# An Algebraic Multigrid Method Based on Matching in Graphs

James Brannick[1], Yao Chen[1], Johannes Kraus[2], and Ludmil Zikatanov[1]

[1] Department of Mathematics, The Pennsylvania State University, PA 16802, USA; Email: {brannick|chen_y|ltz}@math.psu.edu.
[2] Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria; Email: johannes.kraus@oeaw.ac.at.

## 1 Introduction

We present an Algebraic Multigrid (AMG) method for graph Laplacian problems. The coarse graphs are constructed recursively by pair-wise aggregation, or matching as in [3] and we use an Algebraic Multilevel Iterations (AMLI) [1, 6] for the solution phase.

The two-level method constructs a splitting of the underlying vector space into two subspaces $V_S$ and $V_P$ and then corrects the error successively on $V_S$ and $V_P$. The coarse space $V_P$ is obtained using matching on the underlying graph. Such a two-level method is shown to be uniformly convergent. In the AMLI method (multilevel), $m$ coarse level corrections are applied on each level. For large $m$, while the convergence rate of the method is comparable to that of the two-level method and, hence, uniformly convergent, it is clear that the overall complexity of such method could be too high for large values of $m$. In our approach, the AMLI convergence rate is estimated solely based on the underlying two-level method, which allows us to show that $m = 2$ gives a balance between the complexity and the desired convergence rate, thus, resulting in an efficient algorithm.

The paper is organized as follows. In Sect. 2 the graph Laplacian problem is described. In Sect. 3, the graph matching algorithm is introduced and it is indicated that the $\ell_2$ projection on the coarse space is the key quantity for obtaining the multilevel estimates of the AMLI method. In Sect. 4, an analysis of a specific two-level method is presented and in Sect. 5 its convergence and complexity are estimated. In the following section, numerical results are reported.

## 2 Graph Laplacian Problems

Graph Laplacian solvers can be used as preconditioners for various discrete numerical models, e.g., ones arising from discretizations of partial differential equations,

R. Bank et al. (eds.), *Domain Decomposition Methods in Science and Engineering XX*, Lecture Notes in Computational Science and Engineering 91,
DOI 10.1007/978-3-642-35275-1_15, 

machine learning algorithms, and spectral clustering of images. Consider a connected unweighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ and $\mathcal{E}$ are the sets of vertices and edges. The graph Laplacian $A \in \mathbb{R}^{n\times n}$, where $n = |\mathcal{V}|$ (cardinality of $\mathcal{V}$), corresponding to the graph $\mathcal{G}$, can be defined as follows:

$$(Au, v) = \sum_{k=(i,j)\in\mathcal{E}} (u_i - u_j)(v_i - v_j).$$

The matrix $A$ is symmetric and positive semi-definite. The null space of $A$ is one dimensional, and its basis is given by $\{\mathbf{1}\}$, where $\mathbf{1}$ is a vector whose components are all equal to 1. Our aim here is to solve graph Laplacian problems, or to find $u$, such that $(u, \mathbf{1}) = 0$ and

$$Au = f,$$

for a given $f$ satisfying $(f, \mathbf{1}) = 0$.

We want to find an AMG method to solve graph Laplacians with simple settings, so that we can estimate the performance of the AMG method, with as few assumptions introduced as possible. The construction of this AMG method can also help us to derive similar methods for weighted graph Laplacian problems, which come from finite element or finite difference discretizations of elliptic partial differential equations, circuit simulations, and in general, network flow simulations.

## 3 Graph Matching

Given a graph $\mathcal{G}$, assume that we can find a set of aggregates $\mathcal{M}$ called a *matching*, where each aggregate contains exactly two vertices, and every vertex of $\mathcal{G}$ is contained in exactly one aggregate. For a certain aggregate that contains vertices $i$ and $j$, we merge the two vertices, and the newly formed vertex, named $k$, is considered connected to the vertex $l$ if and only if $l$ is connected to $i$ or $j$ on graph $\mathcal{G}$. By merging vertices in each aggregate, a reduced graph of the graph $\mathcal{G}$ is formed. Applying such a matching algorithm recursively will result in a sequence of graphs. We then construct a solver for the graph Laplacian of $\mathcal{G}$ based on the sequence of reduced graphs.

In the matching $\mathcal{M}$, we consider the $k$-th aggregate as a graph $\mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$. Let $Q$ be the $\ell_2$-orthogonal projection on the coarse space, which consists of vectors that are piecewise constant on each set $\mathcal{V}_k$. An alternative definition of $Q$ is as follows.

$$(Qu)_i = \frac{1}{|\mathcal{V}_k|} \sum_{j\in\mathcal{V}_k} u_j, \quad i \in \mathcal{V}_k.$$

Classical AMG theory suggests that the coarse space should cover, or approximate algebraically smooth error components. Detailed explanations can be found, e.g., in the appendix of [5]. In the following section, we will compute how well piecewise constant vectors can approximate smooth vectors and will discuss the properties of two-level and multilevel methods using the subspace(s) associated with the projection $Q$.

## 4 A Two-Level Method

Define matrices $P$ and $S$ for a given matching $\mathscr{M}$, such that

$$P \cdot e_k = e_i + e_j, \quad S \cdot e_k = e_i - e_j, \quad (i,j) \in \mathscr{V}_k,$$

where $e_i$ and $e_j$ are Euclidean basis vectors. Since a prerequisite for designing an efficient AMLI method is an efficient two-level method, in this section we focus on two-level methods and their convergence rates. Given an initial guess $u_0$, a typical two-level algorithm which takes as input $u_k$ and returns the next iterate $u_{k+1}$ is as follows:

1. $v = u_k + SR^{-1}S^T(f - Au_k)$,
2. $w = v + PA_c^{-1}P^T(f - Av)$,
3. $u_{k+1} = w + SR^{-T}S^T(f - Aw)$.

Here the matrix $R$ is a preconditioner of $S^TAS$, which is the restriction of $A$ on the space $\text{range}(S) = [\text{range}(P)]^{\perp}$. The matrix $A_c$ is an approximation of the restriction of $A$ on the coarse space $V_c = \text{range}(P)$. In our algorithm, $A_c$ is first defined as the graph Laplacian of the unweighted coarse graph and thus $A_c \neq P^TAP$. We then scale $A_c$ such that $(v^TA_cv)/(v^TP^TAPv) \in [1, c_c]$. A proper scaling results in $c_c = 2$ for $P$ that corresponds to an aligned matching and $A$ that is a structured grid of any dimension. The matrix representation of this two-level method, denoted by $G$, can be deduced via the error propagation matrix given as follows.

$$E = (I - SR^{-T}S^TA)(I - PA_c^{-1}P^TA)(I - SR^{-1}S^TA) = I - G^{-1}A. \tag{1}$$

We now derive an estimate on the angle between the spaces $\text{range}(S)$ and $\text{range}(P)$, which in our setting amounts to obtaining a bound on the energy norm of $Q$, the $\ell_2$-orthogonal projection onto $\text{range}(P)$. Let $\gamma$ be the C.B.S. constant such that it is the smallest number satisfying $(Sw, Pv)_A \leq \gamma |Sw|_A |Pv|_A$, then (cf. [6, Corollary 3.7]):

$$|Q|_A^2 = 1/(1 - \gamma^2).$$

Using [2, Theorem 4.2] we can show that, if the symmetrized smoother $\widetilde{R} = R + R^T - S^TAS$ is positive definite, and $(w^T\widetilde{R}w)/(w^TS^TASw) \in [1, \kappa_s]$, then

$$\frac{v^TGv}{v^TAv} \in [1, |Q|_A^2(\kappa_s + c_c - 1)].$$

If a two-level method using a certain matching is already given, then both $|Q|_A$ and $\kappa_s$ can be estimated using the properties of the underlying graph. The norm $|Q|_A$ is estimated as follows:

$$u^TQAQu = \sum_{(i,j)\in\mathscr{E}} ((Qu)_i - (Qu)_j)^2 \leq 2d \sum_{(i,j)\in E} (u_i - u_j)^2 \leq (2d)u^TAu$$

where $d$ is the maximum degree of the graph. This implies that $|Q|_A^2 \leq 2d$. Assuming that the matching $\mathscr{M}$ is perfect, we show that the smallest eigenvalue of $S^TAS$ is larger or equal to 4, by computing

$$w^T S^T A S w \geq \sum_{(i,j)\in\mathscr{M}} ((Sw)_i - (Sw)_j)^2 = \sum_{(i,j)\in\mathscr{M}} 4(Sw)_i^2 = 4\|w\|_{\ell_2}^2.$$

According to the Gershgorin theorem, the largest eigenvalue of $S^T AS$ is bounded by a function of $d$ and for a simple smoother $R$, such as Richardson iteration, $\kappa_s$ is also bounded by a function of $d$. From the above results (i.e, the stability estimate of $Q$ in the $A$-seminorm and the lower bound on the smallest eigenvalue of $S^T AS$) it follows that the two-level method is uniformly convergent with respect to the size of the matrix $A$. Based on the two-level convergence estimate, AMLI cycles with low complexity and predictable convergence is then constructed.

## 5 Algebraic Multilevel Iterations

An estimate of the two-level convergence rate does not automatically carry over to an estimate of the convergence of a multilevel V-cycle, and in general, for piece-wise constant coarse spaces, it can be shown that the convergence rate degrades exponentially with respect to the number of levels. A remedy for this issue is to use more complicated cycles such as AMLI, and keep a balance between complexity of a cycle and its convergence rate so that the resulting algorithm is optimal or nearly optimal.

We describe an AMLI method by first rewriting the two-level preconditioner $G$, as well as $\widehat{G}$ which is $G$ under the hierarchical basis $(S,P)$, in block form:

$$\widehat{G}^{-1} = \widehat{L}^{-T} \begin{pmatrix} (R+R^T - S^T AS)^{-1} & 0 \\ 0 & A_c^{-1} \end{pmatrix} \widehat{L}^{-1},$$
$$G = (S,P)^{-1} \widehat{G} (S,P)^{-T},$$

where

$$\widehat{L} = \begin{pmatrix} I & 0 \\ P^T ASR^{-1} & I \end{pmatrix}.$$

Then define an AMLI preconditioner $B$ as follows.

$$\widehat{B}^{-1} = \widehat{L}^{-T} \begin{pmatrix} (R+R^T - S^T AS)^{-1} & 0 \\ 0 & B_c^{-1} q(A_c B_c^{-1}) \end{pmatrix} \widehat{L}^{-1},$$
$$B^{-1} = (S,P)^T \widehat{B}^{-1} (S,P).$$

Here $A_c$ is the scaled unweighted graph Laplacian of the coarse graph and $B_c$ is a preconditioner of $A_c$, and $q(t)$ is a polynomial. When $q(t) = 1$, the action $\widehat{B}^{-1}$ stands for a V-cycle with an inexact solver $B_c^{-1}$ on the coarse level. In the case of a W-cycle, we have $q(t) = 2 - t$.

The following lemma shows how well the AMLI preconditioner $B$ approximates the two-level preconditioner $G$.

**Lemma 1.** *If $\lambda_1 \leq \lambda(B_c^{-1} A_c) \leq \lambda_2$ and $tq(t) > 0$ for $t \in [\lambda_1, \lambda_2]$, then*

$$\min(1, \min_{\lambda_1 \leq t \leq \lambda_2} \frac{1}{tq(t)}) \leq \frac{v^T G^{-1} v}{v^T B^{-1} v} \leq \max(1, \max_{\lambda_1 \leq t \leq \lambda_2} \frac{1}{tq(t)}).$$

This lemma suggests that, the AMLI method is spectrally equivalent to a two-level method, given that the coarse-level preconditioner is spectrally equivalent to the coarser-level matrix. The upper and lower bounds in the lemma above are related to estimates on $|tq(t)|$ for $t$ in a given interval. As shown in [1, 6], using higher order polynomials $q(t)$, the matrix $B^{-1}$ can approximate $G^{-1}$ arbitrarily well and thus we will have a method with excellent convergence rate. However, a higher order polynomial $q(t)$ leads to a much more expensive computation of the coarser level correction, and the resulting multilevel methods can have a very high complexity and one should be careful in the choice of the polynomial degree.

Assume that a multilevel hierarchy is formed by a recursive application of the matching algorithm. Denote the graph Laplacians on each level, and the corresponding two-level preconditioners by $A_k$ and $G_k$. Following the ordering of levels in [1, 6] we set $A = A_0$ and denote by $A_J$ the coarsest matrix. Define a sequence of solvers as

$$\begin{aligned}
\widehat{B}_J^{-1} &= \widehat{A}_J^{\dagger} = (S_J, P_J)^{-T} A_J^{\dagger} (S_J, P_J)^{-1}, \\
B_k^{-1} &= (S_k, P_k)^T \widehat{B}_k^{-1} (S_k, P_k), \quad k = 0, \ldots, J, \\
\widehat{B}_k^{-1} &= \widehat{L}_k^{-T} \begin{pmatrix} (R_k + R_k^T - S_k^T A_k S_k)^{-1} & 0 \\ 0 & B_{k+1}^{-1} q(A_{k+1} B_{k+1}^{-1}) \end{pmatrix} L_k^{-1}, \quad k = 0 \ldots J-1.
\end{aligned}$$

Then, a multilevel proof of convergence follows.

**Lemma 2.** *Assume that there is a constant $c_g$, $1 \le c_g < 4$, such that the following relation holds.*

$$v^T \widehat{A}_k v \le v^T \widehat{G}_k v \le c_g v^T \widehat{A}_k v, \quad \forall v \text{ and } k = 0, \ldots, J.$$

*Then there exists a linear function $q(t)$, such that*

$$\frac{2}{\sqrt{c_g}} - 1 \le \frac{v^T B_k^{-1} v}{v^T A_k^{-1} v} \le 1, \quad \forall v \text{ and } k = 0, \ldots, J.$$

Here $q(t)$ is a scaled and shifted Chebyshev type polynomial (see [1]).

This lemma shows that, if $c_g$ is strictly less than 4, then the action $B_0^{-1}$ is an uniformly convergent AMLI cycle with $O(n \log n)$ complexity. Even if $c_g = 4$ on all levels, one may prove that the condition number of $B_J^{-1} A_J$ for the case of second order $q(t)$ (similar to a W-cycle) grows linearly with respect to the number of levels $J = \log n$. This results in a convergence factor $1 - 1/\log n$ at a complexity of $O(n \log n)$ for each cycle.

The two-level method we suggest is based on graph matching, thus $c_g \le |Q|_A^2 (\kappa_s + c_c - 1)$. In a simple case where the graph $\mathscr{G}$ is a two-dimensional uniform grid, an aligned regular matching yields $|Q|_A^2 \le 2$, $\kappa_s = 1 + \varepsilon$ for arbitrary small $\varepsilon$, and $c_c \le 2$.

This yields $c_g \le 4$ and thus the W-cycle AMLI preconditioner will result in a nearly optimal order method (cf. Lemma 2 and the discussion below). For unstructured or higher dimensional grids, numerical experiments indicate that random matching may still result in two-level methods for which $c_g \le 4$.

## 6 Numerical Results

We use the matching based AMLI method to solve a family of unweighted graph Laplacians, corresponding to graphs that represent structured grids or unstructured triangulations.

**Structured grids.** In the structured grid case on a rectangular domain, we match in a fixed direction. After several levels of matching the graph corresponding to the coarsest grid is a line. For the test on L-shaped domain, we still use matching in a fixed direction until a part of the coarsest graph becomes a tree. In such case, the unknowns can be ordered so that the fill-in during LU factorization on the coarsest grid is small.

A similar strategy can be used for graph Laplacians corresponding to three-dimensional structured grids. The matching procedure is applied only in two fixed directions.

Convergence analysis indicates that, choosing as a smoother $R^{-1} = (S^T AS)^{-1}$ guarantees the bound $c_g \leq 4$, for a matching based two-level method on structured grids. In the numerical experiments, we instead use a Gauss-Seidel smoother for all structured grid problems. Using such a smoother retains a convergence rate $\sim(1 - 1/\log n)$ and $O(n \log n)$ computational complexity.

**Unstructured grids.** Each of the unstructured grids in our tests are constructed by first perturbing the coordinates of vertices of a structured grid, followed by Delaunay triangulation of the resulting set of vertices. For unstructured grids, we use a random matching algorithm. Numerical results show that the maximum degree of the coarser graphs grow only during the first few coarsening steps. Hence, smoothers such as Gauss-Seidel can approximate well $(S_k^T A_k S_k)^{-1}$ on all levels and the application of such a smoother has a complexity proportional to the number of degrees of freedom (DOF) on level $k$. We use the CG method to perform the action of $(S_k^T A_k S_k)^{-1}$ on a vector. Such approach is practical since $S^T AS$ is equally well conditioned on all levels.

Instead of using the same AMLI polynomial $q(t)$ on all levels, we determine the polynomials $q_k(t)$ on each level recursively, starting from the second coarsest level. After constructing a multilevel hierarchy, we use 6 AMLI two level cycles (level $(J-1)$ and level $J$) and a Lanczos algorithm to estimate the condition number of $B_{J-1}^{-1} A_{J-1}$. We apply this procedure recursively (and with 6 AMLI *multilevel* cycles from level $(k+1)$ to $J$) to estimate the condition number of $B_k^{-1} A_k$ on level $k$, for $k = 1, \ldots, J-2$. When all polynomials are determined, they are used in the AMLI cycle during the solving phase.

**Numerical tests.** We use the AMLI cycle as a preconditioner of Conjugate Gradient (CG) method. We stop the iterations when the relative residual becomes smaller than $10^{-10}$. The results are summarized in Table 1. The number of CG iterations is denoted by $M$, and the average convergence rate of the last five iterations is denoted by $r_a$. The CG coefficients are also used to estimate the condition number $\kappa(B_0^{-1} A_0)$, as suggested in [4]. The operator and grid complexities are less than 2 in all the examples presented below.

(a) 2D unit square

| DOF | $\kappa$ | $r_a$ | $M$ |
|---|---|---|---|
| $256^2$ | 18.4 | 0.55 | 32 |
| $512^2$ | 24.8 | 0.61 | 36 |
| $1024^2$ | 32.9 | 0.69 | 40 |

(b) 3D unit cube

| DOF | $\kappa$ | $r_a$ | $M$ |
|---|---|---|---|
| $32^3$ | 7.8 | 0.36 | 21 |
| $64^3$ | 11.4 | 0.45 | 25 |
| $128^3$ | 19.2 | 0.51 | 29 |

(c) 2D L-Shaped

| DOF | $\kappa$ | $r_a$ | $M$ |
|---|---|---|---|
| $(3/4)\cdot 256^2$ | 17.8 | 0.56 | 33 |
| $(3/4)\cdot 512^2$ | 23.9 | 0.64 | 36 |
| $(3/4)\cdot 1024^2$ | 31.7 | 0.69 | 38 |

(d) 3D Fichera

| DOF | $\kappa$ | $r_a$ | $M$ |
|---|---|---|---|
| $(7/8)\cdot 32^3$ | 7.5 | 0.40 | 22 |
| $(7/8)\cdot 64^3$ | 11.1 | 0.48 | 25 |
| $(7/8)\cdot 128^3$ | 15.8 | 0.55 | 29 |

(e) 2D unit square (ug)

| DOF | $\kappa$ | $r_a$ | $M$ |
|---|---|---|---|
| $256^2$ | 31.4 | 0.58 | 35 |
| $512^2$ | 36.7 | 0.63 | 39 |
| $1024^2$ | 42.0 | 0.58 | 41 |

(f) 3D unit cube (ug)

| DOF | $\kappa$ | $r_a$ | $M$ |
|---|---|---|---|
| $32^3$ | 29.5 | 0.51 | 35 |
| $64^3$ | 37.6 | 0.68 | 46 |
| $128^3$ | 48.3 | 0.72 | 52 |

**Table 1.** Results for structured grids on square, cubic, L-shaped and Fichera domain, and for unstructured grids (ug) on square and cubic domain. Here, $\kappa$ is an estimate (from CG) of $\kappa(B_0^{-1}A_0)$.



Note that for the 2D and 3D unstructured grid problems, the number of levels for a given unstructured grid is the same as that of a structured grid with the same degrees of freedom. We observe a logarithmic growth of the condition numbers with respect to the size of the grids, and fast convergence rates of the preconditioned CG method in all cases.

## 7 Conclusions

We present an AMLI (AMG) method based on graph matching with a nearly optimal convergence rate and computational complexity. We have also presented numerical tests which confirming our estimates. Our ongoing research is on extending the estimates to general aggregation algorithms and aggregates configurations and we are also investigating improvements of the AMLI method components.

**Acknowledgments** The authors gratefully acknowledge the support by the Austrian Academy of Sciences and by the Austrian Science Fund (FWF), Project No. P19170-N18 and the support from the National Science Foundation under grants NSF-DMS 0810982 and NSF-OCI 0749202.

## Bibliography

[1] O. Axelsson and P. S. Vassilevski. Algebraic multilevel preconditioning methods. II. *SIAM J. Numer. Anal.*, 27(6):1569–1590, 1990.

[2] R. D. Falgout, P.S. Vassilevski, and L.T. Zikatanov. On two-grid convergence estimates. *Numer. Linear Algebra Appl.*, 12(5–6):471–494, 2005.

[3] H. Kim, J. Xu, and L. Zikatanov. A multigrid method based on graph matching for convection-diffusion equations. *Numer. Linear Algebra Appl.*, 10(1–2):181–195, 2003.

[4] Yousef Saad. *Iterative methods for sparse linear systems*. SIAM, Philadelphia, PA, 2nd edition, 2003.

[5] U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press Inc., San Diego, CA, 2001.

[6] P. S. Vassilevski. *Multilevel block factorization preconditioners*. Springer, New York, 2008.