MULTILEVEL SOLVERS FOR UNSTRUCTURED SURFACE MESHES ∗

BURAK AKSOYLU †, ANDREI KHODAKOVSKY †, AND PETER SCHRÖDER †

Abstract. Parameterization of unstructured surface meshes is of fundamental importance in many applications of Digital Geometry Processing. Such parameterization approaches give rise to large and exceedingly ill-conditioned systems which are difficult or impossible to solve without the use of sophisticated multilevel preconditioning strategies. Since the underlying meshes are very fine to begin with, such multilevel preconditioners require mesh coarsening to build an appropriate hierarchy. In this paper we consider several strategies for the construction of hierarchies using ideas from mesh simplification algorithms used in the computer graphics literature. We introduce two novel hierarchy construction schemes and demonstrate their superior performance when used in conjunction with a multigrid preconditioner.

Key words. multilevel preconditioning, multigrid, hierarchical basis multigrid, BPX, computer graphics, unstructured surface mesh, surface parameterization, harmonic weights, mean value weights, mesh coarsening.

AMS subject classifications. 65Y20, 65F10, 65N55, 65D18, 65M50

1. Introduction. Unstructured triangle meshes which approximate surfaces of arbitrary topology (genus, number of boundaries, number of connected components) appear in many application areas. Examples range from iso-surfaces extracted [49] from volumetric imaging sources and scientific simulations [53] to surfaces produced through range scanning techniques (e.g., [10, 46]) in areas as varied as historical preservation, reverse engineering, and entertainment. These meshes can be quite detailed: 100,000 samples, i.e., point positions on the surfaces, are quite common with many datasets ranging into the millions and some even into billions [46] of samples. Processing such meshes efficiently, in particular when numerical simulation algorithms are involved, requires sophisticated solvers.

One of the most fundamental issues in the processing of such geometry is the establishment of a parameterization, i.e., the construction of functions from sections of the surface to regions in $\mathbb{R}^2$. For example, parameterizations are critical for the approximation of one surface by another (e.g., [44, 43, 70]), numerical simulation of the mechanics of such surfaces (e.g., [26, 5]), their resampling (e.g., [18, 45, 31]), editing [42], or just plain decoration (“texture mapping”) [60].

Most parameterization algorithms define a good parameterization as one which minimizes some energy functional [18, 37, 62, 29, 27, 64, 25, 55, 39]. The energy serves to encode measures such as low distortion [59, 60], conformality [33, 13, 47, 28], area preservation [13], or elastic energy [50]. Others define the solution to satisfy barycentric coordinate conditions [19, 21] which take the original triangle shapes into account.

A basic building block of all of these parameterization algorithms is the solution of sparse linear systems which are defined relative to the input mesh, i.e., they contain as many degrees of freedom (DOFs) as vertices. Such systems tend to be ill-conditioned for large meshes and standard iterative solver methods typically take very long to
converge or even fail to converge. This comes as no surprise since the leading order terms of most of these parameterization algorithms correspond to discretizations of 2nd order elliptic PDEs. For such operators appropriate preconditioning methods have long been studied and optimal methods are well known [32].

These preconditioners are generally based on hierarchical representations, which are easy to come by in the standard structured refinement setting (e.g., quadrasection or longest edge bisection refinement [56, 57]). However, in the cases of interest to us we are confronted with a setting in which we are handed a very fine, unstructured mesh and any kind of hierarchical solver machinery must address the issue of coarsening, not refinement.

1.1. Contributions. In this paper we construct effective multilevel preconditioners in the unstructured, 2-manifold (with boundary) triangle mesh setting. As example problems we consider the linear systems arising in mesh parameterization algorithms, both symmetric and non-symmetric. Our hierarchy construction is based on mesh simplification methods using vertex removal [16] as their primitive operations. These simplification methods guarantee a logarithmic number of levels, i.e., geometric decay of DOFs from finest to coarsest. To avoid the problems of retriangulation of polygonal holes in \( \mathbb{R}^3 \) we implement vertex removal through half-edge contractions [35, 15]. We propose three hierarchy constructions, which exhibit different decay rates in the number of DOFs. Two of these constructions are entirely novel and perform exceedingly well on problems of interest. The different hierarchies are compared vis-a-vis different multilevel preconditioning strategies. Our novel, fast decaying, unstructured MIS hierarchy, coupled with the multigrid (MG) preconditioner, exhibits the by far best overall runtime over a range of problem sizes.

1.2. Outline. In Section 2 we present the mathematical framework for surface parameterization and elaborate how the linear systems arises from the parameterization process. The parameterization weights of interest are harmonic and mean value weights, and we provide some background on them. The mesh coarsening process is described in Section 3, together with the methods we use to construct fast, medium and slow decaying hierarchies using the notion of maximal independent sets in a variety of ways. The construction of prolongation matrices is elaborated in Section 4. The different preconditioners we consider are described in Section 5, together with their relations to each other. To understand the observed performance better we analyze the sparsity fill-in for the different hierarchies in Section 6. Finally, Section 7 documents our numerical experiments and the superiority of the MG preconditioner coupled with our novel fast decaying MIS hierarchy. Conclusions in Section 8 close the paper.

2. Parameterization. We begin by fixing our notation in describing the basic parameterization setup. A 2-manifold triangle mesh of arbitrary genus, possibly with boundary, is given as a simplicial complex \( M = (V, E, T) \). This mesh is (typically) embedded in \( \mathbb{R}^3 \), i.e., each vertex \( v_i \in V \) has associated with it a point position \( p_i \in \mathbb{R}^3 \). The point positions are extended in a piecewise linear fashion using the incidence relations given by the set of edges, \( e_{ij} \in E \) and triangles \( t_{ijk} \in T \). For surfaces of arbitrary topology, not equivalent to a disk, parameterizations are typically computed by combining a sequence of individual parameterization problems for disk like subsections of the surface [18, 29, 39]. To simplify the exposition in this paper we will consider \( M \) to be a (sub-)mesh topologically equivalent to a disk.

A parameterization of \( M \) is a piecewise linear injective map from the embedded
surface to a region $\Omega \subset \mathbb{R}^2$

$$\psi : \mathbb{R}^3 \supset M \to \Omega \subset \mathbb{R}^2.$$ 

$\Omega$ may be thought of as a flattened version of the same mesh (see Figure 2.1), i.e., it shares the same combinatorial structure $(V, E, T)$. Consequently $\psi$ is fully specified by the values it takes on the vertices $v_j \in V$, $\psi(v_j) \in \Omega$. Note that the injectivity of $\psi$ implies that no triangle is “flipped” under the parameterization.

Fig. 2.1: Igea face (Model 2) and its parameterization with harmonic weights.

Let $V_I$ and $V_B$ denote the interior and boundary vertices of $V$ and $1-R_i$, the $1$-ring, i.e., the set of edge neighbors, of $v_i$, $1-R_i = \{ v_j | e_{ij} \in E \}$. The commonly described approaches to solving for a $\psi$ with desired properties all reduce to solving a sparse linear system. These are defined by specifying for each vertex $v_i \in V_I$ a set of weights $\lambda_{ij}$, one for each vertex $v_j \in 1-R_i$. The parameter points $\psi(v_i) \in \Omega$ are then found by solving the linear system

$$\psi(v_i) \sum_{v_j \in 1-R_i} \lambda_{ij} - \sum_{v_j \in 1-R_i \cap V_I} \lambda_{ij} \psi(v_j) = \sum_{v_j \in 1-R_i \cap V_B} \lambda_{ij} \psi(v_j), \quad v_i \in V_I. \quad (2.1)$$

This can be written as the matrix equation

$$Ax = b, \quad (2.2)$$

where $x = \{ \psi(v_i) \}_{v_i \in V_I}$ is the vector of DOFs in some arbitrary order, and $b$ the right hand side of equation (2.1). The matrix $A = \{ a_{ij} \}_{v_i, v_j \in V_I}$ has dimension $N \times N$ with $N = |V_I|$ and entries

$$a_{ij} = \begin{cases} 
\sum_{v_k \in 1-R_i} \lambda_{ik}, & i = j, \\
-\lambda_{ij}, & v_j \in 1-R_i, \\
0, & \text{otherwise}.
\end{cases}$$

1In those cases in which $\psi$ is given as the solution of some non-linear functional we generally get a sequence of linear systems of the same basic structure.
Note that $\lambda_{ij}$ may or may not be equal to $\lambda_{ji}$ leading respectively to symmetric or non-symmetric $A$. If all $\lambda_{ij}$ are non-negative, $\psi(v_i)$ lies in the convex hull of its neighbors $\psi(v_j)$ ($v_j \in 1-R_i$), which, taken together with a convex boundary mapping for all $V_B$, ensures the injectivity of the solution [67] (see also Floater [20] for a more recent, simpler proof). For example, the simple set of weights $\lambda_{ij} = 1$ has this property. Unfortunately these weights take none of the geometry (angles, lengths, sampling, etc.) of the original embedding into account. If the map is to have properties such as angle or area preservation the weights must depend on the geometry of the embedded mesh.

For purposes of this article we consider two different types of weights: harmonic [54] and mean value [21]. Harmonic weights are defined by

$$\lambda_{ij} = \lambda_{ji} = \cot a_k + \cot a_l,$$

where $a_k$ and $a_l$ are the angles opposite $e_{ij}$ in the two incident triangles $t_{kji}$ and $t_{ijl}$ (see Figure 2.2). Note that boundary edges ($v_i, v_j \in V_B$), which have only one incident triangle, do not appear, ensuring that the $\lambda_{ij}$ are well defined. Boundary conditions can be of the Dirichlet type—with prescribed mappings of the $V_B$—or of a Neumann ("natural") type [13].

The harmonic weights arise from the standard piecewise linear finite element approximation to the Laplace-Beltrami operator and are based on the fact that harmonic functions minimize the Dirichlet energy [54] with $f : \Omega \rightarrow M$:

$$E(f) = \int_{\Omega} \|\nabla f\|^2_g,$$

where $g$ indicates the metric used on the surface $M$ and $\|\nabla f\|^2_g = \text{trace } g(\partial f, \partial f)$. The Laplace-Beltrami equation is a self-adjoint nonlinear 2nd order elliptic PDE and is the generalization of the usual Laplace operator to a smooth surface with the given metric. The system arising from the harmonic weights is symmetric and positive definite [54]. However, a given triangulation may lead to negative coefficients $\lambda_{ij}$ in the presence of obtuse angles. In that case the matrix may become positive semi-definite. The weights depend smoothly on the points $p_i \in \mathbb{R}^3$, i.e., the embedding.

Mean value weights are derived from an application of the mean value theorem for harmonic functions and defined as:

$$\lambda_{ij} = \frac{\tan(b_i/2) + \tan(c_i/2)}{\|p_i - p_j\|}.$$
Angles used in the construction of the mean value weights do not allow a symmetric system. However, the weights are guaranteed to be positive. Taken together with a convex embedding of the boundary of $M$ into $\partial \Omega$ this property guarantees an injective $\psi$. Assuming that all edge lengths are bounded away from zero, so are the $\lambda_{ij}$, which may help in the conditioning of the linear system of equation (2.2). As in the harmonic weights case, the weights depend smoothly on the points $p_i \in \mathbb{R}^3$.

3. Mesh Coarsening. Given that our linear systems are too large to begin with, hierarchy construction implies coarsening strategies. We borrow algorithms from the well developed area of mesh simplification for this purpose and give a brief background in this section.

The development of mesh simplification methods was initiated to deal with increasingly fine sampled surface meshes. Hoppe [35] introduced progressive meshes, which are built through a greedy strategy of topology preserving edge contractions [15] (see Figure 3.1), prioritized according to the geometric error introduced between the coarser mesh and the original mesh [22]. Criteria which also take triangle shape and curvatures into account were examined by Kobbelt and co-workers [41]. Such modifications may be particularly useful to ensure good mesh conditioning for the finite element method [4, 63].

![Fig. 3.1: Edge contraction removes one vertex, three edges, and two faces by making the end points of one edge coincident.](image)

Each edge contraction removes one vertex, three edges, and two triangles. Such a sequence of edge contractions can be “played” back and forward allowing for a traversal of a linear number of mesh approximations. A partial order of topological dependencies between the contractions—a given contraction cannot be undone before some neighboring contractions have been undone—can be used to induce a discrete set of “levels” of meshes between finest and coarsest [71, 36]. While the number of levels is “reasonable” in practice, no guarantees are made as to the asymptotic number of such levels. This is not surprising since the choice of simplification order and criteria in these algorithms depend on the geometry (embedding). Other simplification strategies such as vertex removal [61] as well as triangle removal [24].

Guarantees as to the number of levels of a fine to coarse hierarchy—and the cost of its construction—can be made by vertex removal strategies based solely on combinatorial considerations. Such methods have been employed in computational geometry for the construction of asymptotically optimal planar point location [40] and spatial geometric queries [16]. The so-called Dobkin-Kirkpatrick (DK) hierarchy is constructed by removing a maximal, independent set of vertices from a given triangle mesh (see Figure 3.2, left). A subset of vertices, $V_o \subset V$, is said to be independent if for any $v_i, v_j \in V_o$, $e_{ij} \notin E$. It is maximal if addition of any vertex $v_k \in V \setminus V_o$ to the set $V_o$ would violate its independence property. Each removed vertex leaves
a polygonal hole which is subsequently retriangulated. This process removes one vertex, three edges, and two triangles per independent vertex. Repeatedly removing a maximal independent set one arrives at a full hierarchy.

There are many different ways to implement the DK hierarchy. Selection of the maximal independent set of vertices is typically performed through some sweep and mark algorithm. The classical approach is purely combinatorial, i.e., it does not take any property of the embedding into account. This can lead to deteriorating aspect ratios as pointed out by Miller et al. [51]. To enable control over the quality of the coarser triangulation we prioritize half-edge contractions based on edge length (effectively removing the shortest edges first) to favor uniform triangle sizes at a given level of the hierarchy.

For a given edge we check the incident vertices for their status: unmarked, independent, or dependent. Initially all vertices are unmarked. If one incident vertex of an edge is independent, the other is marked dependent if still unmarked. Else we mark an unmarked vertex as independent and the other endpoint as dependent. If both endpoints are unmarked we favor the vertex of higher valence as independent. This biasing towards higher valence tends to result in smaller maximally independent sets of vertices. While this leads to slower decaying DK hierarchies it also leads to faster decay in our MIS hierarchy (see below). After the marking sweep, independent vertices are removed through a half-edge contraction into one of their dependent neighbors (there is always at least one such neighbor). For the construction of prolongation operators we also record all dependent vertices in the 1-ring of an independent vertex.

Using the Euler characteristic of a planar mesh one can show that $|V_o| \geq c|V|$ for some $1 > c \geq c_0 > 0$ for any planar mesh with $|V| > n_0 > 0$ and $c_0$ independent of the particular mesh. Four colorability of a planar graph ensures that $c_0 = 1/4$ is possible. Randomized greedy selection strategies can achieve this bound in empirical practice though the theoretical guarantees only assert $c_0 \geq 1/24$ for such strategies. These strategies favor low valence vertices for removal. With our strategy of favoring high valence vertices we observe an empirical $c_0 \approx .21$. Whatever the exact factor, it implies that a sequence of coarser meshes can be constructed with $J = O(\log |V|)$ levels.

Fig. 3.2: Marked vertices in the left mesh denote a maximal independent set. On the right a retriangulation of the independent set, which is used in our fast MIS hierarchy construction.

The decay rate—going from fine to coarse—of the thusly constructed hierarchies is much slower than the rates achieved in quadrasection refinement: $3/4$ compared to $.21$. In analogy to the regular refinement setting of coarse to fine mesh hierarchies one would like to use coarsening strategies that remove a larger fraction of all vertices.

Consider quadrasection refinement of a triangle mesh. In that case each old
vertex is surrounded by a set of new vertices in the finer mesh: the old vertices form a maximal independent set with respect to the finer mesh. This observation suggests the construction of a hierarchy which turns the DK removal strategy on its head. Instead of removing the maximal independent set we form the coarser level by removing all other vertices. We will refer to this novel hierarchy construction as the MIS hierarchy.

Exchanging dependent and independent markings the coarser level is once again built by performing half-edge contractions after a sweep and mark pass. Note that the prolongation list will always contain at least one vertex, but may not contain more than that. In practice we found that approximately 1/3 of all removed vertices fall into this category. For such vertices the prolongation matrices have rows with a single non-zero entry. To evaluate the impact of these degenerate rows on the solver convergence we consider a medium hierarchy in between the slow DK and fast MIS hierarchies. The medium hierarchy is constructed in exactly the same way as MIS but we prohibit the removal of vertices which have only a single entry on their prolongation list. Naturally, such a hierarchy decays slower then MIS but still faster then DK. We observed a decay rate of .51, compared to .21 and .80 for the MIS and DK hierarchies respectively. Figure 3.3 shows the decay rate for two models (other models exhibit the same rates).

Boundary vertices are further qualified in the removal strategy. A half-edge contraction is not performed if it would contract a boundary vertex into an interior vertex. This ensures a better representation of the boundary during coarsification. In some settings it may also be desirable to fix some boundary vertices, typically corners, as unremovable throughout the hierarchy. Other cases in which an edge contraction is not performed are those which would change the global topology of the mesh (see [15]). In some applications it may also be desirable to incorporate geometric measurements into the removal criteria. Examples include quality measures such as triangle aspect ratio or whether the coarser surface has self intersections. Such modifications can of course change the observed decay rates.

Fig. 3.3: The rate of decay in the number of DOFs for Model 2 (left) and Model 3 (right). Average decay rates of .21, .51, .80 are observed for fast (MIS), medium (modified MIS), and slow (DK) hierarchies respectively.
Alternative approaches such as geometrical coarsening of unstructured planar meshes [4, 8, 52] or algebraic coarsening techniques [38, 68] have been studied in the numerical PDE community. These techniques usually employ agglomeration and aggregation strategies. The approach taken by Bank and Xu [4] is fundamentally different from our approach and it is unique in the sense that they force an arbitrary unstructured mesh into a nonuniform and locally refined mesh, enabling them to impose a logically nested (but not physically nested) structure on the mesh. Except for the algebraic coarsening techniques these algorithms are generally more complicated than our approach and it is unclear whether they can be generalized to the non-planar setting. Purely algebraic techniques offer an alternative, albeit at the cost of giving up knowledge of the embedding in choosing the best coarsification steps.

4. Prolongation Operators. We now have three different hierarchies at our disposal, slow (DK), medium (modified MIS), and fast (MIS) and will examine their behavior vis-a-vis different multilevel preconditioning strategies. As in the standard multigrid framework, coarser representations of the finest level system matrix are formed algebraically by the triple matrix product (also known as the variational condition)

\[ A^{(j)} = (P_{j-1}^j)^T A^{(j-1)} P_{j-1}^j, \quad j = 1, \ldots, J, \]  

(4.1)

where \( j \) corresponds to a coarser level than \( j-1 \) and \( P_{j-1}^j \) is the prolongation operator from level \( j \) to \( j-1 \). \( P_{j-1}^j \) is of size \( N_{j-1} \times N_j \) where \( N_j \) denotes the number of DOFs at level \( j \). There are many possible predictor choices which can be used for the prolongation operator: centroid, inverse edge length, Guskov [30] (divided differences), Desbrun [14], or energy minimizing [69] predictors. Gieng reports [23] that the performance of the multilevel preconditioners is not dramatically affected by the predictor choice. Hence we use the simplest one: centroid prediction (see Figure 4.1) (for fourth order problems however the predictors of Guskov and Desbrun may be more appropriate).

![Fig. 4.1: For this example, the prolongation matrix row corresponding to \( v_i \) has entries \( 1/|1-R_i| = 1/5 \).](image)

The rows corresponding to the coarse DOFs will form an identity matrix \( I \) of size \( N_{j-1} \times N_{j-1} \) and a row in \( R_{j-1}^j \) corresponding to a newly introduced DOF—or vertex \( v_i \)—will contain entries equal to \( 1/|1-R_i| \) (where \( |1-R_i| \) is the number of DOFs in the prolongation list of \( v_i \)) in the appropriate columns. All other entries vanish.

\[ P_{j-1}^j = \begin{bmatrix} I \\ R_{j-1}^j \end{bmatrix}. \]  

(4.2)
5. Preconditioners. Most of the parameterization work published in the computer graphics literature has focused on different approaches to the formulation of the system matrix. Little effort has been devoted to effective numerical solvers, though the need to go beyond simple diagonal preconditioning has been pointed out repeatedly (see for example [47]). For example, the work of Liesen et al. [48] concentrates on a constrained minimization approach with single level preconditioning in the form of a Krylov subspace method (for more details see [11]). In contrast our preconditioners exploit a full multilevel hierarchy. Duchamp and co-workers [17] did use a full hierarchical approach to compute piecewise linear harmonic embeddings. They construct lazy wavelets [66, 65] induced by a DK hierarchy and consider a conjugate gradient solver in the wavelet domain. Empirically, this reduced the number of iterations from linear to logarithmic, similar to what is found when using a hierarchical basis [73] for the solution of 2nd order elliptic problems. In our construction we observe a constant number of iterations for the MG preconditioner as expected.

The systems arising in parameterization problems are well known to be ill-conditioned, with (bi-)conjugate gradient methods often failing for systems beyond $\approx 50,000$ DOFs. Some simple hierarchical preconditioning (naive BPX, see below) has been used since it is easy to integrate with mesh simplification approaches without the need to build coarser level system matrices (see, e.g., [42]). In our comparison of methods we will include this strategy.

Multilevel preconditioners are usually classified as either multiplicative or additive Schwarz methods. Multiplicative preconditioners are designed to update the residual at every level throughout the multilevel hierarchy, whereas the additive ones update the residual after a full sweep of the multilevel hierarchy. We are going to employ two multiplicative Schwarz methods: the multigrid (MG) and the hierarchical basis multigrid (HBMG) [3] preconditioners. Essentially MG and HBMG are the same preconditioner with the difference lying in the DOFs used for the smoothing iteration (see Algorithm 6.1). MG sweeps all DOFs at a given level, whereas HBMG sweeps the ones that are newly introduced at that level. Hence, HBMG is very attractive for adaptive regimes where the storage complexity is optimal and the computational complexity is close to optimal (in two spatial dimensions). Based on the decay rate of the hierarchies, different multilevel preconditioners are appropriate. Traditional MG is most appropriate for fast decaying hierarchies while the HBMG method is more appropriate for slow decaying hierarchies. (In the traditional refinement setting they are used for highly adaptive refinement with only a small constant number of DOFs added when going from coarse to fine.) In either case the system is solved with the help of the (bi-)conjugate gradient method as an outer accelerator. The iteration counts we report give the number of iterations of the (bi-)conjugate gradient method. The computational complexities of several multilevel preconditioners were discussed in detail in [1]. Both MG and HBMG act as standard V-cycle iteration with one symmetric Gauss-Seidel iteration as smoother.

Additive versions of MG and HBMG preconditioners are the Bramble-Pasciak-Xu (BPX) [6] and the hierarchical basis (HB) [73] preconditioners respectively. The action of the classical BPX preconditioner [1, 72] in 2D can be written in matrix form as:

$$X_{\text{BPX}} = \sum_{j=0}^{J-1} P_j S_j P_j^T + P_J (A_J)^{-1} P_J^T,$$  \hspace{1cm} (5.1)

where $(A_J)^{-1}$ represents a coarsest level direct solve. The prolongation matrix from
level $j$ is denoted by

$$P_j \equiv P_0^j = P_0 \ldots P_j^{-1} \in \mathbb{R}^{N_0 \times N_j}.$$ 

$P_0^j$ is the identity matrix $I \in \mathbb{R}^{N_0 \times N_0}$, $P_j^{-1}$ is the prolongation matrix from level $j$ to $j-1$, and $S_j$ is the smoother. The HB preconditioner in 2D can be expressed as:

$$X_{\text{HB}} = J^{-1} \sum_{j=0}^{J-1} H_j S_j H_j^T + H_J (A^J)^{-1} H_J^T,$$  \hspace{1cm} (5.2)

where the $H_j$ are the tails of the $P_j$ corresponding to newly introduced DOFs. In other words, with $J+1 = 0$, $H_j \in \mathbb{R}^{N_0 \times (N_j - N_{j+1})}$, $j = 0, \ldots, J$ is given by only keeping the columns that correspond to new DOFs (the last $N_j - N_{j+1}$ columns of $P_j$). Notice that the smoother $S_j$ in equation (5.2) acts only on the newly introduced DOFs.

The additive method which we denote as the naïve nBPX preconditioner is a variant of BPX. It takes only the finest level term in equation (5.1), drops the smoother, and the coarsest level direct solve:

$$X_{\text{naive}} = P_J P_J^T.$$ \hspace{1cm} (5.3)

As mentioned in Section 2, the harmonic and mean value weights give rise to symmetric positive definite and non-symmetric systems, hence the preconditioners above are coupled with conjugate gradient and bi-conjugate gradient methods respectively. The solvers for the system formed by using the harmonic weights are provably guaranteed to converge. We have full rank prolongation matrices and convergent smoothing iterations, then through the use of the variational conditions applied to the symmetric positive definite system the convergence of the solver is guaranteed [58].

All the above preconditioners are mesh oriented in the sense that the prolongation operators are created using the geometry and connectivity information of the mesh. An alternative would be the use of algebraic multigrid (AMG) methods [7] which do not use any geometric knowledge about the mesh, though the connectivity of the mesh enters at the finest level through the associated sparsity structure of the system matrix (2.2). AMG methods may be attractive for parameterization problems. However, the original AMG theory was developed for Stieltjes matrices (i.e., $A \in \mathbb{R}^{N_0 \times N_0}$ is symmetric positive definite, and has non-positive off-diagonal entries). While Stieltjes matrices are not a requirement for AMG to work, matrices $A$ which are far from being Stieltjes may render AMG less effective [7]. We do not pursue AMG methods further here, in particular since they do not provide explicit control over the construction and quality of the coarsification hierarchy. (For some recent work employing AMG for mesh partitioning and multilevel surface editing see the work of Clarenz et al. [9].)

6. Sparsity. When comparing different preconditioning strategies one can evaluate their performance in terms of the number of iterations required by the solver. However, this does not tell the whole story since the time to solution is a function of both the number of iterations and the sparsity structure of the matrices appearing in the multilevel hierarchy. In this section we take a closer look at the fill-in of the matrices involved in the actions of MG and HBMG preconditioners. (Note that the question of sparsity does not arise for the nBPX since it does not use any matrix structures.) The sparsity will help to explain the overall superior performance of the MG preconditioner in conjunction with the fast (MIS) hierarchy.
To evaluate the sparsity we consider the average number of non-zero entries per row. We observe relatively small and uniform standard deviation of this measure making it a good predictor of performance. We plot the number of non-zero entries in the matrices formed by the variational conditions (equation 4.1) for fast, medium, and slow hierarchies in Figure 6.1. For fast and medium hierarchies, fill-in increases approximately linearly with level (two models are shown, others exhibit the same behavior). The slow hierarchy fill-in is also linear with respect to level over a wide range. But it is very rapid and the curves flatten out as the matrices become dense. At the finest level, the matrices have seven non-zeros on average as a consequence of the Euler characteristic. At the coarsest level, the fast, medium and slow hierarchies produce an average of 12, 30 and 190 non-zero entries. This easily indicates that sparsity is the primary factor which is responsible for the poor performance of slow hierarchies. Also note how the medium hierarchy leads to significantly more (almost three times more) fill-in compared to the fast hierarchy.

For slow decaying hierarchies it is more appropriate to use the HBMG preconditioner which only operates on some of the DOFs at each level. Let $A^{(j)}$ be represented by a two-by-two block form [2]:

$$A^{(j-1)} = \begin{bmatrix} A^{(j)} & A_{12}^{(j-1)} \\ A_{21}^{(j-1)} & A_{22}^{(j-1)} \end{bmatrix}$$

(6.1)

where $A^{(j)}, A_{12}^{(j-1)}, A_{21}^{(j-1)},$ and $A_{22}^{(j-1)}$ correspond to coarse-coarse, coarse-fine, fine-coarse, and fine-fine interactions respectively. The HBMG method utilizes a change of basis operation resulting in the so-called stabilized blocks (represented with hats). Dropping the superscripts for simplicity, the blocks are expressed as:

$$\begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} = \begin{bmatrix} I & R^T \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ R & I \end{bmatrix},$$

Fig. 6.1: Sparsity of the system matrix for Models 2 and 3 using fast, medium, and slow hierarchies (left). Close up of the sparsity between levels 0 and 6 (right). Notice the rapid fill-in corresponding to the slow hierarchy.
In terms of these stabilized blocks the HBMG algorithm can be interpreted as an iterative process for solving the system \((2.2)\) where \(x = P\hat{x}, \hat{b} = P^Tb\).

**Algorithm 6.1.**
1. smooth \(A_{22}\hat{x}_2 = \hat{b}_2\)
2. form residual \(\hat{r}_1 = \hat{b}_1 - (\hat{A}_{11}\hat{x}_1) - \hat{A}_{12}\hat{x}_2\)
3. solve \(\hat{A}_{11}\hat{x}_1 = \hat{r}_1\)
4. prolongate \(\hat{x} = \hat{x} + P\hat{x}_1\)
5. smooth \(A_{22}\hat{x}_2 = \hat{b}_2 - (\hat{A}_{21}\hat{x}_1)\)

This can be further simplified by transforming the linear system \((2.2)\) into the equivalent system

\[
A(x - x_j) = b - Ax_j,
\]

with an initial guess of \(x_j\). In this setting, the initial guess is zero, and the HBMG algorithm recursively iterates towards the error with given residual on the right hand side. In that case the terms in parentheses in Algorithm 6.1 are zero.

In the HBMG method—a multiplicative preconditioner—the residual is computed at every level (see step 2 in Algorithm 6.1), therefore the sparsity of the \(A_{12}\) block plays a crucial role. Similarly, the cost of the smoothing step is related to the sparsity of \(A_{22}\). Figure 6.2 shows the sparsity of the \(A_{12}\) and \(A_{22}\) blocks. As before the fill-in increases linearly with level over a wide range. Eventually the matrices become quite dense so that the curves flatten out, similar to what we observed for \(A\) in the slow hierarchies. At the finest level, \(A_{12}\) contains six non-zero entries and \(A_{22}\) has only one non-zero. As the hierarchy reaches the coarsest level, HB stabilization produces 173 and 38 non-zeros in the \(A_{12}\) and \(A_{22}\) blocks respectively. The ratio of non-zero entries in \(A_{12}\) over \(A_{22}\) is roughly 4.5 with an almost identical slope. This ratio is due to the \(A_{12}\) block having many initial non-zero entries before HB stabilization is in effect.
7. Numerical Experiments. The preconditioners employed in this paper have been implemented as library extensions to the freely available Finite Element ToolKit (FETk) [34] (more information about FETk can be found at: http://www.fetk.org/). The code was compiled using gcc-2.96 with O2 optimization and all timings were taken on a 2.8GHz P4 Xeon with 4GB of RAM running Linux.

We present experiments with six different unstructured surface meshes (see the images in Figure 7.1) of varying size (see Table 7.1). Note the rapidly deteriorating condition numbers as the number of DOFs increases.

Fig. 7.1: Models used: Igea face (Model 5), skull (Model 6), David face (Model 2) and David head (Models 1, 3, and 4 with descending order in the number of DOFs).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1-David head</td>
<td>579649</td>
<td>4</td>
<td>8</td>
<td>17</td>
<td>1.22e−5</td>
<td>3.19e+4</td>
<td>2.60e+9</td>
</tr>
<tr>
<td>2-David face</td>
<td>223654</td>
<td>4</td>
<td>7</td>
<td>21</td>
<td>1.29e−4</td>
<td>2.26e+3</td>
<td>1.74e+7</td>
</tr>
<tr>
<td>3-David head</td>
<td>94220</td>
<td>4</td>
<td>7</td>
<td>22</td>
<td>7.61e−5</td>
<td>2.82e+2</td>
<td>3.71e+6</td>
</tr>
<tr>
<td>4-David head</td>
<td>46094</td>
<td>4</td>
<td>5</td>
<td>18</td>
<td>1.44e−4</td>
<td>2.03e+2</td>
<td>1.40e+6</td>
</tr>
<tr>
<td>5-Igea face</td>
<td>8038</td>
<td>3</td>
<td>4</td>
<td>16</td>
<td>4.23e−3</td>
<td>2.34e+2</td>
<td>5.53e+4</td>
</tr>
<tr>
<td>6-Skull</td>
<td>1203</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2.70e−2</td>
<td>3.54e+1</td>
<td>1.31e+3</td>
</tr>
</tbody>
</table>

Table 7.1: Collection of models used in our experiments showing the number of levels used for fast, medium, and slow hierarchies, the maximum and minimum eigenvalue and the resulting condition number (for harmonic weights).

The goal of the numerical experiments is to determine the minimum solve time as a function of hierarchy and preconditioner used. Another factor concerns the sparse-cutoff, i.e., the level of the hierarchy at which a direct solve (such as SuperLU [12]) is performed. This is typically not the coarsest level the hierarchy creation could produce. Among all possible levels, we numerically choose the coarsest level with respect to the best solve time. The number of levels reported in Table 7.1 is obtained by the sparse-cutoff (in fast hierarchy we typically find a cutoff after four levels). In our experiments the solver iteration was stopped when the norm of the residual fell below 5.0e−5. Vector entries are expressed by double datatypes, whereas matrix entries are of float type. Interestingly, going from double to float precision only reduced runtimes by 5% for the largest model, indicating that the runtime is dominated by memory latency, not bandwidth issues.
Figure 7.2 shows all nine combinations of fast, medium, and slow hierarchies with MG, HBMG, and nBPX preconditioners acting on both symmetric (harmonic weights) and non-symmetric (mean-value weights) systems. For each preconditioner the thick line indicates the winning hierarchy. For example, MG with the fast hierarchy outperforms MG with medium or slow hierarchies. HBMG timings are comparable for fast and medium hierarchies, but asymptotically medium hierarchy timings are favorable (see also the Model 1 timings in Table 7.2). HBMG is designed for adaptive meshes and slow hierarchies are the closest to that pattern. Among the preconditioners used with the slow hierarchy, HBMG turns out to be the most effective as expected. nBPX favors the fast hierarchy, making the fast hierarchy the best for all preconditioners. The least effective preconditioner turns out to be nBPX especially with the slow hierarchy. Although, nBPX is effective on small systems, it rapidly loses its performance on large systems even though each iteration is very cheap. The main advantage of nBPX—and reason for its use—is its implementation simplicity. All matrix/vector operations can be implemented directly on the mesh datastructure with no need to construct matrices explicitly or compute triple matrix products. With the fast hierarchy, MG becomes such an effective preconditioner that it outperforms all other preconditioners on all the models. This holds true for both symmetric and nonsymmetric systems.

The geometric decay rate in the number of DOFs produces $O(N)$ solve times for MG and $O(N \log N)$ for HBMG and nBPX albeit with very different constants. This can be seen in the number of iterations which remain constant for MG whereas a logarithmically growing count can be observed for HBMG and nBPX (see Figure 7.3). This is very much the same behavior as observed for elliptic PDEs, which comes as no surprise when the parameterization is obtained by discretizing the Laplace-Beltrami operator. We remark that the iteration counts of nBPX are an order magnitude larger than those for HBMG. This difference is the primary reason why sophisticated preconditioners become superior to the naïve ones even though they require significantly more work per iteration. In order to exhibit the striking performance difference between the preconditioners and the hierarchies used, we insert the results for the largest system in Table 7.2. Notice that iteration counts for the nonsymmetric system are two to three times less than the symmetric one. This is particularly noticeable for the HBMG experiments. The slow hierarchies consistently give rise to the least number of iterations for MG and HBMG preconditioners. However, the depth of the slow hierarchies create large fill-in as discussed in Section 6. This fill-in makes slow hierarchies the most expensive to use when considering the overall timings.

8. Conclusion. Using different coarsening hierarchies, we presented a systematic performance analysis for MG, HBMG, and nBPX preconditioners on systems arising from the parameterization of unstructured surface meshes. The parameterization schemes of interest use harmonic weights and mean value weights giving symmetric and nonsymmetric systems respectively. Iteration counts indicate that for the same preconditioner a better conditioning of the system is obtained by using more levels, i.e., a slow (DK) hierarchy. But since the cost of one iteration on the slow hierarchy is relatively expensive, fast hierarchies result in the best runtimes.

Our experiments consistently show that the fast hierarchy is favorable over medium and slow hierarchies for all preconditioners, with the best results achieved when using the MG preconditioner. If the slow hierarchy is chosen, the best solve time is achieved by the HBMG preconditioner. We should emphasize that we observed convergence in all the experiments and no restarts for the bi-conjugate gradient method. The prov-
Fig. 7.2: CPU time for symmetric (top) and nonsymmetric (bottom) systems.
Fig. 7.3: Iteration counts for symmetric (left) and nonsymmetric (right) systems.

<table>
<thead>
<tr>
<th>Method</th>
<th>Sym time</th>
<th>Nsym time</th>
<th>Sym iter</th>
<th>Nsym iter</th>
</tr>
</thead>
<tbody>
<tr>
<td>fast+MG</td>
<td>37.3</td>
<td>55.9</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>med+MG</td>
<td>65.0</td>
<td>70.7</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>med+HBMG</td>
<td>184.0</td>
<td>133.8</td>
<td>100</td>
<td>37</td>
</tr>
<tr>
<td>fast+HBMG</td>
<td>230.1</td>
<td>149.2</td>
<td>141</td>
<td>44</td>
</tr>
<tr>
<td>slow+HBMG</td>
<td>326.8</td>
<td>356.0</td>
<td>60</td>
<td>35</td>
</tr>
<tr>
<td>slow+MG</td>
<td>370.2</td>
<td>500.2</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>fast+nBPX</td>
<td>597.9</td>
<td>877.5</td>
<td>1156</td>
<td>869</td>
</tr>
<tr>
<td>med+nBPX</td>
<td>656.6</td>
<td>960.8</td>
<td>1171</td>
<td>838</td>
</tr>
<tr>
<td>slow+nBPX</td>
<td>1132.7</td>
<td>1202.3</td>
<td>1357</td>
<td>696</td>
</tr>
</tbody>
</table>

Table 7.2: Solve time in seconds (ascending order in time) and iteration counts for the system with 579,649 DOFs (Model 1).

able guarantee of convergence (for the harmonic weights) is evidenced. In contrast, the (bi-)conjugate gradient solver does not converge at all for large models when no preconditioning is used. We conclude that the preconditioners described in this paper are robust.

There are a number of avenues for interesting future work. For example, a comparison of our work to AMG approaches would be of great interest. A more complete analysis of convergence properties for the mean value weights, which do not arise from the discretization of an elliptic PDE, would also be desirable. Finally, we look forward to applying our solvers to surface parameterization problems involving more than a single disk-like region.
Acknowledgments. The authors would like to thank M. Holst for providing FEtk, S. Bond for his help on the components of the preconditioner code, and I. Guskov for parts of the parameterization and coarsening code. We would also like to thank them for many enlightening discussions. The David head model is courtesy of the Digital Michelangelo Project at Stanford University. The Igea and skull models are courtesy of Cyberware, Inc. and Headus, Inc. respectively.

REFERENCES


18

B. AKSOYLU, A. KHODAKOVSKY, AND P. SCHRÖDER


