PERFORMANCE OF ALGEBRAIC MULTIGRID METHODS FOR NON-SYMMETRIC MATRICES ARISING IN PARTICLE METHODS

BENJAMIN SEIBOLD

Abstract. Large linear systems with sparse, non-symmetric matrices arise in the modeling of Markov chains or in the discretization of convection-diffusion problems. Due to their potential to solve sparse linear systems with an effort that is linear in the number of unknowns, algebraic multigrid (AMG) methods are of fundamental interest for such systems. For symmetric positive definite matrices, fundamental theoretical convergence results are established, and efficient AMG solvers have been developed. In contrast, for non-symmetric matrices, theoretical convergence results have been provided only recently. A property that is sufficient for convergence is that the matrix be an M-matrix. In this paper, we present how the simulation of incompressible fluid flows with particle methods leads to large linear systems with sparse, non-symmetric matrices. In each time step, the Poisson equation is approximated by meshfree finite differences. While traditional least squares approaches do not guarantee an M-matrix structure, an approach based on linear optimization yields optimally sparse M-matrices. For both types of discretization approaches, we investigate the performance of a classical AMG method, as well as an AMLI type method. While in the considered test problems, the M-matrix structure turns out not to be necessary for the convergence of AMG, problems can occur when it is violated. In addition, the matrices obtained by the linear optimization approach result in fast solution times due to their optimal sparsity.

1. Introduction

The efficient and robust numerical approximation of the Poisson equation is a key problem in many applications. In some special cases (such as spectral methods that use the fast Fourier transform), the Poisson equation can be solved in a single step. However, in most scenarios, the solution process is done in two consecutive steps: first, a discretization transforms the Poisson equation into a finite linear system; second, the linear system is solved. While in most investigations on multigrid methods the discretization is assumed as canonically given, in this paper we present an application in which many possible discretization strategies exist, and the interplay between discretization and performance of multigrid solvers is of interest.

We consider Poisson problems that arise in Lagrangian particle methods for the simulation of incompressible fluid flows. Particle methods solve the governing fluid equations on a set of nodes that move with the flow. Thus the discretization is done in the more natural Lagrangian frame of reference. As a consequence, convective terms are solved exactly. One of the earliest Lagrangian particle methods is smoothed particle hydrodynamics (SPH) [29, 18], which was first applied to astrophysical fluid dynamics, and later expanded to a wider class of hydrodynamic equations [32, 33]. Generalization of the SPH approach, that use moving least squares approximations for derivatives, have been presented by Dilts [12] and Kuhnert [22]. Both approaches have been successfully applied to problems with complex, time-dependent geometries [23] and free surface flows [24]. Further examples are shown in [14]. The problems considered in this paper are mostly formulated in the terminology of those latter two approaches. However, many aspects transfer to other types of particle methods. A fundamental difficulty with Lagrangian particle methods is that one gives away the control over the positions of the computational nodes. While one can
(and in fact should) remove particles where not desired, and insert new particles where required, the points are in general completely unstructured and without connections between each other.

A classical solution strategy for incompressible flows is the Chorin projection method [7], which has been successfully applied to Lagrangian particle methods [9, 10]. In each time step, the flow is first evolved neglecting pressure, which yields a velocity field that is not incompressible. Then the velocity field is modified by a gradient field that is chosen so that incompressibility is restored. This correction step requires the solution of a Poisson equation. As noted above, the arrangement of the particles is in a general, unstructured cloud of points. And the Poisson equation has to be approximated on this geometry.

One possibility to approximate the Poisson equation is using a finite element approach. For instance, one can generate a triangulation [46] (or another type of mesh) between the points, and apply a traditional finite element approach [47]. However, mesh generation is an expensive task, especially in 3d, and a good mesh generator may not always be available. In the context of particle methods, one rarely is willing to make such an investment. An alternative is the use of meshfree finite element approaches, such as the element-free Galerkin method (EFGM) [5], or the partition of unity finite element method (PUFEM) [3, 4]. Similar approaches [19] have been applied in a multilevel context [20]. A key advantage of all these approaches is that they are based on inner products between (meshfree) basis functions, hence they lead to symmetric linear systems to be solved. A fundamental drawback is that the computation of the inner products requires quadratures, which in many cases can only be achieved by introducing a regular background grid. This step is not only technically inconvenient, it is also costly. Another difficulty with finite element approaches is the treatment of general boundary conditions, which arise in multi-phase flows in complex, time-dependent domains.

In contrast, meshfree finite difference methods do not require basis functions or quadrature, and boundary conditions can be implemented directly. At each point, the approximation of the Laplace operator is based on a local Taylor expansion. A first approach by Perrone and Kao [36] was later extended by Liszka, Orkisz, et al. [28, 11] to the generalized finite difference method (GFDM). A detailed presentation of common particle methods and meshfree approaches is provided in [43]. In this paper, we outline the fundamental conditions for a consistent approximation of the Laplace operator in Sect. 2.

A key drawback of meshfree finite difference approaches is that the resulting linear systems are in general non-symmetric. As we explain in Sect. 3 this non-symmetry cannot be overcome by any simple means. In the same section, we outline theoretical results on the convergence of algebraic multigrid approaches for non-symmetric matrices. A property sufficient for convergence is an M-matrix structure. We present fundamental properties of M-matrices and their relevance for multilevel methods in Sect. 4.

In addition to the non-symmetry of the meshfree finite difference matrices classical least squares approaches, as outlined in Sect. 3 involve two other difficulties. First, the arising matrices are about six times as dense as finite difference matrices on rectangular grids. Second, an M-matrix structure is not guaranteed, and is—unless the point geometry is sufficiently nice—in general violated. Unlike the non-symmetry, these latter two problems can be overcome, by a new approach that is based on linear sign-constrained optimization. This approach is presented in Sect. 5, and conditions for its successful application are provided in Sect. 6. In Sect. 7 theoretical predictions about the computational cost to generate the Poisson matrices with least squares vs. linear optimization approaches are given. In Sect. 8 we numerically investigate the performance of various multigrid approaches for the matrices constructed by the different types of minimization approaches. Conclusions and an outlook are given in Sect. 9.
2. Meshfree Finite Differences for the Poisson Equation

Consider the Poisson equation to be solved inside a domain $\Omega \subset \mathbb{R}^d$

$$\begin{cases}
-\Delta u = f & \text{in } \Omega \\
u = g & \text{on } \Gamma_D \\
\frac{\partial u}{\partial n} = h & \text{on } \Gamma_N
\end{cases}$$

(1)

where $\Gamma_D \cup \Gamma_N = \partial \Omega$. We consider the case of $\Omega$ possessing a unique smooth solution.

Let a point cloud $X = \{\vec{x}_1, \ldots, \vec{x}_n\} \subset \overline{\Omega}$ be given, which consists of interior points $X_i \subset \Omega$ and boundary points $X_b \subset \partial \Omega$. The point cloud is meshfree, i.e. no information about connection of points is provided. Meshfree finite difference approaches convert problem (1) into a linear system

$$A \cdot \vec{u} = \vec{f},$$

(2)

where the vector $\vec{u}$ contains approximations to the values $u(x_i)$. The $i$-th row of the matrix $A$ consists of the stencil corresponding to the point $\vec{x}_i$.

In the following, we derive conditions on finite difference stencils to yield an approximation to the Laplace operator. Consider a function $u \in C^2(\Omega \subset \mathbb{R}^d, \mathbb{R})$. At each interior point $\vec{x}_0$, we wish to approximate $\Delta u(\vec{x}_0)$ using the function values of a finite number of points in a circular neighborhood $(\vec{x}_0, \vec{x}_1, \ldots, \vec{x}_m) \in B(\vec{x}_0, r)$, where $B(\vec{x}_0, r) = \{\vec{x} \in \overline{\Omega} : ||\vec{x} - \vec{x}_0||_2 < r\}$. Note that the Euclidean norm is a natural choice in meshfree methods due to its rotational symmetry. Define the distance vectors $\vec{s}_i = \vec{x}_i - \vec{x}_0 \forall i = 0, \ldots, m$. The function value at each neighboring point $u(\vec{x}_i)$ is given by a Taylor expansion as

$$u(\vec{x}_i) = u(\vec{x}_0) + \nabla u(\vec{x}_0) \cdot \vec{x}_i + \frac{1}{2} \nabla^2 u(\vec{x}_0) : (\vec{x}_i \cdot \vec{x}_i^T) + e_i .$$

Here we use the Frobenius inner product $A : B = \sum_{i,j} a_{ij} b_{ij}$. The error in the expansion is of order $e_i = O(r^3)$. A linear combination with coefficients $(s_0, \ldots, s_m)$ equals

$$\sum_{i=0}^{m} s_i u(\vec{x}_i) = u(\vec{x}_0) \left( \sum_{i=0}^{m} s_i \right) + \nabla u(\vec{x}_0) \cdot \left( \sum_{i=1}^{m} s_i \vec{x}_i \right) + \nabla^2 u(\vec{x}_0) : \left( \frac{1}{2} \sum_{i=1}^{m} s_i (\vec{x}_i \cdot \vec{x}_i^T) \right) + \left( \sum_{i=1}^{m} s_i e_i \right).$$

This approximates the Laplacian, i.e. $\sum_{i=0}^{m} s_i u(\vec{x}_i) = \Delta u(\vec{x}_0) + O(r^3)$, if exactness for constant, linear, and quadratic functions is satisfied

$$\sum_{i=0}^{m} s_i = 0 \quad , \quad \sum_{i=1}^{m} \vec{x}_i s_i = 0 \quad , \quad \sum_{i=1}^{m} (\vec{x}_i \cdot \vec{x}_i^T) s_i = 2I .$$

(3)

**Definition 1.** A stencil $(s_0, \ldots, s_m)$ to a set of points $(\vec{x}_0, \vec{x}_1, \ldots, \vec{x}_m) \in B(\vec{x}_0, r)$ is called consistent (with the Laplace operator), if the constraints (3) are satisfied.

The linear and quadratic constraints in (3) can be formulated as a linear system of equations

$$V \cdot \vec{s} = \vec{b},$$

(4)

where $V \in \mathbb{R}^{k \times m}$ is the Vandermonde matrix given by $\vec{x}_1, \ldots, \vec{x}_m$, and $\vec{s} \in \mathbb{R}^m$ is the stencil vector. The number of constraints is $k = \frac{d(d+3)}{2}$. The constant constraint in (3) yields $s_0 = -\sum_{i=1}^{m} s_i$. Neumann boundary points can be treated in a similar manner. Approximating $\frac{\partial u}{\partial n}(\vec{x}_0)$ by $\sum_{i=0}^{m} s_i u(\vec{x}_i)$ leads to the constraints

$$\sum_{i=0}^{m} s_i = 0 \quad , \quad \sum_{i=1}^{m} \vec{x}_i s_i = \vec{n} .$$

(5)
For each point, a meshfree finite difference approximation consists of two steps: First, define which points are its neighbors. Typically, more neighbors than constraints are chosen. Second, select a stencil. If (3) is underdetermined, a minimization problem is formulated to select a unique stencil. The results in Sect. 6 guarantee the existence of a solution to the above linear systems (3) and (5).

**Definition 2.** A consistent stencil \((s_0, \ldots, s_m)\) is called **minimal**, if \(m \leq k\).

Minimal stencils are beneficial for the sparsity of the system matrix, resulting in a minimal memory consumption and a fast application of the matrix. In fact, the total number of neighboring points is proportional to the effort of multiplying the matrix with a vector. This promises a fast solution of system (2) with (semi-)iterative linear solvers, assumed the convergence properties are reasonable. More importantly, having minimal stencils results in an optimally sparse matrix graph. This property can be advantageous with respect to coarsening in algebraic multigrid methods.

**Definition 3.** A consistent stencil \((s_0, \ldots, s_m)\) is called **positive**, if \(s_1, \ldots, s_m \geq 0\). Due to (3) and (5), this implies for the central point \(s_0 < 0\).

Having positive stencils implies that the system matrix in (2) is an L-matrix (Def. 4), which gives rise to an M-matrix structure (see Sect. 4). The desirability of positive stencils has been pointed out by Demkowicz, Karafiat and Liszka [11]. Classical approaches do in general not yield positive stencils. An “optimal star selection” [15] makes positive stencils likely, but they are not guaranteed. In Sect. 5 we present a strategy that approximates the Poisson equation (1) on a point cloud by minimal positive stencils. In Sect. 6 conditions on a point cloud are given which guarantee the existence of positive stencils.

### 3. Non-Symmetric Matrices and Algebraic Multigrid

Meshfree finite difference approaches, as described in Sect. 2, approximate the Poisson equation by non-symmetric matrices. Large linear systems with non-symmetric matrices have received considerable attention in Markov chain modeling. For instance, the computation of a page ranking in internet search engines leads to sparse non-symmetric matrices [26]. The arising matrices typically have an M-matrix structure. In the approximation of partial differential equations, non-symmetric matrices traditionally arise when convective terms are approximated. As described in [49], classical AMG methods frequently are observed to converge for such non-symmetric matrices, if the advection does not dominate the problem. When approximating the Poisson equation, traditional approaches lead to symmetric matrices. In contrast, meshfree finite difference approaches approximate the Poisson equation by non-symmetric matrices. At first glance this sounds surprising, since the Laplace operator itself is self-adjoint. However, it can be proved and explained.

The formal argument for non-symmetry is as follows. Consider two points \(\vec{x}_i\) and \(\vec{x}_j\), each being a neighbor of the other, and a third point \(\vec{x}_k\) which is a neighbor of \(\vec{x}_i\) but not a neighbor of \(\vec{x}_j\). Since each stencil entry depends on *all* its neighbors, the point \(\vec{x}_k\) influences the matrix entry \(a_{ij}\), but not the matrix entry \(a_{ji}\).

The non-symmetry of the matrix can also be understood in that a self-adjoint operator is approximated on a discrete set of points that is unstructured, and thus does not possess any local symmetry. Note that on unstructured geometries, classical finite element approaches [17] approximate the Laplacian by a non-symmetric matrix as well. While in the finite element system \(L \vec{u} = M \vec{f}\), the stiffness matrix \(L\) alone is always symmetric, the actual approximation matrix \(M^{-1}L\) is in general non-symmetric.

The non-symmetry is inherent in the finite difference approximation, and to the author’s knowledge, there is no straightforward way to symmetrize the linear system (2). This has interesting implications for the choice of efficient solvers for the arising linear systems. For instance, in the
realm of (preconditioned) Krylov subspace methods, the classical conjugate gradient method has to be replaced by more costly solvers, such as GMRES [41] or BiCGstabl [50].

Of fundamental interest are multigrid solvers for the arising systems (2). In most practical applications with a sufficient number of particles, the largest part of the computational time is spent on the generation and solution of the linear systems in the pressure correction (and possibly in an implicit treatment of viscosity, which leads to quite similar system matrices). Hence it is of crucial interest to achieve the optimal computational effort $O(n)$, where $n$ is the number of particles, that is promised by multigrid methods.

In the context of particle methods, algebraic multigrid (AMG) approaches are most promising, since the application of geometric multigrid approaches on an unstructured, non-connected cloud of points is very difficult. While in the symmetric matrix case, various convergence results for AMG are well known [6, 48, 49], theoretical results for the non-symmetric matrix are much fewer in number. An analysis of AMG for non-symmetric matrices has been given by Notay [34, 35]. A complementary analysis for methods of algebraic multilevel iteration (AMLI) type [1, 2] has been given by Mense and Nabben [30, 31]. AMLI methods are based on a block incomplete factorization of the matrix, partitioned in hierarchical form. Many AMG approaches can be interpreted as AMLI methods.

A common assumption in all the above theoretical results is that the matrix be an M-matrix (see Sect. 4). The M-matrix structure implies certain properties that positive definiteness would imply for symmetric matrices. Since the negative Laplace operator in (1) is positive definite, the M-matrix structure is natural for approximations of it. In above references, proofs of convergence of two-grid methods for the M-matrix case are given; some results can be extended to the full multigrid case. In particular, it yields a discrete maximum principle, and guarantees the convergence of the Gauß-Seidel iteration (see Sect. 4.2). In the above referenced theoretical results, the M-matrix structure is sufficient for the convergence of AMG and AMLI methods. The importance of the M-matrix structure for AMG is discussed in [34]. The analysis of the two-grid case is given in [35]. A key step in the generalization of theoretical AMG results to the non-symmetric matrix case is the suitable understanding of the size of an eigenvalue as its modulus in the complex case. In Sect. 4.3, recent theoretical results on the convergence of AMLI methods in the M-matrix case are outlined.

Of course, the M-matrix property is not necessary for the convergence of multigrid approaches. In fact, convergence is frequently observed even if the M-matrix structure is violated. However, unless theoretical results for weaker assumptions are present, one cannot be sure. This would not be too problematic if a single Poisson equation had to be solved. However, if the Poisson solves are just substeps in a larger computation (as it is the case in the particle method), it is unsatisfactory to not have a guarantee of convergence. In fact, one may observe a rapid AMG convergence in the beginning of a long computation (when the particles are distributed nicely), but convergence may deteriorate as the particles become less evenly distributed. It is our hope that for meshfree finite difference methods, the M-matrix property can be a key instrument in successful and efficient solution approaches for the Poisson equation. In addition, it would be desirable to obtain an understanding of the connection between the performance of AMG and the corresponding particle geometry.

4. Properties of M-Matrices

In this section, we provide a brief overview of the definition of the M-matrix property, a practical sufficient condition that relates to meshfree finite difference approximations, and some well-known benefits of the M-matrix structure.

Definition 4. A square matrix $A = (a_{ij})_{ij} \in \mathbb{R}^{n \times n}$ is called Z-matrix, if $a_{ij} \leq 0 \forall i \neq j$. A Z-matrix is called L-matrix, if $a_{ii} > 0 \forall i$. 

We write $A \geq 0$ if $a_{ij} \geq 0 \ \forall i, j$. The same notation applies to vectors.

**Definition 5.** A regular matrix $A$ is called *inverse positive*, if $A^{-1} \geq 0$.

**Definition 6.** A Z-matrix is called *M-matrix*, if it is inverse positive.

We use the M-matrix property, since it yields a sufficient condition for inverse positivity.

**4.1. A Sufficient Condition for the M-Matrix Structure.** Conditions that imply the M-matrix property are required, since the inverse matrix is typically not directly available. Let the unknowns be labeled by an index set $I$. We consider square matrices $A \in \mathbb{R}^{I \times I}$.

**Definition 7.** The *graph* $G(A)$ of a matrix $A$ is defined by $G(A) = \{(i, j) \in I \times I : a_{ij} \neq 0\}$. The index $i \in I$ is called connected to $j \in I$, if a chain $i = i_0, i_1, \ldots, i_k = j \in I$ exists, such that $(i_{\nu-1}, i_{\nu}) \in G(A) \ \forall \nu = 1, \ldots, k$.

**Remark 1.** For a finite difference matrix, each index $i \in I$ corresponds to a point $\vec{x}_i$. The index $i \in I$ being connected to $j \in I$ means that the point $\vec{x}_i$ connects (indirectly) to the point $\vec{x}_j$ via stencil entries.

**Definition 8.** A finite difference matrix is called *essentially irreducible* if every point is connected to a Dirichlet boundary point.

**Remark 2.** A finite difference matrix that is not essentially irreducible is singular, since the points that are not connected to a Dirichlet point form a singular submatrix.

**Definition 9.** A matrix $A \in \mathbb{R}^{I \times I}$ is called *essentially diagonally dominant* if it is weakly diagonally dominant ($\forall i \in I : |a_{ii}| \geq \sum_{k \neq i} |a_{ik}|$), and every point $i \in I$ is connected to a point $j \in I$ which satisfies the strict diagonal dominance relation $|a_{jj}| > \sum_{k \neq j} |a_{jk}|$.

**Theorem 1.** An L-matrix arising as a finite difference discretization of (1) is essentially diagonally dominant if it is essentially irreducible.

*Proof.* For an L-matrix the constant relation in (3) implies the weak diagonal dominance relation for every interior and Neumann point. Each row corresponding to a Dirichlet point satisfies the strict diagonal dominance relation. □

**Theorem 2.** An essentially diagonally dominant L-matrix is an M-matrix.

*Proof.* The proof is given in [21, p. 153]. □

If problem (1) can be discretized by positive stencils and every point is connected to a Dirichlet point, then the resulting matrix is an M-matrix.

**4.2. Some Benefits of the M-Matrix Structure.** The Poisson equation satisfies maximum principles. For instance, consider (1) with Dirichlet boundary conditions only. If $f \leq 0$ and $g \leq 0$, then the solution satisfies $u \leq 0$ [10]. A discretization by an M-matrix mimics this property in a discrete maximum principle.

**Theorem 3.** Let $A$ be an M-matrix. Then $A\vec{x} \leq 0$ implies $\vec{x} \leq 0$. Conversely, a Z-matrix satisfying $A\vec{x} \leq 0 \Rightarrow \vec{x} \leq 0$ is an M-matrix.

*Proof.* A is an M-matrix, thus $A^{-1} \geq 0$ by definition. Let $\vec{y} = A\vec{x}$. Then $\vec{x} = A^{-1}\vec{y}$. The component-wise inequalities $A^{-1} \geq 0$ and $\vec{y} \leq 0$ imply $\vec{x} \leq 0$. The reverse statement is proved in [37, p. 29]. □

**Theorem 4.** If $A$ is an M-matrix, and $D$ its diagonal part, then $\rho(I - D^{-1}A) < 1$, thus the Jacobi and the Gauß-Seidel iteration converge.
Proof. The convergence of the Jacobi iteration is given in [21]. The Gauß-Seidel convergence follows from the Stein-Rosenberg-Theorem [52]. □

Further classical benefits of an M-matrix structure with respect to linear solvers can be found in [52].

4.3. Relevance of the M-Matrix Structure for Multilevel Methods. Recent theoretical convergence results for AMG in the non-symmetric M-matrix case have been given in [34, 35], and for AMLI type approaches in [30, 31]. Here, we outline the key results for the latter type of approaches. The M-matrix structure is a sufficient condition for the convergence of additive Schwarz (AMLI) methods for non-symmetric matrices. Consider the system matrix be partitioned into blocks

\[ A = \begin{bmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{bmatrix}, \]

where \( F \) denotes the fine grid unknowns, and \( C \) the coarse grid unknowns. If \( A_{FF} \) is regular, then \( A \) can be factorized as

\[ A = \begin{bmatrix} I_F & 0 \\ A_{CF}A_{FF}^{-1} & I_C \end{bmatrix} \cdot \begin{bmatrix} A_{FF} & 0 \\ 0 & S \end{bmatrix} \cdot \begin{bmatrix} I_F & A_{FF}^{-1}A_{FC} \\ 0 & I_C \end{bmatrix}, \]

(6)

where

\[ S = A_{CC} - A_{CF}A_{FF}^{-1}A_{FC} \]

(7)

is the Schur complement. An approximation \( \tilde{A} \approx A \) is obtained by multiplying \( \tilde{A} \) with the approximations \( \tilde{A}_{FF} \approx A_{FF} \) and \( \tilde{S} \approx S \). A simple example for \( \tilde{S} \) is to replace \( A_{FF} \) by \( A_{FF}^{\alpha} \) in (7).

As outlined in [31], these block Schwarz factorization approaches are closely related to two-grid AMG methods. The AMLI method can be described as the stationary iteration with the iteration matrix

\[ T = I - M^{-1}A. \]

Definition 10. A splitting \((M, N)\) of \( A = M - N \) with \( M \) regular is called weak regular of the first type, if \( M^{-1} \geq 0 \) and \( M^{-1}N \geq 0 \).

The convergence of the AMLI method is ensured by the

Theorem 5. If the splittings \((\tilde{A}_{FF}, \tilde{A}_{FF} - A_{FF})\) and \((\tilde{S}, \tilde{S} - S)\) are weak regular of the first type, then \( T \geq 0 \), and \( \rho(T) < 1 \).

It is shown in [31] that if \( A \) is an M-matrix, then the assumptions in Thm. 5 are satisfied for approximations \( \tilde{A}_{FF} \) given by the Jacobi and the Gauß-Seidel methods and the incomplete LU factorization. However, the convergence proof fails in general, if \( A \) does not have an M-matrix structure. In [31], further methods have been presented (MAML, SMAML), that are based on other types of Schwarz factorizations. Under similar types of assumptions on the splittings, one has the convergence result \( \rho(T_{SMAML}) \leq \rho(T_{MAML}) \leq \rho(T_{AMLI}) \).

5. Least Squares vs. Linear Minimization Approach

Classical approaches for meshfree derivative approximation are moving least squares methods, based on scattered data interpolation [25], and local approximation methods, based on generalized finite difference methods [28]. Their application to meshfree settings has been analyzed in [15, 27].

Around a central point \( \vec{x}_0 \), points inside a radius \( r \) are considered. A distance weight function \( w(\delta) \) is defined, which is small for \( \delta > r \). For instance, one can consider \( w(\delta) = \delta^{-\alpha} \), where \( \alpha \geq 1 \).
Each point in the neighborhood $\vec{x}_i \in B(\vec{x}_0, r)$ is assigned a weight $w_i = w(\|\vec{x}_i - \vec{x}_0\|_2)$. A unique stencil is defined via a quadratic minimization problem

$$\min \sum_{i=1}^{n} \frac{s_i^2}{w_i}, \text{ s.t. } V \cdot \vec{s} = \vec{b}. \tag{8}$$

Using $W = \text{diag}(w_i)$, the solution of (8) is given by

$$\vec{s} = WV^{-1}(VWV^T)^{-1} \cdot \vec{b}. \tag{9}$$

Least squares approaches do not yield minimal stencils, unless exactly $k$ neighbors are considered. In general, they also do not yield positive stencils, as the following example shows.

**Example 1.** Consider $\vec{x}_0 = (0, 0)$ and 6 neighbors on the unit circle $\vec{x}_i = (\cos(\pi/2 \phi_i), \sin(\pi/2 \phi_i))$, where $(\phi_1, \ldots, \phi_6) = (0, 1, 2, 3, 0, 1, 0.2)$. Since all neighbors have the same distance from $\vec{x}_0$, the distance weight function does not play a role. Formula (9) yields the non-positive least squares stencil $\vec{s} = (0.846, 1.005, 0.998, 1.003, 0.312, -0.164)$. However, the configuration admits a positive stencil, namely $\vec{s} = (1, 1, 1, 1, 0, 0)$.

As outlined in Sects. 3 and 4.3, we can be sure that multigrid approaches converge, if an M-matrix structure is generated. Hence, as a new approach, we enforce the positivity of stencils, i.e. we search for solutions in the polyhedron

$$P = \{\vec{s} \in \mathbb{R}^m : V \cdot \vec{s} = \vec{b}, \vec{s} \geq 0\}. \tag{10}$$

This is the feasibility problem of linear optimization [51]. In Sect. 6 we provide conditions for the point geometry that guarantee the existence of solutions. If $P$ is non-void and not degenerate, there are infinitely many feasible stencils. To single out a unique stencil we now formulate a linear (or $\ell^1$) minimization problem

$$\min \sum_{i=1}^{m} s_i, \text{ s.t. } V \cdot \vec{s} = \vec{b}, \vec{s} \geq 0, \tag{11}$$

where the weights $w_i = w(\|\vec{x}_i - \vec{x}_0\|_2)$ are defined by an appropriately decaying (i.e. $w(\delta) = o(\delta^{-2})$, see [45]) non-negative distance weight function $w$. Problem (11) is a linear program (LP) in standard form. It is bounded, since we have imposed sign constraints and the weights $w_i$ are all non-negative.

**Theorem 6.** If the polyhedron (10) is non-void, then the linear minimization approach (11) yields minimal positive stencils.

**Proof.** The sign constraints in (11) ensure that the selected stencil is positive. The existence of a minimal solution is ensured by the fundamental theorem of linear programming [51]. If the LP (11) has a solution, then it also has a basic solution, in which at most $k$ of the $m$ stencil entries $s_i$ are different from zero. \[\Box\]

In contrast to least squares methods, the LP approach yields non-zero values only for a few selected points. This is in line with the work of Donoho [14, 13], who shows that a key advantage of $\ell^1$ minimization is sparsity.

**Remark 3.** For regular grids, the linear minimization approach selects standard finite difference stencils. For instance, for a regular Cartesian grid, the classical 5-point stencil (2d) or 7-point stencil (3d), respectively, is obtained. In these cases, the basic solution is degenerate, i.e. some of the basis variables are zero. Also for the configuration given in Example 1, the linear minimization approach selects the classical 5-point Laplace stencil.
6. Existence of Positive Stencils and Matrix Connectivity

The question of the existence of positive stencils, i.e. whether the polyhedron (10) is non-void, is nontrivial. In [45], the connection between the local point cloud geometry and the existence of a positive stencil has been presented. The key idea is to consider the dual linear program to (11) by using Farkas’ Lemma [51]. This yields a simple necessary criterion, as well as a simple sufficient criterion, as follows.

**Theorem 7 (Necessary Criterion for Positive Stencils).** Let the considered central point w.l.o.g. lie in the origin. If a set of points $X \subset \mathbb{R}^d$ around the origin admits a positive Laplace stencil, then they must not lie in one and the same half space (with respect to an arbitrary hyperplane through the origin).

**Theorem 8 (Sufficient Criterion for Positive Stencils).** Let the considered central point w.l.o.g. lie in the origin. Consider $d \in \{2, 3\}$, and let $\beta = \sqrt{2} - 1$ in 2d, and $\beta = \sqrt{\frac{1}{3} \left(3 - \sqrt{6}\right)}$ in 3d. Assume the set of points $X \subset \mathbb{R}^d$ has the property that in any (i.e. $\vec{v}$ with $\|\vec{v}\|_2 = 1$ arbitrary) cone $\vec{v} \cdot \vec{x} > \left(1 + \beta^2\right)^{-\frac{1}{2}} \|\vec{x}\|_2$, at least one neighboring point $\vec{x}_i \in X$ is contained. Then a positive Laplace stencil exists.

We call this criterion cone criterion. In 2d, the cone has a total opening angle 45°, and in 3d, the total opening angle is 33.7°. Using the cone criterion, the existence of positive stencils can be guaranteed by applying the $\ell^1$ minimization (11) to a large enough candidate set of neighboring points. As in [27], we define

**Definition 11.** Let $\Omega \subset \mathbb{R}^d$ be a domain and $X = \{\vec{x}_1, \ldots, \vec{x}_n\}$ a point cloud. The mesh size $h$ is defined as the minimal real number, such that $\bar{\Omega} \subset \bigcup_{i=1}^{n} \bar{B}(\vec{x}_i, \frac{h}{2})$, where $\bar{B}(\vec{x}, r)$ is the closed ball of radius $r$ centered in $\vec{x}$ and $\bar{\Omega}$ is the closure of $\Omega$.

**Theorem 9 (Condition on Point Cloud to Guarantee Positive Stencils).** Consider a point cloud of mesh size $h$, and let $\gamma$ be defined as in Thm. 8. If the radius of considered candidate points satisfies $r > \frac{1}{\sin(\gamma/2)} \frac{h}{2}$, then for every interior point which is sufficiently far from the boundary, a positive stencil exists.

The proof, and a description of special conditions near the boundary, are given in [45]. The ratio of candidate radius to maximum hole size radius $r/h$ equals 2.61 in 2d, and 3.45 in 3d. In practice, one generally observes that already much smaller smaller ratios yield positive stencils.

Due to Thm. 1 and Thm. 2, the matrix composed of positive stencils is an M-matrix, if every interior and Neumann boundary point is connected to a Dirichlet boundary point. As discussed in [45], the connectivity of every point to the boundary follows from Thm. 7, and the connectivity to a Dirichlet point can be ensured under comparably mild technical assumptions. Hence, if the implementation of the minimal positive stencil method ensures that Dirichlet boundary points are used in the stencils of nearby points, then the approach is guaranteed to generate M-matrices.

7. Computational Effort to Generate the System Matrix

In meshfree methods, the actual generation of the linear system that approximates the Poisson equation (1) carries a non-negligible computational cost. For each interior point, a minimization problem has to be solved.

Classical least squares methods are based on the quadratic minimization (8). The computation of the stencil by expression (9) requires first the set-up of the $k \times k$ matrix $VWV^T$, then the solution of the the linear system $(VWV^T) \cdot \vec{v} = \vec{b}$, and lastly the computation of the product $\vec{s} = VWV^T \cdot \vec{v}$. This requires $k(k + 1)m + \frac{k^3}{3}$ floating point operations [43, p. 150].
In contrast, the linear minimization approach \[11\] does not admit an explicit solution formula. Instead, a small linear program (LP) has to be solved. To our knowledge, there are no general results about efficient methods for such small LPs, especially any that would consider the special structure of the Vandermonde matrix. A numerical comparison of various methods has been presented in [43, p. 148]. Simplex methods \[8\] perform best for the arising LPs. A basis change corresponds to one stencil point replacing another. The theoretical worst case performance of simplex methods is not observed. Typical runs find the solution in about \(1.5k\) steps, resulting in a complexity of \(O(k^2m)\), which equals the asymptotic effort of least squares approaches.

8. Numerical Results

The two discretization approaches, least-squares vs. minimal positive stencils, have been compared in \[45\] in terms of accuracy of the finite difference approximation. For a sequence of point clouds, the global truncation errors with both approaches have been estimated numerically. The observation is that the approximation errors of the two approaches differ by at most 20\%. An interesting aspect is that the measured order of convergence is generally second order, although the constraints \[3\] enforce only first order accuracy. This effect is also observed in other types of meshfree approaches.

Since the two approaches do not differ significantly in the accuracy of approximation, a comparison of their computational effort when generating and solving the linear system is of interest. The presented linear minimization approach generates optimally sparse M-matrices, and we can hope that the solution of the arising linear systems is faster than for systems that are generated with classical least squares approaches.

We consider two types of multigrid/multilevel approaches. In Sect. 8.1 we investigate the performance of SAMG ("Algebraic Multigrid Methods for Systems") \[42\], by the Fraunhofer Institute for Algorithms and Scientific Computing. It is a library of AMG subroutines, based on the code RAMG05 \[19\] App. A], which is a successor of the public domain code AMG1R5 \[38\]. SAMG is implemented as a preconditioner for a BiCGstab \[50\] iteration. It performs a full multigrid cycle with Gauss-Seidel relaxation, and sparse Gaussian elimination on the coarsest level. The numerical tests presented here are performed using standard coarsening \[49\] pp. 473], though aggressive coarsening strategies \[49\] pp. 476] frequently turn out to improve the performance. In Sect. 8.2 we consider an AMLI type method, implemented by C. Mense and R. Nabben, TU Berlin, 2005.

8.1. Performance of SAMG. For a 3d Poisson test problem, the classical least squares approach is compared to the linear minimization approach in terms of the computational cost. We consider the geometry shown in Fig. 1. It is motivated by the simulation of the flow of a viscous material in a geometry with a thin outlet. More detail is given in \[43\] p. 183. The considered Poisson equation \[11\] has a smooth right hand side \(f\). Neumann boundary conditions are prescribed everywhere except for the thin outlet, at which Dirichlet boundary conditions are prescribed. For the considered geometry, the classical least squares approach is compared to the linear minimization approach in terms of the computational cost. For a sequence of point clouds with increasing resolution, the geometry is discretized. One approximation is generated using a weighted least squares approach \[13\], with about 40 neighbors for each interior point. Another approximation is generated from the \(\ell^1\) minimization approach \[11\]. By construction, each interior point has only 10 neighbors. Consequently, the \(\ell^1\) matrices possess about 4 times fewer non-zero entries than the LSQ matrices.

Fig. 2 shows the CPU times for the setup of the system matrices. The solid curve represents the costs with the least squares approach, which is based on expression \[10\]. The dashed curve shows the cost of the linear optimization approach, for which the simplex method is used. As predicted in Sect. 7 asymptotically the two approaches are equally expensive. In the considered example, the
linear optimization approach is about 20% more expensive. Note that the implemented simplex algorithm is not optimized. Better linear programming methods, and/or smart heuristics may accelerate the matrix setup phase.

Fig. 3 shows the CPU times for the solution of the arising system with a BiCGstab method [50], preconditioned by ILU(0) [40, Chap. 10]. In comparison, Fig. 5 shows the CPU times for the solution of the same system with SAMG [42]. Three aspects can be observed:

- The cost of the BiCGstab method grows superlinearly with the number of unknowns. In contrast, the SAMG cost is almost perfectly linear in the number of points.
• The SAMG method solves the arising linear systems significantly faster than the BiCGstab method. For low resolutions, SAMG is 2-3 times as fast as BiCGstab. For high resolutions, it is 4-5 times as fast. While the used BiCGstab method is certainly not fully optimized, the performance of SAMG is still compelling.

• For both solution methods, the linear systems that arise from the $\ell^1$ minimization approach are 3-4 times faster to solve than the systems generated by the LSQ approach.

Fig. 4 shows the memory consumption in an SAMG solve. As one can expect, it is perfectly linear in the number of unknowns, and proportional to the number of non-zero matrix entries.

The first conclusion we draw from this numerical experiment is that SAMG performs very well on both LSQ matrices and $\ell^1$ matrices. In particular, it is always observed to converge, even though the code is primarily designed for symmetric matrices, and even though most of the considered LSQ matrices are not M-matrices. SAMG is significantly faster than a classical BiCGstab method with ILU(0) preconditioning. In addition, the computational effort of SAMG is linear in the number of unknowns.

The second conclusion is that the matrices obtained from the $\ell^1$ minimization approach are solved significantly faster than classical LSQ matrices. The observed speed-up equals the increase in sparsity that the $\ell^1$ minimization approach yields. In another test (not shown here), we store the non-basis neighbors, as obtained by the simplex method, as actual zero-entries in the matrix. The solution times with those matrices turn out to be within 10% of the solution times with the LSQ matrices. This implies that the M-matrix property itself does not increase the actual convergence rate of linear solvers. On the other hand, very sparse Poisson stencils are often times observed to be less robust than larger ones. In this sense, one can interpret the M-matrix property as an ingredient that allows the selection of a minimal stencil, without actually worsening convergence properties of linear solvers.

8.2. Performance of an AMLI Type Method. We consider a 2d Poisson test problem on a circular domain $\Omega = \{ \vec{x} \in \mathbb{R}^2 : \| \vec{x} \|_2 < 1 \}$ as

\[
\begin{align*}
-\Delta u &= f \quad \text{in } \Omega \\
  u &= f \quad \text{on } \partial \Omega
\end{align*}
\]

where $f(x, y) = \sin(4x + 0.1) + x \cos(2y + 0.4)$. The analytical solution is $u(x, y) = 16 \sin(4x + 0.1) + 4x \cos(2y + 0.4)$. This problem is purposefully chosen to have a simple geometry, so that the focus lies on effects due to the unstructured point cloud. In [43, p. 173], this problem is used to analyze discretization errors on various point clouds. Here, we restrict to one specific point cloud, shown in Fig. 6. Boundary points are placed equidistantly on the unit circle. In the interior, 4000 points are placed. For the given point cloud, we generate four types of discretization matrices: the $\ell^1$ minimization matrix, and three LSQ matrices, with the closest 5 / 12 / 36 neighboring points considered in each stencil. For each matrix, the linear system is solved by various versions of AMLI type two-grid methods, all implemented by C. Mense and R. Nabben, TU Berlin, 2005.

Two versions of coarsening are considered. In default coarsening, the first half of the unknowns is chosen as coarse scale variables, and the other ones are selected as fine scale variables. Since here the points are not sorted, this is equivalent to a random selection of coarse variables. In the Ruge-Stüber coarsening [39], the coarsening of the matrix graph is performed preferably in the directions of large couplings, where the coupling strength (for an M-matrix) is defined as the magnitude of the corresponding off-diagonal entries.

Four versions of AMLI type methods are considered, which are based on different splittings [30, 31]. Specifically, besides the standard additive AMLI, a multiplicative (MAMLI), a reverse multiplicative (RMAMLI), and a symmetrized multiplicative (SMAMLI) splitting are considered.
Figure 6. Point cloud on a circular domain

Figure 7. Relative CPU times for various types of AMLI approaches

Fig. 7 shows the relative CPU times for the various test cases. The left third corresponds to the \( \ell^1 \) matrix, the middle third shows the LSQ approach with 12 neighbors for each point, and the right third represents the LSQ approach with 36 neighbors. The CPU times for the approach with the 5 nearest neighbors is not shown, since the AMLI methods failed to yield the correct solution. Needless to say, the obtained matrices are not M-matrices. However, also for the LSQ(12) and the LSQ(16) case, the Poisson matrices are in general not M-matrices. Nevertheless, the AMLI methods converge. In each case, the height of the dark part of the bar shows the setup time, and the light part shows the actual solution time.

One can observe that the \( \ell^1 \) matrices lead to the shortest solution times. However, unlike with the SAMG method, the speed-up factor with respect to the LSQ(12) case is not significant, although the \( \ell^1 \) matrix is only half as dense. With respect to the LSQ(36) matrices, a noticeable speed-up is observed, although again not as significant as one would expect from the sparsity factors. Another observation is that for the considered example, there is no significant difference between the various splitting versions. Finally, for larger stencils, the Ruge-Stuiben coarsening is more expensive than the default coarsening. While this could be expected for the setup phase, it is surprising for the iteration phase. We cannot provide a satisfying explanation for this observation.

9. Conclusions and Outlook

With the pressure correction step in particle methods for incompressible fluid flows, we have presented an application that leads to large, sparse, non-symmetric matrices. These matrices that approximate the Poisson equation are constructed using a meshfree finite difference method. They have a different structure from traditionally considered non-symmetric matrices, such as Markov chain representation matrices. In particular, classical meshfree least squares approaches yield in general matrices that do not have an M-matrix structure.

However, the M-matrix property is a key ingredient in recent convergence proofs of AMG and AMLI type methods for non-symmetric matrices. While it is not a necessary condition, convergence is not ensured if the M-matrix property is violated. In particle methods, Poisson equations are solved in every time step with ever changing geometries, and one single failure of convergence could spoil a whole long-term computation. Therefore, until multigrid convergence can be shown under weaker assumptions, the M-matrix structure is a property that should be guaranteed. The presented approach does so. Using an \( \ell^1 \) minimization formulation, and linear programming methods, optimally sparse M-matrices are generated. Conditions are provided that ensure the success of this approach.

\footnote{The absolute CPU times are not representative here and are thus omitted.}
The performance of multigrid methods for meshfree finite difference approaches is numerically investigated in various test cases. With SAMG, a sophisticated AMG method is applied, and with AMLI, an approach is used in which convergence for M-matrices is proved. The main conclusion is that the $\ell_1$ minimization approach guarantees convergence and generates matrices that are significantly sparser than classical LSQ matrices, and this sparsity results in a speed-up in multigrid solvers.

Currently, the setup of the system matrices with the $\ell_1$ minimization approach is slightly slower than with LSQ approaches. However, more efficient linear programming methods may turn the tide towards the linear optimization approach. The efficient solution of the linear programs is a crucial point in the presented method, and worth a deeper analysis. Another interesting question is whether a weaker requirement than the M-matrix property already implies convergence of AMG and AMLI methods. Numerical tests presented in [13] indicate that the convergence of AMLI methods tends to be related to the inverse positivity of an LSQ matrix, rather than an M-matrix structure. In fact, LSQ matrices are frequently observed to be inverse positive, while not having an L-matrix structure. Alternative characterizations of inverse positive matrices [17] may help answer this question.

Acknowledgments

The author would like to thank Dr. Jörg Kuhnert, Dr. Christian Mense, Dr. Klaus Stüben, and Dr. Sudarshan Tiwari for fruitful discussions and support on meshfree finite difference methods and on the application of algebraic multigrid codes. The support by the National Science Foundation is acknowledged. The author was partially supported by NSF grant DMS–0813648.

References

[1] O. Axelsson and P. Vassilevski. Algebraic multilevel preconditioning methods, Part I. *Numer. Math.*, 56:157–177, 1989.
[2] O. Axelsson and P. Vassilevski. Algebraic multilevel preconditioning methods, Part II. *SIAM J. Numer. Anal.*, 27:1569–1590, 1990.
[3] I. Babuska and J. M. Melenk. The partition of unity finite element method: Basic theory and applications. *Comput. Methods Appl. Mech. Engrg.*, 139:289–314, 1996.
[4] I. Babuska and J. M. Melenk. The partition of unity method. *Internat. J. Numer. Methods Engrg.*, 40:727–758, 1997.
[5] T. Belytschko, Y. Y. Lu, and L. Gu. Element-free Galerkin methods. *Internat. J. Numer. Methods Engrg.*, 37:229–256, 1994.
[6] A. Brandt. Algebraic multigrid theory: The symmetric case. *Appl. Math. Comput.*, 19:23–56, 1986.
[7] A. J. Chorin. Numerical solution of the Navier-Stokes equations. *Math. Comput.*, 22:745–762, 1968.
[8] V. Chvátal. *Linear programming*. W. H. Freeman and Company, 1983.
[9] S. J. Cummins and M. Rudman. Truly incompressible SPH. In *ASME Fluids Engineering Division Summer Meeting*, Washington DC, USA, 1998.
[10] S. J. Cummins and M. Rudman. An SPH projection method. *J. Comput. Phys.*, 152(2):584–607, 1999.
[11] L. Demkowicz, A. Karafiat, and T. Liszka. On some convergence results for FDM with irregular meshes. *Comput. Methods Appl. Mech. Engrg.*, 42:343–355, 1984.
[12] G. A. Dilts. Moving least squares particles hydrodynamics I, consistency and stability. *Internat. J. Numer. Methods Engrg.*, 44:1115–1155, 1999.
[13] D. L. Donoho. For most large underdetermined systems of linear equations, the minimal $\ell_1$-norm near-solution approximates the sparsest near-solution. *Comm. Pure Appl. Math.*, 59(7):907–934, 2006.
[14] D. L. Donoho. For most large underdetermined systems of linear equations, the minimal $\ell_1$-norm solution is also the sparsest solution. *Comm. Pure Appl. Math.*, 59(6):797–829, 2006.
[15] L. Duarte, T. Liszka, and W. Tworszyko. hp-meshless cloud method. *Comput. Methods Appl. Mech. Engrg.*, 130(1–7):263–288, 1996.
[16] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998.
[17] T. Fujimoto and R. R. Ranade. Characterization of inverse-positive matrices: The Hawkins-Simon condition and the Le Chatelier-Braun principle. *Electron. J. Linear Algebra*, 11:59–65, 2004.
Department of Mathematics, Temple University, 1801 N. Broad Street, Philadelphia, PA 19122

E-mail address: seibold@temple.edu

URL: http://www.math.temple.edu/~seibold