

1 **TOWARD A MULTILEVEL METHOD FOR THE HELMHOLTZ**
 2 **EQUATION ***

3 CLÉMENT RICHEFORT †

4 In collaboration with: Matthieu Lecouvez, Rob Falgout, Pierre Ramet

5 **Abstract.** It is well known that multigrid methods are very competitive in solving a wide
 6 range of SPD problems. However achieving such performance for non-SPD matrices remains an
 7 open problem. In particular, two main issues may arise when solving a Helmholtz problem. Some
 8 eigenvalues become negative or even complex, requiring the choice of an adapted smoothing method
 9 for capturing them. Moreover, since the near-kernel space is oscillatory, the geometric smoothness
 10 assumption cannot be used to build efficient interpolation rules. We present some investigations
 11 about designing a method that converges in a constant number of iterations with respect to the
 12 wavenumber. The method builds on an ideal reduction-based framework and related theory for SPD
 13 matrices to correct an initial least squares minimization coarse selection operator formed from a set
 14 of smoothed random vectors. We also present numerical results at the end of the paper.

15 **Key words.** Multigrid, Helmholtz, Linear Algebra, Polynomial Smoother

16 **1. Introduction.** The numerical simulation of various physical phenomena gives
 17 rise to potentially very large linear systems of equations written $Au = b$ in matrix
 18 form. These systems can be solved directly by a convenient factorization of A , or
 19 iteratively by computing and refining an approximation of the solution u starting
 20 from an initial guess u_0 . Multigrid methods [15, 8] work iteratively and are known
 21 to be scalable and quasi-optimal for solving sparse linear systems of equations for
 22 many classes of problems. To simplify the discussion in what follows, we use the term
 23 "small/large eigenvector" to mean an eigenvector with small/large eigenvalue. We
 24 similarly say "positive/negative eigenvector" when referring to the eigenvalue sign.

25 **1.1. General aspects on multigrid methods.** The core idea in multigrid
 26 methods is to accelerate the computation of u by way of a hierarchy of coarse problems
 27 $A_l u_l = b_l$, l being the level in the grid hierarchy. A restriction operator R_l transfers
 28 the information from a level l to a coarser one $l + 1$, while an interpolation operator
 29 P_l transfers the information from level $l + 1$ to l . In most symmetric applications,
 30 $R_l = P_l^T$ and coarse matrices are constructed following the Galerkin formula $A_{l+1} =$
 31 $P_l^T A_l P_l$. Two-grid methods actually need both types of solvers: a direct method for
 32 the coarse correction, and an iterative method called a smoother on the fine level.
 33 The error propagation matrix for the coarse correction is

34 (1.1)
$$E = I - P(P^T A P)^{-1} P^T A.$$

35 The error propagation matrix for the smoother is

36 (1.2)
$$E_M = I - M^{-1} A$$

37 where M^{-1} is an approximation of A^{-1} , for instance the diagonal inverse or a poly-
 38 nomial of A . The smoother is applied before each restriction and after each inter-
 39 polation. Finding a smoother and a coarse correction that are complementary is a

*This work was funded by CEA.

†Commissariat à l’Energie Atomique et aux Energies Alternatives (CEA),
 (richefort.clement@protonmail.com).

major concern in the design of the method. The interpolator must propagate the coarsest information back to the finest, and transferred errors should be eliminated by the smoother. The smoother targets the large eigenvectors and the coarse correction targets the small eigenvectors. The near-kernel space of smallest eigenvectors is especially important in the design of interpolation. A multi-level method can be created by recursively applying the two-level method to solve the coarse system, where a direct solver is used on the coarsest level. The context in which a multigrid method is applied determines what kind of operators should be used in the method. In elliptic problems, where the convergence of multigrid methods is well known, the matrix A is symmetric positive definite, so smoothers like weighted-Jacobi or Gauss-Seidel are known to be good smoothers since they damp the high frequencies without modifying the low frequency eigenvectors. Likewise, interpolators are designed to target slowly varying components as in classic algebraic multigrid methods (AMG) [16].

1.2. Why Helmholtz problems are difficult for multigrid. The Helmholtz problem involves indefinite matrices with potentially wide and oscillatory near-kernel space [7]. This complication breaks the near-kernel space geometric smoothness assumption, a keystone of many multigrid methods. To satisfy the complementarity principle in this context, interpolation rules must reproduce the near-kernel oscillation, and smoothers have to deal with both positive and negative eigenvalues. More importantly, finding a recurring process to build a scalable multilevel method is still an open question. The Helmholtz equation (1.3) is our target in this paper.

$$(1.3) \quad (\text{Continuous Helmholtz problem}) \Leftrightarrow \begin{cases} -\Delta u - k^2 u = f & \text{on } \Omega \\ + \text{ b. c.} & \text{on } \partial\Omega \end{cases}$$

Since the Helmholtz equation can be seen as a shifted Poisson equation, the geometrically smooth components (ie. low Fourier modes) become negative. Because of the shift, the smallest eigenvectors are higher in frequency. Multigrid interpolators must now focus on this more oscillatory spectrum interval. Multiple correction [12] and wave-ray [4, 11] approaches have already been investigated to address this issue. In this paper, we present an approach built on ideal reduction-based ideas, and demonstrate its potential for solving the Helmholtz problem in constant iteration count independent of the wavenumber k . In Section 2, we present a normal equation polynomial smoother specifically designed to damp the desired proportion of highest amplitude eigenvalues, while interpolation rules for propagating oscillatory near-kernel information are established in Section 3. Finally, Section 4 contains benchmarks of this new multigrid cycle for different Helmholtz problems, with a varying wavenumber k .

2. Polynomial Smoothers for Indefinite Problem. Working with a smoothing method whose behavior on each portion of the spectrum is *a priori* known is interesting to guarantee the effectiveness of the cycle. Here, the smoother must drop large positive and negative eigenvalues, which is problematic for most standard methods. Generally, a polynomial method with degree greater than one can work. Krylov iterations are good polynomial smoother in the indefinite case but they minimize the global residual norm regardless of the eigenvalues and are non-linear because of their right-hand side dependence. A linear polynomial is more convenient for generating the set of smoothed candidates vectors needed to construct the interpolation operator described in Section 3.

2.1. General considerations on polynomial smoothers. One way to ensure that both positive and negative eigenvectors are damped is to consider a normal

87 equation polynomial smoother. In general, the degree d of the polynomial must
 88 be greater than 1 to damp positive and negative eigenvectors, as the polynomial
 89 illustrated in Figure 2.1 does. This condition is guaranteed here since the normal
 90 equations lead to an even polynomial degree. In the future, it might be interesting
 91 to investigate more general polynomials to avoid normal equations and consider odd
 92 degrees. In this first approach, we use the convenient symmetric property enabled by
 93 normal equations in the Chebyshev framework. For any positive x and given some
 94 positive spectrum interval $\mathcal{I} := [x_{\min}, x_{\max}]$, we look for a polynomial $p(x)$ of degree d
 95 maximizing the damping of all components lying in \mathcal{I} . In that direction, and according
 96 to Equation (1.2), let

$$97 \quad (2.1) \quad q(A^H A) := I - p(A^H A)A^H A$$

98 be the associated error propagation operator. Then, for any right singular vector v of
 99 A and σ the associated singular value,

$$100 \quad (2.2) \quad q(A^H A)v = q(\sigma^2)v = (1 - p(\sigma^2)\sigma^2)v$$

101 According to Equation (2.2), the lower the amplitude of a component, the less the
 102 polynomial smoother will damp it. This intact portion of low components enables
 103 the construction of a set of smoothed vectors approximating the near-kernel space, as
 104 detailed in Section 3. Let L be a set of points lying in this interval such that

$$105 \quad (2.3) \quad \forall x_i \in L = (x_1, \dots, x_d), x_i \neq 0, q(x_i) = 0 \Leftrightarrow p(x_i) = \frac{1}{x_i}.$$

106 Then the polynomial smoother p is constructed following the Lagrangian formula

$$107 \quad (2.4) \quad x_i \neq x_j, p(x) = \sum_{j=0}^d \frac{1}{x_j} \prod_{i=0, i \neq j}^d \frac{x - x_i}{x_j - x_i}.$$

108 However, such interpolation points should not be selected randomly within the inter-
 109 val, but in order to minimize the polynomial amplitude and avoid Runge's phenom-
 110 enon. Those best interpolating points are defined by the scaled first kind Chebyshev
 111 polynomial roots

$$112 \quad (2.5) \quad x_i := \frac{x_{\max} + x_{\min}}{2} + \frac{x_{\max} - x_{\min}}{2} \cos\left(\frac{(2i-1)\pi}{2d}\right).$$

113 Large intervals \mathcal{I} require a higher polynomial degree to flatten its oscillations.

114 **2.2. Constructing an appropriate target interval \mathcal{I} .** One way to determine
 115 a good interval without preliminary information [2, 1] is to compute a few power iter-
 116 ations to determine x_{\max} , overestimate the result by 10%, and choose the lower bound
 117 x_{\min} according to x_{\max} , for example $x_{\min} = 0.5x_{\max}$. To respect the complementarity
 118 principle, the percentage of damped eigenvalues by the smoother must approximate
 119 the proportion of non-coarse variables. For instance, if a coarse level space is defined
 120 by a selection of a quarter of the finer level's degrees of freedom, then three-quarters
 121 of the largest amplitude components should be damped by smoothing steps, while
 122 the coarse correction deals with the other part. Consequently, since eigenvalues are
 123 not necessarily uniformly separated, x_{\min} should be determined so that such a pro-
 124 portion of eigenvalues belongs to the interval \mathcal{I} . In this paper and following [10], we

125 first compute a rough approximation of the matrix *spectral density* defined by the
 126 distribution function $\phi(t)$, which represents the probability of finding an eigenvalue
 127 at each point t of a given interval. We set the lower bound x_{\min} of the Chebyshev
 128 nodes interval in a second step so that the probability within the interval is equal to
 129 the target proportion, for instance half of the total area in a scenario of exact balance
 130 between \mathcal{C} and \mathcal{F} points. The squared singular values should be scaled by setting
 131 $B := \frac{2}{x_{\max}} A^H A - I$, so the distribution function ϕ can be approximated by a linear
 132 combination of orthogonal Chebyshev polynomial functions

$$133 \quad (2.6) \quad \phi(t) = \sum_{k=0}^{\infty} \mu_k T_k(t) \approx \sum_{k=0}^M \mu_k T_k(t), \quad \mu_k = \frac{2 - \delta_{k0}}{n\pi} \times \text{Trace}(T_k(B))$$

134 where $T_k(t) = \cos(k \arccos(t))$ and the coefficients μ_k are determined by a moments
 135 matching procedure. Here, n corresponds to the matrix size and δ_{k0} the Kronecker
 136 symbol. The trace can be estimated using a set U of n_{vec} random and orthogonal
 137 vectors u , where each element of these vectors is chosen following a normal distribution
 138 with zero mean and a unit standard deviation. Any vector u can be written as a
 139 linear combination of B eigenvectors v , giving the expression $u = \sum_{j=1}^n \beta_j v_j$. As a
 140 consequence, $\mathbb{E}[u] = 0$ and $\mathbb{E}[\beta_i \beta_j] = \delta_{ij}$, giving the following convenient property

$$141 \quad (2.7) \quad \mathbb{E}[u^T T_k(B) u] = \mathbb{E}\left[\sum_{j=1}^n \beta_j^2 T_k(\lambda_j(B))\right] = \sum_{j=1}^n T_k(\lambda_j(B)) = \text{Trace}(T_k(B)).$$

142 According to (2.7), each trace can be estimated by a sample mean of n_{vec} products
 143 $u^T T_k(B) u$, and the M vectors $T_k(B) u$ can be computed from the three-term recur-
 144 rence relation $T_{k+1}(B) u = 2B T_k(B) u - T_{k-1}(B) u$, with initial terms $T_0(B) u = u$,
 145 $T_1(B) u = B u$. Once the distribution function ϕ is approximated following the ex-
 146 pression given in (2.6), a rough area approximation by trapezoid rule yields a correct
 147 lower bound x satisfying a proportion around $\text{card}(\mathcal{F})/n$. Computing x_{\min} by remap-
 148 ping x to the initial squared singular values scaling leads to a satisfying and purely
 149 algebraic interval in which our polynomial smoother will be the most efficient. The
 150 bounds x_{\min} and x_{\max} are represented in Figure 2.1, where $x_{50\%}$ represents the the-
 151 oretical lower bound targeted.

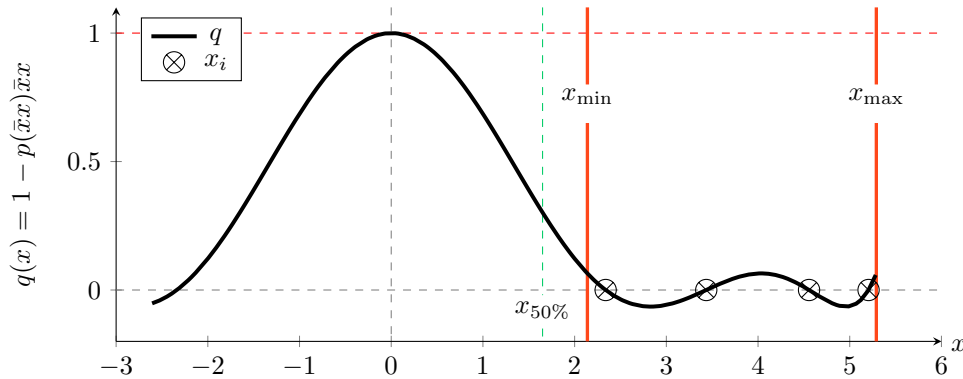


FIG. 2.1. *Spectrum of the polynomial smoother error propagation matrix*

152 **3. Constructing good interpolation rules.** Interpolators are used both to
 153 construct the coarse level matrices and to transfer information across levels. SPD
 154 and geometric smoothness assumptions cannot be used to determine appropriate in-
 155 terpolators in our case. Some methods such as smoothed aggregation [6, 13] and
 156 bootstrap-AMG [3] use candidate vectors that are close to the near-kernel space to
 157 design the interpolation rules. These test vectors are either deduced from geometric
 158 information [4] or algebraically as in adaptive multigrid methods [5]. Here, we prefer
 159 to stick to a fully algebraic and recurring process to create our interpolators. Can-
 160 didate vectors will be generated from random vectors smoothed by the polynomial
 161 presented in Section 2, and used by the least squares minimization framework to de-
 162 termine good fine variable interpolation rules. This initial least squares interpolator
 163 is used as a coarse selection operator in the ideal reduction-based framework [9].

164 **3.1. Ideal framework.** Even though the ideal framework in [9] requires an SPD
 165 assumption and has not been generalized to indefinite problems, removing orthogonal
 166 information from the interpolation range still improves its accuracy. Furthermore,
 167 assuming the smoother captures this orthogonal information, we can still expect good
 168 convergence as exposed by Equation (3.3) below. Following [9], let \mathcal{C} and \mathcal{F} be
 169 complementary coarse and fine variables subsets of Ω . Let $R^T : \mathbb{R}^{n_{\mathcal{C}}} \rightarrow \mathbb{R}^n$ and
 170 $S : \mathbb{R}^{n_{\mathcal{F}}} \rightarrow \mathbb{R}^n$ be coarse and fine selection operators respectively, such that $RS = 0$,
 171 for instance, the orthogonal matrices

$$172 \quad (3.1) \quad R^T = [0 \ I_{\mathcal{C}}]^T, \quad S = [I_{\mathcal{F}} \ 0]^T.$$

173 The space defined by the coarse selection operator R^T must be handled by the coarse
 174 correction, whereas the fine variables selection operator S defines a space where
 175 smoothing must operate in order to respect the complementarity principle. The *Ideal*
 176 *Interpolator* is a theoretical operator that is the best interpolator satisfying $RP = I_{\mathcal{C}}$
 177 in the sense that it minimizes the difference between variables and interpolated coarse
 178 variables, within a space that is the most complementary to the range of the smoother.
 179 For this reason, P_* is called *ideal* and is defined by

$$180 \quad (3.2) \quad P_* = (I - S(S^T AS)^{-1} S^T A) R^T.$$

181 The left operator in (3.2) removes all complementary \mathcal{F} -related information from
 182 R^T . Such information is irrelevant at a coarse level and should be handled by the
 183 smoother. Under the assumption that the smoother captures \mathcal{F} -related information,
 184 the best coarse matrix, according to the complementarity principle, is a matrix where
 185 fine variable information is removed. Furthermore, a simple development of a two-grid
 186 cycle combining the ideal coarse error propagation operator E_* , as initially defined in
 187 (1.1), with an \mathcal{F} -relaxation matrix $E_{\mathcal{F}}$ as a pre-smoother shows that

$$188 \quad (3.3) \quad E_* E_{\mathcal{F}} = (I - P_*(P_*^T A P_*^T)^{-1} P_*^T A)(I - S(S^T AS)^{-1} S^T A) = 0.$$

189 In (3.3), the final error propagation matrix is null, meaning that in this idealistic
 190 scenario, one iteration of \mathcal{F} -relaxation and one coarse correction is equivalent to a
 191 direct method. As shown in [9], we can extend this principle by separating coarse
 192 and fine spaces respectively in directions of low and high frequencies. Let $V_0 =$
 193 $(v_1, \dots, v_{\text{Card}(\mathcal{C})})$ and $V_+ = (v_{\text{Card}(\mathcal{C})+1}, \dots, v_n)$ respectively be low frequency and
 194 high frequency eigenvectors sets, assuming here that $\lambda_i \leq \lambda_{i+1}$. Then we define
 195 $R^T = V_0$ and its counterpart $S = V_+$. Naturally, since eigenvectors are orthonormal,
 196 the necessary condition $RS = 0$ is still satisfied. It finally gives

$$197 \quad (3.4) \quad P_* = R^T \text{ and } A_{\mathcal{C}} = \text{Diag}(\lambda_0, \dots, \lambda_{\text{Card}(\mathcal{C})}).$$

198 This example proposes another idealistic dichotomy enabled by P_* , maximizing the
 199 complementarity principle. The near-kernel space is solved directly at the coarsest
 200 level while the high frequencies remain in the smoothing space. Even if $(S^T AS)^{-1}$
 201 is most of the time impossible to use in practice, it gives insight on an idealistic
 202 convergence scenario.

203 **3.2. Least Squares Minimization Interpolator.** As mentioned at the begin-
 204 ning of Section 3.1, demonstrating that interpolator (3.2) is ideal in the theoretical
 205 framework of [9] requires A to be symmetric positive-definite. However, the reduction
 206 viewpoint of Equation (3.3) is still valid, hence removing the orthogonal fine informa-
 207 tion from the coarse selection operator is a viable approach. Numerical experiments
 208 show that the coarse selection operator in (3.1) is not a good option for Helmholtz.
 209 Using the lowest components V_0 from Section 3.1 to guarantee the representation of
 210 the near-kernel space within the interpolation range is not practical. Instead, we con-
 211 struct a set of vectors approximating an oscillatory and potentially large near-kernel
 212 space by using the normal equations polynomial smoother developed in Section 2.

213 In this section, we present a coarse selection operator R^H constructed by a least
 214 squares minimization strategy [3]. Let K be a set of κ vectors that approximate the
 215 near-kernel space, and assume some \mathcal{C}/\mathcal{F} splitting with $n_{\mathcal{C}}$ and $n_{\mathcal{F}}$ their respective
 216 size. Coarse variables are interpolated to the finer level with a simple injection rule,
 217 meaning the coarse interpolation block in \hat{R}^H corresponds to a $n_{\mathcal{C}} \times n_{\mathcal{C}}$ identity
 218 matrix, while fine interpolation rules are determined by the least squares minimization
 219 method presented in this section. Let i be a fine variable and r_i the i^{th} row of
 220 \hat{R}^H . The idea consists of constructing each fine interpolation rule by minimizing the
 221 squared difference between fine values of the near-kernel candidate vectors and the
 222 interpolation from their connected coarse variables \mathcal{C}_i . Denote by $K_{:,l}$ the l^{th} test
 223 vector, $K_{i,:}$ a row vector containing the i^{th} values of each test vector, and $K_{\mathcal{C}_i,l}$ a
 224 vector containing the values in $K_{:,l}$ of the coarse variables that are connected to the
 225 i^{th} fine variable. Then

$$226 \quad (3.5) \quad \forall i \in \mathcal{F}, r_i = \arg \min_r \sum_{l=1}^{\kappa} w_l (K_{i,l} - r \cdot K_{\mathcal{C}_i,l})^2 := \arg \min_r \mathcal{L}_i(r)$$

227 where w_l are scaling weights (for instance $w_l = 1/\lambda_l$ if K contains near-kernel eigen-
 228 vectors). Finding the minimum of the convex loss function \mathcal{L}_i is equivalent to solving

$$229 \quad (3.6) \quad \nabla \mathcal{L}_i(r_i) = 0.$$

230 Equation (3.6) can be rewritten element-wise

$$231 \quad (3.7) \quad \frac{\partial \mathcal{L}_i(r_i)}{\partial r_{ij}} = \sum_{l=1}^{\kappa} 2w_l (K_{i,l} - r_i \cdot K_{\mathcal{C}_i,l}) K_{\mathcal{C}_{ij},l} = 0, \forall j \in [1, \text{card}(\mathcal{C}_i)].$$

232 Finally, (3.7) leads to a system of linear equations to solve for each fine variable i

$$233 \quad (3.8) \quad r_i K_{\mathcal{C}_i} W K_{\mathcal{C}_i}^H = K_i W K_{\mathcal{C}_i}^H$$

234 The matrix is full rank and the solution of Equation (3.8) is unique if we have at
 235 least $\kappa = \max_i \{\text{Card}(\mathcal{C}_i)\}$ locally linearly independent test vectors. Even if it
 236 is statistically always the case when starting from random candidate vectors, the
 237 matrix singularity can be detected during the factorization. In that special case, a
 238 pseudo-inverse can be computed to find an optimal solution in the least squares sense.

239 **3.3. Ideal approximation from least squares coarse selection.** In Section
 240 3.2, we presented a better coarse selection operator for Helmholtz designed by a least
 241 squares minimization strategy. Using the framework presented in 3.1, let us define
 242 new coarse and fine selection rules

$$243 \quad (3.9) \quad \hat{R}^H = [R_{\mathcal{F}} \ I_C]^T, \quad \hat{S} = [I_{\mathcal{F}} \ -R_{\mathcal{F}}^H]^T,$$

244 where \hat{R}^H is the least squares operator presented in Section 3.2 and $R_{\mathcal{F}}$ is its fine
 245 variable interpolation block. Note that $\hat{R}\hat{S} = 0$ as required. To simplify the discussion,
 246 let $A_{\mathcal{F}} := \hat{S}^H A \hat{S}$. Beyond the necessity to find appropriate coarse and fine selection
 247 operators, another important concern is related to the inverse of $A_{\mathcal{F}}$ required in (3.2).
 248 Reorganizing the definition of the ideal interpolator,

$$249 \quad (3.10) \quad \hat{P} := (I - \hat{S}A_{\mathcal{F}}^{-1}\hat{S}^H A)\hat{R}^H \approx \hat{R}^H - \hat{S}X_{\mathcal{K}}^{-1}\hat{S}^H A\hat{R}^H,$$

250 where $X_{\mathcal{K}}^{-1}$ is the best polynomial approximating $A_{\mathcal{F}}^{-1}$ within the Krylov subspace
 251 \mathcal{K} . Since we need to ensure our interpolator keeps good sparsity, the subspace must
 252 be constrained according to a given pattern \mathcal{P} as detailed in [14]. In our case, the
 253 matrix inversion approximation will be computed column-wise, giving more flexibility
 254 than by computing the global constrained matrix approximation at once with a single
 255 polynomial. Consequently, let \mathcal{P}_i be some vector sparsity pattern, and define the
 256 associated operator $Z_i : \mathbb{C}^n \rightarrow \mathbb{C}^{\text{Card}(\mathcal{P}_i)}$ filled with ones and zeros that restricts any
 257 full vector to the non-zero pattern \mathcal{P}_i . Also, define

$$258 \quad (3.11) \quad b_i := \hat{S}^H A \hat{R}_{:,i}^H.$$

259 In practice, we choose $\mathcal{P}_i := \mathcal{P}(b_i)$ such that $Z_i^T Z_i b_i = b_i$. Then, we construct for
 260 each right hand side b_i the corresponding constrained Krylov subspace

$$261 \quad (3.12) \quad \mathcal{K}_{\mathcal{P}_i}^m = \{ Z_i b_i, Z_i A_{\mathcal{F}} Z_i^T Z_i b_i, \dots, (Z_i A_{\mathcal{F}} Z_i^T)^{m-1} Z_i b_i \},$$

262 within which we approximate the multiplication of $A_{\mathcal{F}}^{-1}$ with b_i . Since each subspace
 263 is constructed under a sparsity constraint, the solution is approximated locally with
 264 a window of $A_{\mathcal{F}}$. Consequently, it is not guaranteed that the accuracy will increase
 265 with respect to m , however a few iterations are enough to reach a good approxima-
 266 tion in practice. It is still possible to converge toward the best solution in a least
 267 squares sense, but this requires reformulating the problem with normal equations
 268 which increases the cost of construction.

269 **4. Numerical Experiments.** In this section, we present some numerical results
 270 performed with this new multigrid cycle. One pre-smoothing and one post-smoothing
 271 iteration of the normal equations Chebyshev polynomial presented in Section 2 are
 272 computed before each restriction and after each interpolation respectively. The inter-
 273 polator \hat{P} is constructed following Section 3. In practice, the degree d of the normal
 274 equation polynomial smoother p of Equation 2.2 is equal to 3, the x_{\max} is computed
 275 by several power iterations (5 to 10 iterations are enough in practice) and x_{\min} by the
 276 spectral density approximation described in Section 2.2, with the parameters $M = 5$
 277 and $n_{\text{vec}} = 15$. The number κ_l of input random vectors per level l to construct a
 278 correct approximation of near-kernel space follows the arbitrary recursive equation

$$279 \quad (4.1) \quad \kappa_l = 5 \times \max_{i \in \mathcal{F}} \{ \text{Card}(\mathcal{C}_i) \} + \kappa_{l-1} + 10l, \quad \kappa_{-1} = 0$$

280 where \mathcal{C}_i is the set of coarse variables strongly connected to the fine variable i . The
 281 strongly connected variables are selected according to a θ -rule [16] comparing matrix
 282 entries in absolute value. Before being smoothed by the normal equation polynomial,
 283 test vectors are created either by the restriction of finer level test vectors or randomly.
 284 In practice, $\text{Card}(\mathcal{C}_i)$ never exceeds 10. The size m of the Krylov subspaces needed
 285 to approximate $A_{\mathcal{F}}^{-1}$ in the ideal framework is set to 3.

286 **4.1. Benchmarks.** We apply this multigrid method on a 5-points stencil Carte-
 287 sian discretization of the Helmholtz equation (1.3) with absorbing boundary condi-
 288 tions ($\partial_n u - iku = 0$ on $\partial\Omega$), where 10 points per wavelength are used ($h = \lambda/10 \Leftrightarrow$
 289 $kh = 2\pi/10 \approx 0.625$). Multigrid cycles are iterated until the residual norm falls be-
 290 low 10^{-6} . This method will be benchmarked on more difficult geometries in the
 291 future, however this simple discretization already allows us to tackle the oscillating
 near-kernel space problem.

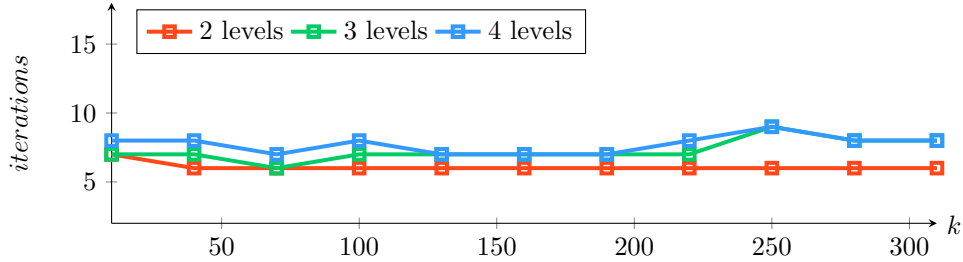


FIG. 4.1. Number of iterations following the wavenumber k

292 The three V-cycle schemes benchmarked in Figure 4.1 converge in a roughly con-
 293 stant number of iterations independent of the wavenumber k . We also present these
 294 results together with the overall complexity of our multigrid method. Let $\overline{nnz}(\cdot)$ be
 295 the average number of non-zeros per row of a given matrix. We measure the sparsity
 296 of each interpolator \hat{P}_l and level matrix A_l in Table 4.1. As expected, matrices on
 297 deeper levels are denser. Even if $A_{\mathcal{F}}^{-1}$ is approximated under pattern constraints and
 298 already allows to find an interesting trade-off between sparsity and good interpolation
 299 properties, it will be necessary to improve the sparsity of \hat{P} to reach a more competi-
 300 tive multigrid method. A thresholding strategy or other heuristics on its pattern will
 301 be one of the main concerns for the future developments of the method.
 302

k	70	100	130	160	190	220	250	280	310
$n (\times 10^5)$	1.28	2.59	4.37	6.60	9.30	12.46	16.08	20.16	24.7
$n_c (\times 10^5)$	0.16	0.32	0.55	0.83	1.16	1.56	2.01	2.52	3.09
$\overline{nnz}(A_0)$	5	5	5	5	5	5	5	5	5
$\overline{nnz}(A_1)$	77	78	79	79	79	80	80	80	80
$\overline{nnz}(A_2)$	299	315	316	322	332	335	337	339	341
$\overline{nnz}(A_3)$	322	342	345	359	376	380	385	386	386
$\overline{nnz}(\hat{P}_1)$	20	20	20	20	20	20	20	20	20
$\overline{nnz}(\hat{P}_2)$	63	64	65	66	66	66	67	67	67
$\overline{nnz}(\hat{P}_3)$	168	178	178	183	188	190	191	192	193

TABLE 4.1

Sparsity measurement of level matrices and interpolators following k

303 **4.2. Large shift experiment.** Even if the approach presented in this paper can
 304 be improved in many ways, it provides a direction for constructing interpolation for
 305 problems with oscillatory near-kernel spaces like the Helmholtz equation. However,
 306 since the problem is indefinite, the matrix A does not provide a norm. As a conse-
 307 quence, there is no guidance on the convergence. Moreover, even under the convenient
 308 assumption where $\hat{P}\hat{P}^T v \approx v$ where v is an eigenvector associated with a very small
 309 eigenvalue, for some large shift problems, the method can be divergent. To illustrate
 310 this issue, let L_s be a scaled laplacian matrix shifted by a large coefficient α (yielding
 $\lambda(L_s) \in [-\alpha, 8 - \alpha]$), such that its near-kernel space is very oscillatory.

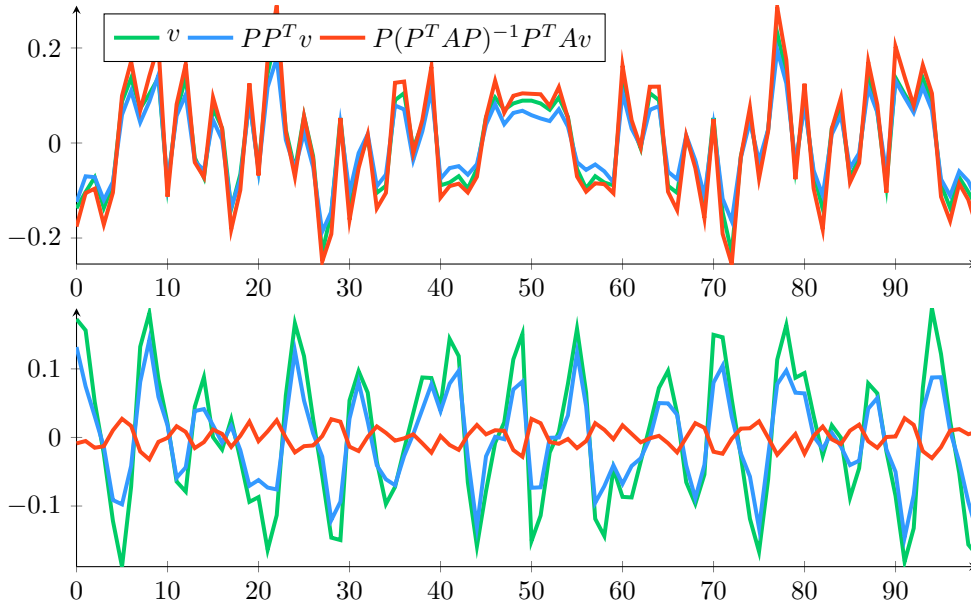


FIG. 4.2. Layering of lowest eigenvector (green), interpolation of restricted lowest eigenvector (blue) and coarse correction applied to lowest eigenvector (red) for $\alpha = 2.68$ (top) vs. $\alpha = 2.98$ (bottom)

311

312

313

314

315

316

317

318

319

320

321

322

323

Figure 4.2 shows the layering of the lowest eigenvector of L_s (v - green), the interpolation of the restricted lowest eigenvector ($\hat{P}\hat{P}^T v$ - blue), and the vector returned by the coarse correction ($\hat{P}(\hat{P}^T A \hat{P})^{-1} \hat{P}^T Av$ - red). As we can see, v and $\hat{P}\hat{P}^T v$ are very close in both large shift experiments, the very oscillatory near-kernel vector v is well approximated by the interpolation range. However, in the second experiment, this is not true for the coarse correction where the red vector seems to be oriented oppositely to v . This experiment shows that capturing the oscillatory near-kernel space with an appropriate set of interpolators, such as \hat{P} , will not necessarily be enough to reach a perfectly recurring method for solving indefinite problems like Helmholtz. For this reason, adding more levels has been challenging. One of our priorities is about finding a framework guaranteeing the convergence of the method in an indefinite context such as this large shift experiment.

324 **5. Conclusions.** Indefinite and oscillatory problems are difficult for multigrid
 325 methods. The negative eigenvalues require an adapted smoother, and the interpolator
 326 should be able to propagate the oscillatory near-kernel space. We presented a method
 327 that reaches those requirements up to a certain limit. The normal equation poly-
 328 nomial smoother is designed to target a desired proportion of components according
 329 to their amplitude, and the range of our interpolator offers a good approximation
 330 of the near-kernel space despite its oscillations. Finding more accurate interpolation
 331 rules, improving the sparsity of our operators, and constructing a polynomial without
 332 resorting to normal equations will be important points in our future investigations.
 333 However, the ultimate objective is to find a proper framework for indefinite problems
 334 guaranteeing the convergence of our multigrid method.

335

REFERENCES

- 336 [1] M. F. ADAMS, M. BREZINA, J. J. HU, AND R. S. TUMINARO, *Parallel multigrid smoothing:*
 337 *polynomial versus gauss–seidel*, Journal of Computational Physics, 188 (2003), pp. 593–
 338 610.
- 339 [2] A. H. BAKER, R. D. FALGOUT, T. V. KOLEV, AND U. M. YANG, *Multigrid smoothers for ultra-*
 340 *parallel computing*, SIAM Journal on Scientific Computing, 33 (2011), pp. 2864–2887, <https://doi.org/10.1137/100798806>, <https://doi.org/10.1137/100798806>, <https://arxiv.org/abs/https://doi.org/10.1137/100798806>.
- 343 [3] A. BRANDT, J. BRANNICK, K. KAHL, AND I. LIVSHITS, *Bootstrap amg*, SIAM Journal of Scien-
 344 tific Computing, 33 (2011), pp. 612–632, <https://doi.org/10.1137/090752973>.
- 345 [4] L. I. BRANDT A., *Wave-ray multigrid method for standing wave equations.*, ETNA. Electronic
 346 Transactions on Numerical Analysis [electronic only], 6 (1997), pp. 162–181, <http://eudml.org/doc/119506>.
- 348 [5] M. BREZINA, R. FALGOUT, S. MACLACHLAN, T. MANTEUFFEL, S. MCCORMICK, AND J. RUGE,
 349 *Adaptive smoothed aggregation (asa)*, SIAM Journal on Scientific Computing, 25 (2004),
 350 pp. 1896–1920, <https://doi.org/10.1137/S1064827502418598>, <https://doi.org/10.1137/S1064827502418598>, <https://arxiv.org/abs/https://doi.org/10.1137/S1064827502418598>.
- 352 [6] P. EK, M. BREZINA, AND J. MANDEL, *Convergence of algebraic multigrid based on smoothed*
 353 *aggregation*, Computing, 56 (1998), <https://doi.org/10.1007/s002110000226>.
- 354 [7] O. G. ERNST AND M. J. GANDER, *Why it is difficult to solve helmholtz problems with classical*
 355 *iterative methods*, (2010).
- 356 [8] R. D. FALGOUT, *An introduction to algebraic multigrid*, Computing in Science and Engineering,
 357 vol. 8, no. 6, November 1, 2006, pp. 24–33, (2006), <https://www.osti.gov/biblio/897960>.
- 358 [9] R. D. FALGOUT AND P. S. VASSILEVSKI, *On generalizing the amg framework*, SIAM J. NUMER.
 359 ANAL, 42 (2003), pp. 1669–1693.
- 360 [10] L. LIN, Y. SAAD, AND C. YANG, *Approximating spectral densities of large matrices*, SIAM Re-
 361 view, 58 (2016), pp. 34–65, <https://doi.org/10.1137/130934283>, <https://doi.org/10.1137/130934283>, <https://arxiv.org/abs/https://doi.org/10.1137/130934283>.
- 363 [11] I. LIVSHITS, *A scalable multigrid method for solving indefinite helmholtz equations with constant*
 364 *wave numbers*, Numerical Linear Algebra with Applications, 21 (2014), <https://doi.org/10.1002/nla.1926>.
- 366 [12] I. LIVSHITS, *Multiple galerkin adaptive algebraic multigrid algorithm for the helmholtz equa-*
 367 *tions*, SIAM Journal on Scientific Computing, 37 (2015), pp. S195–S215, <https://doi.org/10.1137/140975310>, <https://doi.org/10.1137/140975310>, <https://arxiv.org/abs/https://doi.org/10.1137/140975310>.
- 369 [13] L. OLSON AND J. SCHRODER, *Smoothed aggregation for helmholtz problems*, Numerical Linear
 370 Algebra with Applications, 17 (2010), pp. 361 – 386, <https://doi.org/10.1002/nla.686>.
- 372 [14] L. N. OLSON, J. B. SCHRODER, AND R. S. TUMINARO, *A general interpolation strategy*
 373 *for algebraic multigrid using energy minimization*, SIAM Journal on Scientific Comput-
 374 ing, 33 (2011), pp. 966–991, <https://doi.org/10.1137/100803031>, <https://doi.org/10.1137/100803031>, <https://arxiv.org/abs/https://doi.org/10.1137/100803031>.
- 375 [15] G. STRANG, *Multigrid methods*, tech. report, MIT, 2006, <https://math.mit.edu/classes/18.086/2006/am63.pdf>.
- 377 [16] K. STÜBEN, *Algebraic multigrid (amg). an introduction with applications*, (1999).