determining strength of connection in [1] and [2]. The success of these methods hinges on the fact

that the matrix inverse is not actually computed, as this is too expensive. Instead, a local approximation to the matrix inverse is generated and this local approximation is actually better for determining strength of connection than the true matrix inverse. The local nature of the approximation indirectly accounts for the relaxation process used to compute it and ignores distorting boundary effects.

If, for example, we recall (2.5):

$$u_t = -Au + \delta_i, \text{ with } u(0) = 0$$

and solve this system by a forward Euler method. This essentially solves

$$Au = \delta_i \tag{3.1}$$

by a Jacobi scheme. This idea was considered in [1] along with the use of other iterative relaxation schemes.

The approximate column of the inverse was then combined with a particular strength formula involving  $u$  and the  $A$ -norm. This energy-based post-processing step calculates strength of connection between  $i$  and  $j$  by taking the  $i$ -th approximate column of the inverse,  $z_i$ , and evaluating

$$\frac{\|\bar{z}_i\|_A - \|z_i\|_A}{\|z_i\|_A}, \tag{3.2}$$

where  $\bar{z}_i$  is  $z_i$  zeroed out at entry  $j$ . This corresponds to calculating a normalized change in energy for  $z_i$ . Since  $z_i$  is a locally smooth vector, it would make a good interpolation basis function around  $i$ . Hence if one zeroed out an entry,  $j$ , in  $z_i$ , i.e. interpolation from  $j$  is not used, a large change in energy would be expected if  $j$  were important to the interpolation. This indicates a strong connection.

In our experiments with the scheme, we found that it often worked well in practice but produced less accurate strength of connection measures if the iteration was carried on too long. The  $\delta$ -function inverse measure converges to  $y(t) = A^{-1}\delta_i + e^{-At}\delta_i$ . As the iterations count is increased for this method,  $t \rightarrow \infty$  and all information about smooth modes in the exponential is lost. However, for small numbers of iterations, the result should be dominated by the locally smooth modes present in the matrix exponential.

The  $\delta$ -function inverse measure has a number of strengths and worked well in a number of our experiments. It worked as well as a distance-based strength measure on stretched meshes. It avoids any dependence on random starting guesses. Calculation (3.2) experimentally gave useful information, albeit at the cost of  $n$   $A$ -norms.

However, this method is also not without its weaknesses. Iterating  $k$  times to generate the approximate column of the inverse is equivalent to  $kn$  matrix-vector products. The method converges to non-ideal strength information, i.e.  $A^{-1}$ . The method does not indicate when it is appropriate to stop and if one iterates too much, the quality of the strength information degrades.

**4. Proposed Method.** In this section, a new method is proposed by integrating the strengths of CR and the  $\delta$ -function inverse measure while addressing their weaknesses. We give a description for the method using the ODE perspective, show scale invariance and finally give experimental results.

**4.1. Algorithm Description.** The proposed method solves equation (2.1) with Euler's method and  $\delta_i$  as the initial guess. The steady state solution is 0, but the method stops at a "moderate" time,  $t_f$ . The strength information for  $i$  is then in the resulting smoothed vector.

This vector,  $(I - \Delta t D^{-1} A)^k \delta_i$ , where  $D = \text{diag}(A)$  and we have modified (2.3b) to explicitly account for scaling, can be examined directly to determine connection strength or it can be post-processed as in expression (3.2). The only parameters that could be user-defined are  $k$  and  $t_f$ , i.e. the number of time steps and the final time. Here is a simple algorithm that describes our method.

Input:  $A$ : Matrix  
 $t_f$ : Stop time for evolution of  $\delta$ -function  
 $k$ : Time steps for evolution of  $\delta$ -function  
 $drop\text{-}tol$ : Drop tolerance for weak connections  
 $energy$ : Boolean control for post-processing  
 $b$ : Null Space vector that needs to be well approximated on coarse meshes.  $b$  is often taken to be a vector of ones.

Output:  $S$ :  $S(i, j) = i$ 's strength of connection to  $j$   
 $S$  has the same sparsity structure as  $A$

```

for  $i = 1:\text{numRows}$ 
 $z = (I - \frac{t_f}{k} D^{-1} A)^k \delta_i$ 
 $cols = \text{nonzero-pattern}(A(i, :))$ 
for  $j = 1:\text{length}(cols)$ 
if( $energy$ )
%Energy-based Post-processing
 $\bar{z} = z$ 
 $\bar{z}(cols(j)) = 0$ 
 $S(i, cols(j)) = \frac{\|\bar{z}\|_A - \|z\|_A}{\|z\|_A}$ 
else
%No Post-processing
 $S(i, cols(j)) = \frac{z(cols(j))/b(cols(j))}{z(i)/b(i)}$ 
end
Apply( $S(i, :)$ ,  $drop\text{-}tol$ )

```

The calculation  $\frac{z(cols(j))/b(cols(j))}{z(i)/b(i)}$  serves two purposes. One, this calculation ensures scale invariance. Two, the comparison of  $z$  to a null space vector,  $b$ , allows for this method to work for nonconstant null spaces. This calculation measures how well a locally smooth function around  $i$  approximates the null space vector at neighbor  $cols(j)$ . This calculation should accurately determine strength of connection, i.e. how well algebraically smooth error can be interpolated from  $i$  to  $cols(j)$ .

The choice of both  $t_f$  and  $k$  is not entirely clear.  $k$  must at least be chosen so that the iteration is stable. We have, however, found that large  $k$  rarely helps.  $k = 1$  is similar to the classic strength measure in that only matrix coefficients are used. We do find that  $k = 2$  offers significant improvement over  $k = 1$  but that  $k > 2$  does not offer much further improvement. With respect to  $t_f$ , a value too small will result in a measure close to the classic strength measure and is hence undesirable. A value too large will result in information that is too global in scope. Overall, we want a  $t_f$  just large enough to damp discretization error. Also,  $t_f$  and  $k$  should result in an Euler's method whose action is commensurate with the smoother used in the multigrid cycle. Such a choice that meets these constraints and that worked well in our experiments, let  $t_f = \frac{1}{\rho(D^{-1}A)}$  and  $k = 2$ .

**4.2. Scale Invariance.** As with the  $\delta$ -function inverse measure, our method is invariant with respect to an arbitrary symmetric diagonal scaling.

Proof. Let  $\tilde{A} = \hat{D}^{-1/2} A \hat{D}^{-1/2}$ , for an arbitrary nonsingular diagonal matrix  $\hat{D}$ ,  $D = \text{diag}(A)$

and  $\widetilde{D} = \text{diag}(\widetilde{A})$ . Note that  $\widetilde{D} = \hat{D}^{-1}D$ . We first smooth  $\delta_i$  an arbitrary number of times with  $\widetilde{A}$  in order to derive a relationship to smoothing with  $A$ .

$$(I - \omega \widetilde{D}^{-1} \widetilde{A})^k \delta_i = (I - \omega \widetilde{D}^{-1} \widetilde{A}) \hat{D}^{1/2} \hat{D}^{-1/2} (I - \omega \widetilde{D}^{-1} \widetilde{A}) \hat{D}^{1/2} \hat{D}^{-1/2} \dots \quad (4.1a)$$

$$\begin{aligned} & (I - \omega \widetilde{D}^{-1} \widetilde{A}) \hat{D}^{1/2} \hat{D}^{-1/2} \delta_i \\ &= (\hat{D}^{1/2} - \omega D^{-1} \hat{D}^{1/2} A) \hat{D}^{-1/2} (\hat{D}^{1/2} - \omega D^{-1} \hat{D}^{1/2} A) \hat{D}^{-1/2} \dots \\ & (\hat{D}^{1/2} - \omega D^{-1} \hat{D}^{1/2} A) \hat{D}^{-1/2} \delta_i \end{aligned} \quad (4.1b)$$

$$= \hat{D}^{1/2} (I - \omega D^{-1} A) (I - \omega D^{-1} A) \dots (I - \omega D^{-1} A) \hat{D}^{-1/2} \delta_i \quad (4.1c)$$

$$= \hat{D}_{(i,i)}^{-1/2} \hat{D}^{1/2} (I - \omega D^{-1} A)^k \delta_i, \quad (4.1d)$$

where  $\hat{D}_{(i,i)}^{-1/2}$  is a scalar equal to the  $i$ -th diagonal entry. Let  $z_i = (I - \omega D^{-1} A)^k \delta_i$ ,  $\widetilde{z}_i = \hat{D}_{(i,i)}^{-1/2} \hat{D}^{1/2} (I - \omega D^{-1} A)^k \delta_i$ , and  $\widetilde{b} = \hat{D}^{1/2} b$  be the null space vector for  $\widetilde{A}$ .

We first consider the case with no post-processing.

$$S(i, j) = \frac{\frac{\widetilde{z}(j)}{\widetilde{b}(j)}}{\frac{\widetilde{z}(i)}{\widetilde{b}(i)}} = \frac{\frac{\hat{D}_{(i,i)}^{-1/2} \hat{D}_{(j,j)}^{1/2} z(j)}{\hat{D}_{(j,j)}^{1/2} b(j)}}{\frac{\hat{D}_{(i,i)}^{-1/2} \hat{D}_{(i,i)}^{1/2} z(i)}{\hat{D}_{(i,i)}^{1/2} b(i)}} = \frac{z(j)}{b(j)} \quad (4.1e)$$

Now, consider the case of energy-based post-processing. Let  $\widetilde{z} = \widetilde{z}$  but with the entry corresponding to neighbor  $j$  zeroed out.

$$S(i, j) = \frac{\|\widetilde{z}_i\|_{\widetilde{A}} - \|\widetilde{z}_i\|_{\widetilde{A}}}{\|\widetilde{z}_i\|_{\widetilde{A}}} = \frac{\|\widetilde{z}_i\|_{\widetilde{A}}}{\|\widetilde{z}_i\|_{\widetilde{A}}} - 1 \quad (4.1f)$$

$$= \frac{\|\hat{D}_{(i,i)}^{-1/2} \hat{D}^{1/2} \widetilde{z}_i\|_{\widetilde{A}}}{\|\hat{D}_{(i,i)}^{-1/2} \hat{D}^{1/2} z_i\|_{\widetilde{A}}} - 1 \quad (4.1g)$$

$$= \frac{|\hat{D}_{(i,i)}^{-1/2}|}{|\hat{D}_{(i,i)}^{-1/2}|} \frac{\widetilde{z}_i^T \hat{D}^{1/2} \hat{D}^{-1/2} A \hat{D}^{-1/2} \hat{D}^{1/2} \widetilde{z}_i}{z_i^T \hat{D}^{1/2} \hat{D}^{-1/2} A \hat{D}^{-1/2} \hat{D}^{1/2} z_i} - 1 \quad (4.1h)$$

$$= \frac{\|\widetilde{z}_i\|_A - \|z_i\|_A}{\|z_i\|_A} \quad \square \quad (4.1i)$$

**4.3. Experiments.** We implemented our algorithm as part of the existing ML smoothed aggregation framework and implemented the energy minimization algorithm of [3] for prolongator generation. In the below tables, “Energy-based Post-processing” and “No Post-processing” refer to options in the algorithm of Section 4.1 and “Steps” refers to the number of time steps used.  $t_f = \frac{1}{\rho(D^{-1}A)}$ , unless otherwise mentioned.

All of the below strength stencils are for the degree of freedom in the center of a  $31 \times 31$  regular mesh. The data is presented as  $3 \times 3$  arrays of values that represent the strength of connection values between the degree of freedom in question and its geometric neighbors above, below, to the left, to the right and diagonally offset. The degree of freedom in question is represented as “\*\*\*”, as no strength of connection information is needed between a point and itself.

First, we briefly examine the isotropic case with Q-1 elements on a uniform grid. As expected, the strength stencils are isotropic. In Tables 4.1 and 4.2, we present the “No Post-processing” and “Energy-based Post-processing” cases respectively.

TABLE 4.1  
*Isotropic – Strength of Connection Stencils – No Post-processing*

Steps =	1			3		
	0.0836	0.0836	0.0836	0.0547	0.0583	0.0547
Stencils	0.0836	***	0.0836	0.0583	***	0.0583
	0.0836	0.0836	0.0836	0.0547	0.0583	0.0547

TABLE 4.2  
*Isotropic – Strength of Connection Stencils – Energy-based Post-processing*

Steps =	1			3		
	0.0190	0.0381	0.0190	0.0141	0.0183	0.0141
Stencils	0.0381	***	0.0381	0.0183	***	0.0183
	0.0190	0.0381	0.0190	0.0141	0.0183	0.0141

Next, we again examine results for Q-1 elements on a uniform grid, but with anisotropy that is rotated by  $\frac{\pi}{4}$  and vertically aligned anisotropy, corresponding to

$$-(c^2 + \epsilon s^2)u_{xx} - 2(1 - \epsilon)cs u_{xy} - (\epsilon c^2 + s^2)u_{yy} = f, \quad (4.2)$$

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

In Table 4.3, we show the matrix stencils, which can be compared with the computed strength measures. In Tables 4.4–4.7, we present strength stencils for the “No Post-processing” and “Energy-based Post-processing” options for the vertically aligned anisotropy case and then the rotated anisotropy case. For the vertically aligned case,  $t_f = \frac{2}{\rho(D^{-1}A)}$ .

TABLE 4.3  
*Original Matrix Stencils*

Problem	Vertical Anisotropy			Anisotropy Rot. By $\frac{\pi}{4}$		
	-0.1668	-0.6663	-0.1668	0.0829	-0.1668	-0.4166
Stencils	0.3326	1.3346	0.3326	-0.1668	1.3346	-0.1668
	-0.1668	-0.6663	-0.1668	-0.4166	-0.1668	0.0829

TABLE 4.4  
*Vertical Anisotropy – Strength of Connection Stencils – No Post-processing*

Steps =	1			2		
	0.0838	0.3345	0.0838	0.0278	0.2085	0.0278
Stencils	-0.1670	***	-0.1670	-0.0830	***	-0.0830
	0.0838	0.3345	0.0838	0.0278	0.2085	0.0278
Steps =	3			4		
	0.0257	0.1951	0.0257	0.0245	0.1889	0.0245
Stencils	-0.0772	***	-0.0772	-0.0743	***	-0.0743
	0.0257	0.1951	0.0257	0.0245	0.1889	0.0245

It is important that the range for appropriate drop tolerances is large. Weak connections are defined to be less than the drop tolerance times the largest strength of connection value for that degree of freedom. For instance in Table 4.7, the strong connections are along the diagonal from the lower-left to the upper-right and these connections are a factor of 4 greater

TABLE 4.5  
Vertical Anisotropy – Strength of Connection Stencils – Energy-Based Post-processing

Steps =	1			2		
Stencils	-0.0512	-0.1082	-0.0512	-0.0065	0.2157	-0.0065
	-0.1057	***	-0.1057	0.0084	***	0.0084
	-0.0512	-0.1082	-0.0512	-0.0065	0.2157	-0.0065
Steps =	3			4		
Stencils	-0.0037	0.2061	-0.0037	-0.0026	0.2002	-0.0026
	0.0146	***	0.0146	0.0166	***	0.0166
	-0.0037	0.2061	-0.0037	-0.0026	0.2002	-0.0026

TABLE 4.6  
Anisotropy Rotated by  $\frac{\pi}{4}$  – Strength of Connection Stencils – No Post-processing

Steps =	1			2		
Stencils	-0.0347	0.0698	0.1742	-0.0226	0.0552	0.1280
	0.0698	***	0.0698	0.0552	***	0.0552
	0.1742	0.0698	-0.0347	0.1280	0.0552	-0.0226
Steps =	3			4		
Stencils	-0.0205	0.0520	0.1190	-0.0196	0.0506	0.1151
	0.0520	***	0.0520	0.0506	***	0.0506
	0.1190	0.0520	-0.0205	0.1151	0.0506	-0.0196

TABLE 4.7  
Anisotropy Rotated by  $\frac{\pi}{4}$  – Strength of Connection Stencils – Energy-Based Post-processing

Steps =	1			2		
Stencils	-0.0019	0.0287	0.0861	0.0011	0.0181	0.0731
	0.0287	***	0.0287	0.0181	***	0.0181
	0.0861	0.0287	-0.0019	0.0731	0.0181	0.0011
Steps =	3			4		
Stencils	0.0012	0.0161	0.0669	0.0012	0.0152	0.0642
	0.0161	***	0.0161	0.0152	***	0.0152
	0.0669	0.0161	0.0012	0.0642	0.0152	0.0012

than the next strongest connection. Hence a drop tolerance greater than 0.25 would be appropriate. In Table 4.4, the strong connections are in the vertical direction and are a factor of 7-8 greater than the next strongest connections, which imply a drop tolerance greater than  $\frac{1}{7}$  is appropriate.

It is not entirely clear how to interpret the negative entries, but they most likely imply a weak connection. Also, it is important that the separation between weak and strong connections does not change much after 2 time steps. We therefore suggest using only 2 time steps with this method.

As a means of comparison, strength of connection information from the  $\delta$ -function inverse measure is given in Tables 4.8 and 4.9. One of the inherent problems of the  $\delta$ -function inverse measure appears in Table 4.9, where the strength of connection information begins to degrade for higher degrees. This also happens in the vertically aligned anisotropy case, but at higher degrees. It is important that this phenomenon was not observed in our method. The separation between strong and weak connections in Tables 4.8 and 4.9 and is roughly the

same as in our method.

TABLE 4.8  
Vertical Anisotropy – Strength of Connection Stencils –  $\delta$ -function inverse measure

Steps =	1			2		
		0.0020	0.0541	0.0020	0.0017	0.0977
Stencils	0.0113	***	0.0113	0.0161	***	0.0161
	0.0020	0.0541	0.0020	0.0017	0.0977	0.0017
Steps =	3			4		
		0.0011	0.1359	0.0011	0.0006	0.1710
Stencils	0.0187	***	0.0187	0.0205	***	0.0205
	0.0011	0.1359	0.0011	0.0006	0.1710	0.0006

TABLE 4.9  
Anisotropy Rotated by  $\frac{\pi}{4}$  – Strength of Connection Stencils –  $\delta$ -function inverse measure

Steps =	1			2		
		0.0011	0.0073	0.0378	0.0012	0.0151
Stencils	0.0073	***	0.0073	0.0151	***	0.0151
	0.0378	0.0073	0.0011	0.0634	0.0151	0.0012
Steps =	3			4		
		0.0010	0.0230	0.0840	0.0008	0.0308
Stencils	0.0230	***	0.0230	0.0308	***	0.0308
	0.0840	0.0230	0.0010	0.1021	0.0308	0.0008

If appropriate drop tolerance values are chosen, the above strength stencils will yield correct coarse grids. With correct coarse grids, AMG can be used as an effective preconditioner as evidenced in Table 4.10, where an AMG method was used in conjunction with the strength of connection measures computed here.

TABLE 4.10  
PCG Convergence Counts

Problem		Vertical Ani.	Rot. By $\frac{\pi}{4}$ Ani.	Rot. By $\frac{\pi}{8}$ Ani.
Mesh Size	$31 \times 31$	12	9	12
	$63 \times 63$	16	10	16
	$127 \times 127$	17	14	20

**5. Conclusions.** The proposed method performs as well as the  $\delta$ -function inverse measure if the same number of iterations and energy-based postprocessing are both used. The proposed method can also be computationally much cheaper than the  $\delta$ -function inverse measure, especially if no energy-based post processing is used and the number of time steps is 2. In this case, the proposed method only calculates the entries of  $(I - \Delta t A)^2 \delta_i$  for the neighbors of  $i$  in the matrix graph. The entire vector need not be calculated. However in the  $\delta$ -function inverse measure, the entire vector must be calculated so that the energy-based post-processing step can be applied. While the strength information produced by only 2 time steps is inferior to that produced by the  $\delta$ -function inverse measure using 2 time steps and energy-based post-processing, it is much more computationally feasible and is an improvement over just the matrix stencil.

As the iterations of the  $\delta$ -function inverse measure increase,  $t_f \rightarrow \infty$ , and the method converges to the matrix inverse, which is problematic. On the other hand, as the iterations of the proposed method increase,  $\Delta t \rightarrow 0$ , and the method converges to  $e^{-At_f} \delta_i$ , which is a useful combination of locally smooth modes for “moderate”  $t_f$ .

The proposed method is similar to CR in that we relax an initial error for the homogeneous system of equations and our strength decisions are based directly on the action of the smoother. However, we avoid the use of an undetermined number of random starting vectors by choosing point-sources as our initial error.

#### REFERENCES

- [1] J. BRANNICK, M. BREZINA, S. MACLACHLAN, T. MANTEUFFEL, AND S. MCCORMICK, *An energy-based AMG coarsening strategy*, Numer. Linear Algebra Appl., 13 (2006), pp. 133–148.
- [2] O. BRÖKER, *Parallel Multigrid Methods Using Sparse Approximate Inverses*, PhD thesis, Dept. of Computer Science, ETH Zürich, May 2003.
- [3] J. MANDEL, M. BREZINA, AND P. VANEK, *Energy optimization of algebraic multigrid bases*, Computing, 62 (1999), pp. 205–228.
- [4] J. W. RUGE AND K. STÜBEN, *Algebraic multigrid (AMG)*, in Multigrid Methods, S. F. McCormick, ed., Frontiers Appl. Math., SIAM, Philadelphia, 1987, pp. 73–130.
- [5] P. VANĚK, J. MANDEL, AND M. BREZINA, *Algebraic multigrid based on smoothed aggregation for second and fourth order problems*, Computing, 56 (1996), pp. 179–196.