A LEAST SQUARES BASED VARIATIONAL AMG SOLVER FOR HPD SYSTEMS

KARSTEN KAHL

Abstract. This work concerns the development of an efficient algebraic multigrid (AMG) solver for the linear systems arising in lattice simulations of quantum chromodynamics (QCD). We present some first ideas and tests for developing an adaptive algebraic multigrid method for these systems based on Brandt's Bootstrap AMG (BAMG) framework. In our proposed solver, we use a variational adaptive setup algorithm to recursively construct the coarse spaces of the AMG hierarchy, with interpolation built using Brandt's notions of Compatible Relaxation, algebraic distances, and Least Squares fitting of relaxed vectors. In addition, in the resulting solver we use an optimal over-weighting in energy of the coarse-grid correction. We demonstrate the efficacy of these approaches for a set of Gauge Laplacian systems, involving hermitian positive definite matrices with highly disordered coefficients, as encountered in Lattice QCD. We conclude with a brief discussion of future work.

1. Introduction. The linear systems arising in numerical models of quantum theories are very challenging for traditional multigrid methods, due mainly to three properties of the system matrix. The first is the extremely disordered pattern of the matrix elements; each non-zero off-diagonal entry is chosen at random from a specific distribution function, with little correlation between neighboring coefficients. This makes it difficult to distinguish *weak* and *strong* local neighbors with classical strength-of-connection measures. The second is that the fluctuations of the matrix elements cause local oscillations in the low modes, those associated with the smallest magnitude eigenvalues. Hence, the conventional MG assumption that the low modes are locally constant does not hold. The third is that the support of the low modes can be local and thus the classical idea of using the lowest mode to locally represent other algebraically smooth error components is not applicable. Generally, whereas for standard elliptic operators, such as Poisson's equation, accurate local fitting of the slowest to converge mode of relaxation (usually the constant vector) is sufficient to ensure effective reduction of all other algebraically smooth error modes by the multigrid process, this is not the case for the extremely disordered systems that arise in quantum dynamical systems.

Classical AMG (cf. [7, 21]) can be seen as an important milestone on the way to black-box multigrid solvers. Nevertheless, the challenges outlined above are handled relatively inefficiently by classical AMG. Recent advances, including e.g. adaptive Smoothed Aggregation [11], adaptive (reduction-based) AMG [19, 12], indicate that adaptivity can be a key ingredient to efficient black-box solvers for a broader class of problems.

In this paper, we summarize our investigations to develop a method based on the idea of Bootstrap AMG (BAMG) suggested by Brandt [1, 2] for quantum dynamical systems. In our adaptive BAMG solver, the setup algorithm used in recursively computing the interpolation operator, $P \in \mathbb{C}^{n \times n_c}$, consists of two closely interrelated steps:

1. selecting the coarse-level degrees of freedom;
2. computing the entries of P (interpolation weights).

As we define the variational coarse-level system matrix as $A_c = P^* A P$, it is important that P is sparse, i.e. that many of the entries of P are zero. The main part of this paper is devoted to a precise formulation of these two steps, aimed at defining a highly accurate and also sparse P.

Section 3 contains a general description of the Bootstrap AMG setup scheme. Sections 4 and 5 contain details of the individual steps of our setup algorithm and the Section 6 contains numerical experiments for a simplified 2d QCD model, the Gauge Laplacian, described next.

2. The Gauge Laplacian. We consider a simplified two-dimensional model of the systems arising in quantum chromodynamics called the "Gauge Laplacian", as was done previously in [9, 15, 18]. The inverse Gauge Laplacian is the simplest form of a propagator satisfying a gauge theory, and thus provides for a good initial test problem in AMG algorithmic development for quantum dynamical applications.

Given a $U(1)$ gauge field $\mathcal{U} = \{u_j := e^{i\theta_j},\, j = 1, \ldots, n_e\}$ with gauge links u_j that live on the edges of a uniform 2D quadrilateral (torus-) grid $\{(k, \ell), k, \ell = 1, \ldots, N\}$ we aim at solving systems of the form

$$A(\mathcal{U})\varphi = \psi, \quad A(\mathcal{U}) \in \mathbb{C}^{n \times n}.$$

Here, $n = N^2$ denotes the grid size; the number of edges is thus $n_e = 2n$. We will use standard notation from theoretical physics in that φ, ψ denote (complex-valued) vectors from $\mathbb{C}^n$. The symbol x represents a lattice site, i.e. a point (k, ℓ) of the grid, and the operations $x \pm \mu$ for $\mu = 1, 2$ yield the neighboring lattice sites, i.e. $x \pm 1 = (k \pm 1, \ell)$ and $x \pm 2 = (k, \ell \pm 1)$ where all numbers are understood to be mod N.

The gauge links on the edges (one link u_j per edge j) act as coupling coefficients. To be explicit, the 2d Gauge Laplace matrix, $A = A(\mathcal{U})$, expresses a periodic nearest neighbor coupling which, for a pair of lattice sites x, y with corresponding matrix entry A_{xy}, can be described using the Kronecker δ as

$$A_{xy} = -\frac{1}{h^2} \sum_{\mu=1}^{2} \left(u_x^\mu \, \delta_{x+\mu,y} + (u_{x-\mu}^\mu)^* \, \delta_{x-\mu,y} \right) + (\frac{4}{h^2} + m)\delta_{x,y}. \tag{2.1}$$

Here, u_x^μ is the gauge link defined on the edge connecting lattice sites x and $x + \mu$ and $(u_{x-\mu}^\mu)^*$ is the complex conjugate of the gauge link defined on the edge connecting lattice sites $x - \mu$ and x. As usually, $h = 1/N$, and the parameter m can be interpreted physically as a mass. It is common to scale A to have unit diagonal, yielding $A = I - \kappa D$. The parameter $\kappa = \frac{1}{4+h^2 m}$ is known as the "hopping" parameter and matrix D as the hopping matrix, in the related physics literature. We hereafter work with the scaled matrix A.

To be physically relevant, the links u_j of the gauge field $\mathcal{U}$ have to be random variables from a given Boltzmann distribution. The distribution depends on a temperature parameter β. The case $\beta = 0$ yields the so-called cold or "free" configuration where $u_j = 1$ for all j. For $\beta = \infty$ we obtain hot configurations in which the phases θ_j in $u_j := e^{i\theta_j}$ are distributed in $[0, 2\pi)$ according to the corresponding Boltzmann distribution. Physically relevant configurations arise for $\beta \in (0, \infty)$, with choices depending on the lattice spacing.

2.1. Spectral properties of the Gauge Laplacian. It is easy to see from (2.1) that the Gauge Laplacian is hermitian. In our tests we choose the hopping parameter κ to generate matrices with the smallest eigenvalue shifted to $\frac{1}{N^2}$, yielding ill-conditioned, but still positive definite systems.

A key feature to consider when developing solvers for the Gauge Laplace system is the character of the algebraically smooth error, i.e. the near-kernel. In Figure 2.1, the modulus, real and imaginary part of the eigenmode to the smallest eigenvalue of the hpd system (2.1), which is naturally part of the algebraically smooth error, is shown for $\beta = 5$ on a 64×64 grid; the error after 50 Gauss-Seidel iterations is shown in Figure 2.2 as a representative of algebraically smooth error components.

Here we see two reasons why standard algebraic multi-grid approaches break down when applied to the Gauge Laplacian. First the random nature of the algebraically smooth error varies with the gauge field. Second, algebraically smooth error tends to have localized structure, rendering the standard argument of a single global representative for the local appearance of small eigenmodes rather useless.

3. Bootstrap AMG. In this section, we give a general description of Brandt's BAMG setup [1, 2, 4, 5] and provide some heuristic explanations of the key components of the algorithm.

Bootstrap AMG was proposed as a fully adaptive setup. As all algebraic multigrid methods, Bootstrap AMG consists of two closely related stages. First, the set of coarse-grid degrees of freedom (dofs) $\mathcal{C}$, rich enough to accurately represent the coarse-grid system is generated. This can be done for instance, as in our solver, using Compatible Relaxation. Next, inter-grid transfer operators are defined. In the case of Galerkin-type AMG methods for symmetric or hermitian matrices we only have to define interpolation, P as restriction is automatically given by P^* then.

By combining enough initially random, slightly smoothed test-functions we obtain a rich representation of local algebraically smooth error that can be used to define the so-called *algebraic distances*. These can be used to measure strength of connection and define sets of interpolatory points for each fine-grid dof as well as an accurate interpolation using the found sets of interpolatory points.

As we are interested in solving problems where strength of connection in the classical sense is not available, we substantiate the notion of algebraic distances, put forward by Brandt [4], to use as a measure in selecting sets of interpolatory points for each fine-grid dof, the aim being to determine

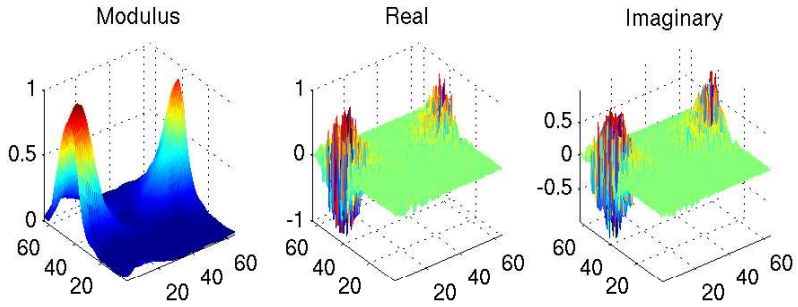

Fig. 2.1: Modulus, real and imaginary part of the eigenmode to the smallest eigenvalue for $\beta = 5$ on a 64×64 grid.

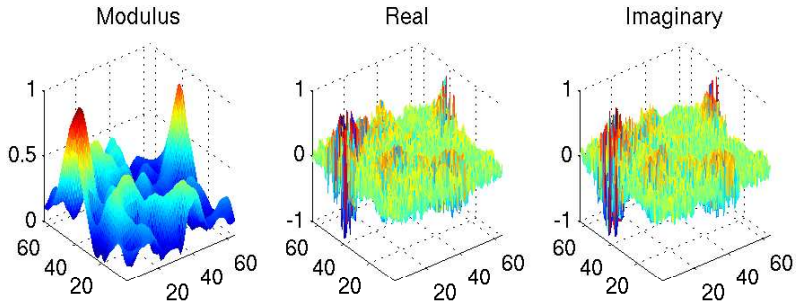

Fig. 2.2: Modulus, real and imaginary part of algebraically smooth error after 50 Gauss-Seidel iterations on a random initial guess for $\beta = 5$ on a 64×64 grid.

which fine-grid variables can be replaced by a combination of certain coarse-grid dofs to represent algebraically smooth error.

Finally, after defining a tentative set of interpolatory points, we compute interpolation weights that minimize a local least squares functional over this set of test-functions. This is done similar to the proposed strategy in [17], in which interpolation for multigrid eigensolvers is defined using a least squares formulation. Furthermore, the least squares functional allows us to adjust the accuracy and sparsity of interpolation in regions where the functional is large and also to drop particular coarse-grid interpolatory points when their weights are too small to yield meaningful information in the interpolation process.

4. Choosing the coarse variables: Compatible Relaxation. Modern approaches for constructing AMG coarse spaces are based on the idea of *Compatible Relaxation (CR)*, introduced by Brandt in [3]. The basic idea of CR is to use the given AMG relaxation scheme, restricted to an appropriately defined subspace, to measure the quality of the given coarse space and thereby iteratively improve it. We proceed with a brief overview of CR and its use in AMG coarsening. A detailed discussion, theory and comparisons between various measures of the quality of coarse spaces and relation with Compatible Relaxation schemes are presented in [3, 8, 10, 13, 14, 15, 16].

4.1. Classical AMG CR-based coarsening. Similarly to the classical AMG setting, we assume that the coarse-level dofs, $\mathcal{C}$, are viewed as a subset of the fine-level degrees of freedom, Ω, in which case CR reduces to $\mathcal{F}$-point relaxation, where $\mathcal{F} = \Omega \setminus \mathcal{C}$. In choosing the coarse variable set, we use the CR-based coarsening algorithm developed by Brannick and Falgout in [10] for SPD problems, modified appropriately for complex valued systems.

The idea of Compatible Relaxation is closely related to a convergence result due to a theorem of Falgout, Vassilevski, and Zikatanov (see [14]), giving the precise convergence rate of a variational

two-level MG algorithm:

$$\|E_{TG}\|_A^2 = 1 - \frac{1}{K(P)}, \qquad K(P) = \sup_v \frac{\|(I - \pi_{\widetilde{M}})v\|_{\widetilde{M}}^2}{\|v\|_A^2},$$

where $E_{TG} = (I - M^{-*}A)(I - PA_c^{-1}P^*A)(I - M^{-1}A)$, is the error propagation operator and $\widetilde{M} := M^*(M^* + M - A)^{-1}M$, is the symmetrization of the smoother M, and $\pi_{\widetilde{M}}$ is a $\widetilde{M}$-orthogonal projection onto the coarse space $V_c = \text{Range}(P)$. Assuming that the set of coarse degrees of freedom has been determined, the remaining task is to define an interpolation operator, $P : \mathbb{C}^{n_c} \longmapsto \mathbb{C}^n$, that (approximately) minimizes $K(P)$. Finding such a P is nontrivial as the dependence of $K(P)$ on P given above is complicated. A common approach is then to instead minimize an upper bound of K, obtained by replacing $\pi_{\widetilde{M}}$ with PR, where R is chosen such that $RP = I_{n_c}$; in this way PR represents a general projection onto Range(P). Here, the choice of coarse variable type determines the structure of P and hence R. For our choice of coarse variable type, for which $\mathcal{C} \subset \Omega$, we have $P = [W, I_{n_c}]^t$ and $R = [0, I_{n_c}]$. The resulting measure for the quality of the coarse space then reads:

$$\mu(P) = \sup_v \frac{\|(I - PR)v\|_{\widetilde{M}}^2}{\|v\|_A^2}.$$

We note that $\mu(P) \geq K(P)$ for all P and that this measure suggests that error components consisting of eigenvectors associated with small eigenvalues (i.e., near null components, not treated well by the relaxation defined via M) must be well approximated by P. The advantage of using $\mu(P)$ vs $K(P)$ is that the minimizer of $\mu(P)$ can be explicitly written as follows:

$$\mu(P_\star) = \min_P \mu(P), \quad \text{where} \quad P_\star = \left[-A_{fc}^* A_{ff}^{-1}\ , I\right]^t,$$

and A_{ff}, A_{fc}, and A_{cc} are the blocks in A corresponding to the $\mathcal{F}$-$\mathcal{C}$ splitting of the fine grid dofs. What is even more interesting in such settings, is that the asymptotic convergence factor of CR provides an upper bound [13] for the above minimum:

$$\mu(P_\star) \leq \frac{c}{1 - \rho_f}, \qquad \rho_f = \|E_f\|_{A_{ff}}^2, \quad E_f = (I - M_{ff}^{-1}A_{ff}).$$

A conclusion that follows immediately from this result is that ρ_f provides a *computable* measure of the quality of the coarse space, that is, a measure of the ability of the set of coarse variables to represent error not eliminated by fine-grid relaxation. The CR coarsening scheme developed in the setting of classical AMG [10] is based on this result which (as mentioned above) illustrates that for a given splitting of Ω into $\mathcal{C}$ and its complement set $\mathcal{F}$, if CR is fast to converge, then there exists a P such that the resulting two-level method is uniformly convergent. The CR algorithm thus provides a way to tie the selection of $\mathcal{C}$ to the relaxation. Generally speaking, in such algorithms the set of coarse variables is constructed using a multistage coarsening algorithm, where a single stage consists of: (1) running several iterations of CR (based on the current $\mathcal{F}$) and (2) if CR is slow to converge, adding an independent set of fine-level variables (not effectively treated by CR) to $\mathcal{C}$. Steps (1) and (2) of the algorithm are applied repeatedly until the convergence of CR is deemed sufficient, giving rise to a sequence of coarse variable sets:

$$\emptyset = \mathcal{C}_0 \subseteq \mathcal{C}_1 \subseteq ... \subseteq \mathcal{C}_m,$$

where, for the accepted coarse set $\mathcal{C} := \mathcal{C}_m$, convergence of CR is below a prescribed tolerance.

4.2. CR for the Gauge Laplacian – an analysis of geometric coarsening. Our aim here is to use CR as a tool to enhance a tentative choice of coarse grids (e.g., taking full coarsening, obtained by doubling the grid spacing in each spatial dimension) for certain choices of the 2d Gauge Laplace operator. Although we apply such techniques only to the simplified hermitian and positive definite (HPD) 2d Gauge Laplace operators, similar strategies are also applicable to the 4d non-HPD systems of QCD. The goal is to begin to study in a more rigorous way the (near) kernel of systems encountered in lattice QCD and their local representations on coarse grids.

CR convergence factors for Gauge Laplacians

β/N	32^2	64^2	128^2
1	.71	.71	.68
5	.68	.68	.68
10	.68	.67	.68

Table 4.1: Computed asymptotic convergence rates of $\mathcal{F}$-point Gauss-Seidel CR iterations for the Gauge Laplacians shifted to $\frac{1}{N^2}$ and full coarsening.

CR convergence factors for Free case

N	32^2	64^2	128^2
ρ	.73	.73	.73

Table 4.2: Computed asymptotic convergence rates, ρ, of $\mathcal{F}$-point Gauss-Seidel CR iterations for the Free case shifted to $\frac{1}{N^2}$ and full coarsening.

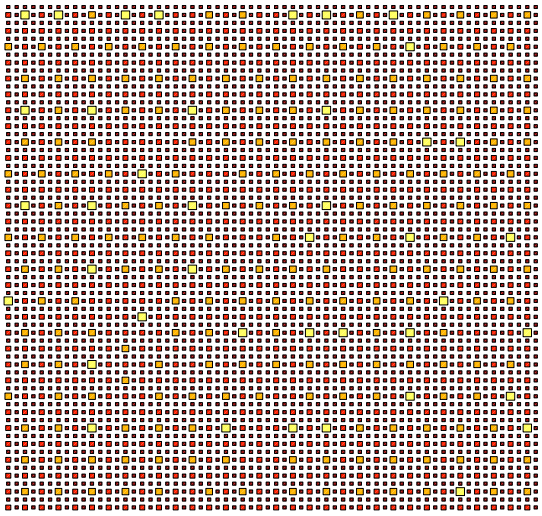

Fig. 4.1: Grid-hierarchy generated by CR and caliber 4 least squares interpolation, as defined in Section 5, for a Gauge Laplacian with $\beta = 5$ on a 64×64 grid on 4 levels. Using full coarsening on the finest grid only.

In Table 4.1, we report asymptotic convergence rates of Compatible Relaxation for a full coarsening, i.e. taking every other grid-point in each dimension, for the free case, i.e. $\beta = 0$, and for Gauge Laplace operators for several temperature parameters β and sizes. This choice of grid is justified as it is known to be a suitable for the free problem, and we can see from the CR rates in Table 4.2 that it is also a suitable choice for the Gauge Laplace operators, and this is independent of size and temperature β.

Generally, we initially choose $\mathcal{C}_0$ to match full coarsening when using CR to generate a $\mathcal{C}$-$\mathcal{F}$-splitting in our multigrid algorithm and add additional points when necessary in order to achieve the desired CR rate.

In Figure 4.1, we see the multigrid coarsening generated by CR for a 64×64 Gauge Laplace operator with $\beta = 5$ and 4 levels. Starting with full-coarsening on the finest grid we generate successive coarser grids with $\mathcal{C}_0 = \emptyset$. The non-global support of the near-kernel is reflected in the choice of coarse grids.

5. Building AMG interpolation: algebraic distance and Least Squares fitting of relaxed vectors. We now proceed with discussions of our methods for defining the nonzero sparsity pattern of interpolation, i.e., the coarse degrees of freedom used to interpolate the coarse-level correction to the finer level, and for computing the weights for these non-zero entries. We hereafter assume that the set of coarse variables is given, that is, the coarse-grid dofs have already been determined using our CR coarsening algorithm.

5.1. Algebraic Distances. A basic concept of classical AMG [6, 7, 22], is building interpolation using an operator-based definition of strength-of-connections. A connection between two degrees of freedom i and j is strong if

$$d_{ij} = \frac{|a_{ij}^-|}{a_{ii}} \frac{(e_i - e_j)^2}{e_i^2} \tag{5.1}$$

is small for algebraically smooth error e and negative a_{ij}^-. This measure provides a good criteria of closeness between i and j assuming that algebraically smooth error [22] is also geometrically smooth. If it is, then small (5.1) means that the same coarse-grid correction will give a good reduction simultaneously at both i and j. If one of these dofs, say the ith, is actually a coarse dof, then its value can be used to significantly reduce the error at j. However, if the algebraically smooth error is no longer geometrically smooth, this measure looses its meaning. There are some notions of strong positive connections in the classical AMG theory, but still no general strategy for measuring strength in cases where strong connections are defined without any assumption on the geometric properties of the algebraically smooth error.

Recently, Brandt [4] proposed an alternate, more general criteria that allows one to determine which points (for example from the coarse set) can accurately transfer information to a given dof. Instead of defining strength-of-connections using the entries in the system matrix, the approach uses the notion of algebraic distances based on a set of relaxed vectors. More specifically, this measure of strength is defined as follows. For a set of functions $u^{(k)}$ and two grid points i and j, we define algebraic distance between i and j as

$$d_{ij}^{\alpha} = \sum_k (u_i^{(k)} - \delta_{ij} u_j^{(k)})^2 \ . \tag{5.2}$$

If the $u^{(k)}$ are all algebraically smooth and (5.2) is small for some δ_{ij} then there exists a constant relation between the dof i and j over a rich representation of algebraically smooth error. Grid points i and j can then indeed be considered as neighbors or, in our terms, algebraically close. Note that the computation of (5.2) only requires so solve a $1 \times k$ least squares problem.

To calculate a rich set of algebraically smooth functions, a set of random functions $u^{(1)}, \ldots, u^{(k)}$ is treated by a few steps of our multigrid relaxation method with the zero right-hand-side. The resulting functions then give a rich local representation of smooth error. In order to avoid situations where d_{ij}^{α} is accidently small, but the two dofs are not algebraically close to each other, a reasonable number of test-functions has to be used.

After computation of d_{ij}^{α}, for each fine dof it is possible to simply choose the algebraically closest neighbors from the coarse-grid set to be taken into account, when defining interpolation. Although it is possible to compute d_{ij}^{α} for all $i \in \mathcal{F}$ and $j \in \mathcal{C}$ we will restrict the computations to pairs (i, j) that are connected over a path of length L in the corresponding graph of the matrix.

Note that the classical meaning of strength-of-connection can be recovered from the definition of algebraic distances. Indeed, if algebraically smooth error is also geometrically smooth, i.e. it changes slowly along large negative couplings, algebraic distance between two degrees of freedom connected by a large negative coupling will give $d_{ij}^{\alpha} \approx 0$ with $\delta_{ij} \approx 1$, and we recover the definition of strength of connection. On the other hand, in the case of algebraically smooth error that is geometrically highly oscillatory, we expect $d_{ij}^{\alpha} \approx 0$ with $\delta_{ij} \approx -1$ for i and j connected by a large positive coupling.

As long as the set of test-functions is in some sense rich enough to represent the nature of smooth error, algebraic distances allow to find strong couplings in the equations without any assumption on the geometrical properties of algebraically smooth error.

5.2. Least Squares fitting of relaxed vectors. We now outline briefly our Least Squares (LS) approach for defining interpolation. We assume that we are given a $\mathcal{C}$-$\mathcal{F}$ separation of variables like in classical AMG. Given the $\mathcal{C}$-$\mathcal{F}$ choice, we aim to compute weights of the interpolation operator.

The new key component of our adaptive (Bootstrap) AMG solver is the idea to use several relaxed vectors, representing algebraically smooth error, to define the MG interpolation weights. The idea of building (LS-based) interpolation on relaxed vectors is heuristically motivated by Brandt in [2], using the observation that only a local representation of the algebraically smooth error is needed to construct an effective (sufficiently accurate) local interpolation and thus the coarse-grid system.

The goal of the AMG interpolation is to accurately transfer algebraically smooth error. In our algorithm it is done by optimally fitting a set of algebraically smooth functions in the LS sense. For a given set of such functions $u^{(k)}$, we are seeking for a prolongation operator P with a fixed sparsity pattern such that

$$P = \text{argmin} \left(\sum_k \omega_k ||u^{(k)} - PRu^{(k)}||^2 \right) , \tag{5.3}$$

where ω_k are positive weights, R is a restriction operator, and $Ru^{(k)}$ is the coarse-grid representation of the fine-grid $u^{(k)}$. While the minimization process (5.3) seems to be global, it is in fact not. Only interpolation weights needed to interpolate to an individual fine dof i are coupled simultaneously, leading to the local least square functional

$$\mathcal{F}(P_i) = \sum_k \omega_k (u_{ik} - \sum_{j \in \mathcal{P}_i} p_{ij} U_{jk})^2 \tag{5.4}$$

for each fine-grid dof i, where $\mathcal{P}_i$ is the set of interpolatory points involved in the interpolatation to the fine dof i.

Calculation of the test-functions $u^{(k)}$ is done starting with a random initial guess, with only a few relaxation sweeps applied to the homogenous problem, similarly to algebraic distances processing. This is sufficient to get a local representation of the smooth error used in the adaptive process to improve the coarse space. In order to enrich this LS approximation, several such test-functions are used to define the interpolation in a least squares fashion. The number of functions needed depends on the *caliber* of P, i.e. the number of coarse-grid points used to interpolate to a fine grid point, chosen a-priori.

Given a tentative restriction operator $\hat{R}$, injection in our implementation, we compute initial coarse-grid functions $U^{(k)}$ by restricting the fine-grid vectors:

$$U^{(k)} = \hat{R} u^{(k)} .$$

Using these coarse-grid functions, the modified linear functional

$$P = \text{argmin} \left(\sum_k \omega_k ||u^{(k)} - PU^{(k)}||^2 \right) , \tag{5.5}$$

is minimized, and a new approximation to P is calculated. Then the process of recalculating $U^{(k)}$ with an improved restriction operator, $R = R(P)$, can be repeated, etc. The process, however, seems to converge fast in computational practice and only few iterations are needed in our algorithm. In our initial tests, we take $\omega_k = 1$, since the random choice of the test-functions and identical treatment makes their information equally valuable, and do not give a reason to bias toward an individual test-function, which could be achieved by choosing a bigger weight.

The LS approach allows an adaptive processing toward improving the accuracy of P. If the minimum of the LS functional (5.4) for the i^{th} row of P is deemed too large, it is possible to add additional interpolatory points to $\mathcal{P}_i$ to increase interpolation accuracy. Furthermore, the LS interpolation weights p_{ij} can be used to reduce the number of interpolatory points. Interpolatory points whose interpolation weights turn out to be much smaller than the rest of the weights might be discarded without loosing much accuracy. However, after adding additional points or dropping interpolatory points the adjusted LS functional has to be minimized again.

To illustrate the efficiency and effectiveness of the Least Squares approach, we consider as a test problem the one-dimensional Poisson equation. In Figure 5.1, we see that the least square fit of several smoothed test-functions yields an accurate interpolation operator that is very close to linear interpolation which is indeed a suitable choice for P in this case.

6. Numerical Results. We test our Bootstrap AMG method for hermitian positive definite Gauge Laplacians. Given the linear system of equations $Au = f$ on the finest grid, $i = 1$ with $A \equiv A_1$, we construct a multigrid hierarchy variationally:

$$A_{i+1} = (P_{i+1}^i)^* A_i P_{i+1}^i , \tag{6.1}$$

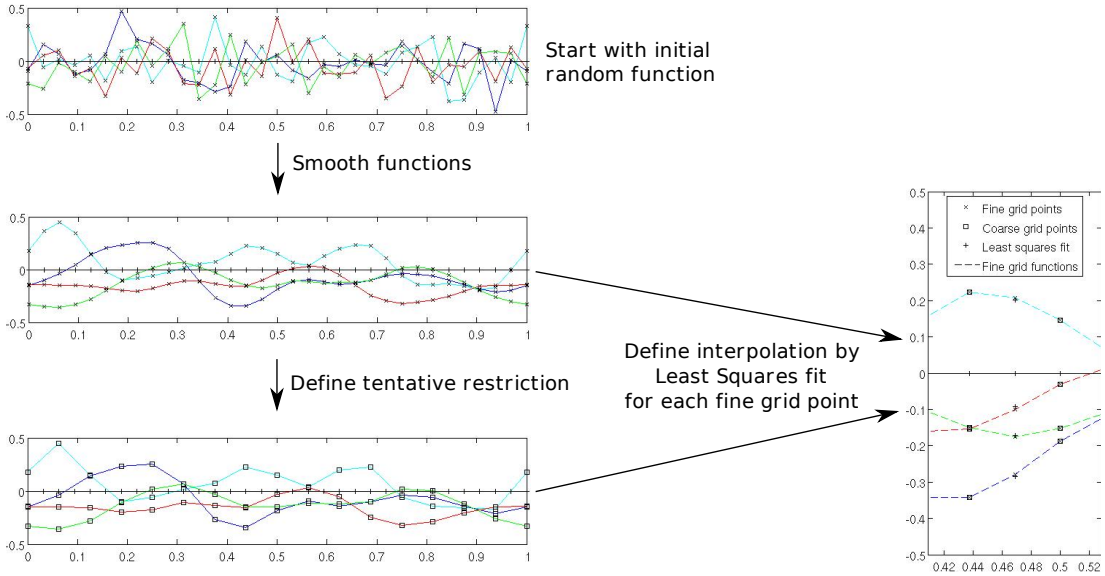

Fig. 5.1: LS fitting for 1D Poisson problem $u'' = 0$ on a 32 grid; 4 test-functions.

where the interpolation operator P^i_{i+1} acts on vectors living on the coarse grid, $i+1$, and transfers them to the related finer grid, i. Once the hierarchy is built, we solve the system with a slightly modified V-cycle. Motivated by work of Brandt [3], and Míka and Vaněk [20], we employ an optimal weight ω_{opt} to the correction in the multigrid cycle that fulfills

$$\omega_{opt} = \text{argmin}\left(\|B\left(I - \omega P A_c^{-1} P^* A\right) Be\|_A^2\right) \ . \tag{6.2}$$

The solution to (6.2), assuming B is hermitian, is given by

$$\omega_{opt} = \frac{\langle B^2 e, BPA_c^{-1}P^*AMe\rangle_A}{\|BPA_c^{-1}P^*ABe\|_A^2} \ .$$

As the optimal weight depends on the error in each step, the resulting multigrid cycle is no longer stationary, so that we omit this weighting when we use the MG solver as a preconditioner for CG.

6.1. Algorithmic Details. We use Gauss-Seidel as our MG relaxation method in the setup and solve phases of the algorithm.

- Compatible Relaxation:
  Initialize $\mathcal{C} = \mathcal{C}_0$ or $\mathcal{C} = \emptyset$
  Initialize $\mathcal{U} = \Omega - \mathcal{C}$
  While $\mathcal{U} \neq \emptyset$
    Perform ν Compatible Relaxation sweeps
    $\mathcal{U} = \{i : \frac{|u_i^\nu|}{|u_i^{\nu-1}|} > \theta\}$
    $\mathcal{C} = \mathcal{C} \cup \{\text{independent set of } \mathcal{U}\}$
  We initialize $\mathcal{C}$ with the set $\mathcal{C}_0$ of full-coarsening; and choose $\theta = .75$.
- Algebraic distances:
  Generate random test-functions $u^{(1)}, \ldots, u^{(k)}$
  Relax each test-function ν times
  for all i in $\mathcal{F}$
    Compute $\mathcal{C}_i^l = \{j : i \text{ connected to } j \text{ by a path of length } \leq l\}$
    Compute $d_{ij}^\alpha = \sum_k (u_i^{(k)} - \delta_{ij} u_j^{(k)})^2$ for all $j \in \mathcal{C}_i^l$
    Set $\mathcal{U}_i = \{j : d_{ij}^\alpha < \gamma \min_{k \in \mathcal{C}_i^l}(d_{ij}^\alpha)\}$
    Take $\min(c, |\mathcal{U}_i|)$ algebraically closest elements of $\mathcal{U}_i$ to define $\mathcal{P}_i$

β/n	32^2	64^2	128^2
1	.57 (10) 1.67	.67 (10) 1.66	.78 (9) 1.86
5	.39 (10) 1.69	.67 (10) 1.76	.79 (9) 1.84
10	.35 (9) 1.69	.55 (10) 1.7	.69 (9) 1.83

Table 6.1: Computed asymptotic convergence rates and PCG iteration count for the Gauge Laplacians shifted to $\frac{1}{N^2}$. Initial caliber of interpolation 4. Full coarsening on the first grid, 4 levels at max.

n	32^2	64^2	128^2
	.19 (8) 1.59	.19 (8) 1.6	.23 (9) 1.6

Table 6.2: Computed asymptotic convergence rates and PCG iteration count for the Free Case shifted to $\frac{1}{N^2}$, Initial caliber of interpolation 4. Full coarsening on all levels, $4, 5$ and 6 levels respectively.

We use $l = 2$, caliber $c = 4$ and $\gamma = 10$. Furthermore, $k = 8$ to $k = 12$ and $\nu = 8$ to $\nu = 12$.

- Least Squares interpolation:
  Generate random test-functions $u^{(1)}, \ldots, u^{(k)}$ (already done for algebraic distances)
  Compute tentative restriction R and coarse-grid functions $U^{(i)} = Ru^{(i)},\ i = 1, \ldots, k$.
  for $\mu = 1, \ldots, n$
    for all i in $\mathcal{F}$
      Minimize $\mathcal{F}(P_i) = \sum_k \omega_k (u_{ik} - \sum_{j \in \mathcal{P}_i} p_{ij} U_{jk})^2$
      Drop entry p_{ij} if $|p_{ij}| < \delta \max_{j \in \mathcal{P}_\setminus} |p_{ij}|$
      Update $\mathcal{P}_i$ if necessary and recompute LS-fit.
    Define $R = (P^*P)^{-1}P^*$ and $U^{(i)} = Ru^{(i)},\ i = 1, \ldots, k$

  We use $n = 2$ to $n = 4$, $\delta = .25$ and $\omega_k = 1$.
- Algebraic Multigrid Solver:
  V$(2, 2)$-cycle with 4 up to 6 levels and optimal coarse-grid overcorrection ω_{opt}.

6.2. BAMG for the Gauge Laplacian. We report the performance of our method for a number of Gauge Laplacians with temperatures $\beta = 1, 5, 10$ and grid-sizes $N = 32^2, 64^2, 128^2$. The smallest eigenvalue of the system matrix is shifted to $\frac{1}{N^2}$ by the appropriate choice of m. In Table 6.1 we report results for a stand-alone V$(2, 2)$ BAMG multigrid method using optimal overweighted coarse-grid correction and the iteration count of the V$(2, 2)$-multigrid preconditioned CG method as well as the operator complexities of the resulting solver. We solve the systems with a point-source right-hand side, given by the first canonical basis vector. We start with a zero initial guess and iterate until the initial residual is reduced by a factor of 10^8 in the $\|.\|_2$. In Table 6.2, we provide results for the over-weighted V$(2, 2)$ BAMG multigrid method and the associated preconditioned CG method for the free case (five-point discrete Laplacian), also shifted such that the smallest eigenvalue is of magnitude $\frac{1}{N^2}$.

Most notably our results show fast convergence, independent of lattice size and gauge field parameter β, at least when BAMG is used as a preconditioner. These results are achieved for almost constant, small operator complexities. Given these results and the relatively cheap and highly customizable setup phase our method can be regarded as an efficient competitor to other adaptive approaches.

7. Conclusions. The proposed BAMG setup produces an effective solver for the highly disordered Gauge Laplace test problems considered. Omitting our choice of full coarsening as an initial grid for the CR coarsening algorithm yields a completely adaptive black-box MG solver. Although research on development of adaptive schemes for the various steps of the AMG setup algorithm have been on-going for many years, our method is not only a new variant of such methods but unique in that it is a fully adaptive scheme. As such, we expect that our BAMG solver (with appropriate modifications) will also be applicable to general anisotropic diffusion problems as well as systems of PDEs, e.g., the 2d $U(1)$ Schwinger systems in QED and 4d $SU(3)$ Dirac Wilson system of QCD, and incompressible linear elasticity. Our current work focuses on convergence analysis of the proposed

method.

Future work will be on extending the BAMG setup to handle non-hermitian and indefinite systems. We are also interested in exploiting further integration of the individual processes of our setup algorithm. For example, using information of smoothed random test-functions and algebraic distances to improve the quality of coarse-grids generated by Compatible Relaxation. Investigating the use of the least squares functional to locally improve interpolation and thereby providing a more direct control over accuracy and sparsity of interpolation is another topic of future research. We also mention that while the results stated here are only preliminary they justify and motivate further effort to better understand and enhance the proposed BAMG ideas.

Acknowledgements. I want to thank Ira Livshits and James Brannick for their encouragement to participate in this year's student paper contest, for their unpayable advice and countless discussions on the subject and for their effort in proofreading and shaping this paper. I also want to thank Matthias Bolten and Andreas Frommer for their input and time spent.

Furthermore, I want to acknowledge Achi Brandt for many inspiring ideas on how to advance adaptive multigrid and Rob Falgout who brought us all together this summer and primed the discussions that lead to this paper.

REFERENCES

[1] A. Brandt, Multi-level adaptive solutions to boundary value problems. Mathematics of Computation 31 (1977), pp. 333–390.
[2] A. Brandt, Multiscale scientific computation: Review 2000 Weizmann Institute tech. report, 2000.
[3] A. Brandt, General highly accurate algebraic coarsening. *Electronic Transactions on Numerical Analysis*, Volume 10, 1–20, 2000.
[4] A. Brandt, Personal communication.
[5] A. Brandt, I. Livshits, Multiscale method for finding many eigenfunctions of PDE operators. In preparation.
[6] A. Brandt, S. McCormick, and J. Ruge, Algebraic multigrid (AMG) for automatic multigrid solutions with application to geodetic computations. Report, Inst. for Computational Studies, Fort Collins, CO, October, 1982.
[7] A. Brandt, S. McCormick and J. Ruge, Algebraic multigrid (AMG) for sparse matrix equations. in Sparsity and Its Applications, D. J. Evans (ed.), Cambridge University Press, Cambridge, 1984.
[8] J. Brannick, Adaptive algebraic multigrid coarsening strategies. Ph.D. thesis, 2005.
[9] J. Brannick, R. Brower, M. Clark, J. Osborn, and C. Rebbi, Adaptive multigrid algorithm for lattice QCD, Phys. Rev. Lett., Issue 4, Volume 100, article 041601, Jan. 2008.
[10] J. Brannick and R. Falgout, Compatible relaxation and coarsening in algebraic multigrid. In preparation.
[11] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge. Adaptive smoothed aggregation (αSA). SIAM J. Sci. Comp., 25(6):1896–1920, 2004.
[12] M. Brezina, R. Falgout, S. MacLachlan, T. Manteuffel, S. McCormick, and J. Ruge. Adaptive AMG (aAMG). SIAM J. Sci. Comp. 26 pp. 1261–1286, 2005.
[13] R. Falgout and P. Vassilevski, On generalizing the amg framework. SIAM J. Numer. Anal., Volume 42 , Issue 4, 1669 - 1693, 2004.
[14] R. Falgout, P. Vassilevski, and L. Zikatanov, On two–grid convergence estimates. Numerical Linear Algebra with Applications, Volume 12 Issue 5-6, 471–494, 2005.
[15] K. Kahl, An algebraic multi-level approach for a model problem for disordered physical systems. Diplomarbeit, Bergische Universität Wuppertal, Fachbereich Mathematik und Naturwissenschaften, 2006.
[16] O. Livne, Coarsening by compatible relaxation. Numerical Linear Algebra with Applications, Volume 11, Issue 2-3, Pages 205 - 227, 2004.
[17] I. Livshits, One-dimensional algorithm for finding eigenbasis of the Schrödinger operator. SIAM Journal on Scientific Computing, Volume 30(1), pp 416-440, 2008.
[18] S. MacLachlan, C. Oosterlee, Algebraic multigrid solvers for complex-valued matrices, SIAM Journal on Scientific Computing, Volume 30(3), pp 1548-1571, 2008.
[19] S. MacLachlan, T. Manteuffel, S. McCormick Adaptive reduction-based AMG, Numerical Linear Algebra with Applications, Volume 13, pp 599-620, 2006.
[20] S. Míka, P. Vaněk, A modification of the two-level algorithm with overcorrection. Applications of Mathematics, Volume 37(1), pp 13-28, 1992.
[21] J. Ruge, K. Stüben, Algebraic multigrid, in Multigrid Methods, S. F. McCormick, ed., vol. 3 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1987, pp. 73-130.
[22] U. Trottenberg, C. Oosterlee, and A. Schüller. Multigrid. Academic Press, San Diego, 2001.