SIAM CSE 2023 - 02/26 - 03/03

Rob FALGOUT (LLNL)
Matthieu LECOUVEZ (CEA/CESTA)
Pierre RAMET (INRIA)
Clément RICHEFORT (CEA/CESTA)

DE LA RECHERCHE À L'INDUSTRIE

Toward a multigrid method for the Helmholtz equation

In order to evaluate the electromagnetic behavior of 3D objects, we need to solve the **Maxwell equations**.

Numerical simulation of **Maxwell Equations** beforehand the conception

↓

Need to solve a large and sparse linear system of equations $Ax = b$

↓

Requires an **efficient** and **scalable** solver

A *domain decomposition* method is used in this context, but limitations appear as the computing resources are growing :

- Increase in the number of sub-domains $\Rightarrow$ Convergence is slower
- Increase in the size of the sub-domains $\Rightarrow$ Increase in computational complexity

We need to investigate an alternative method to domain decomposition :
multigrid methods.

Table of contents

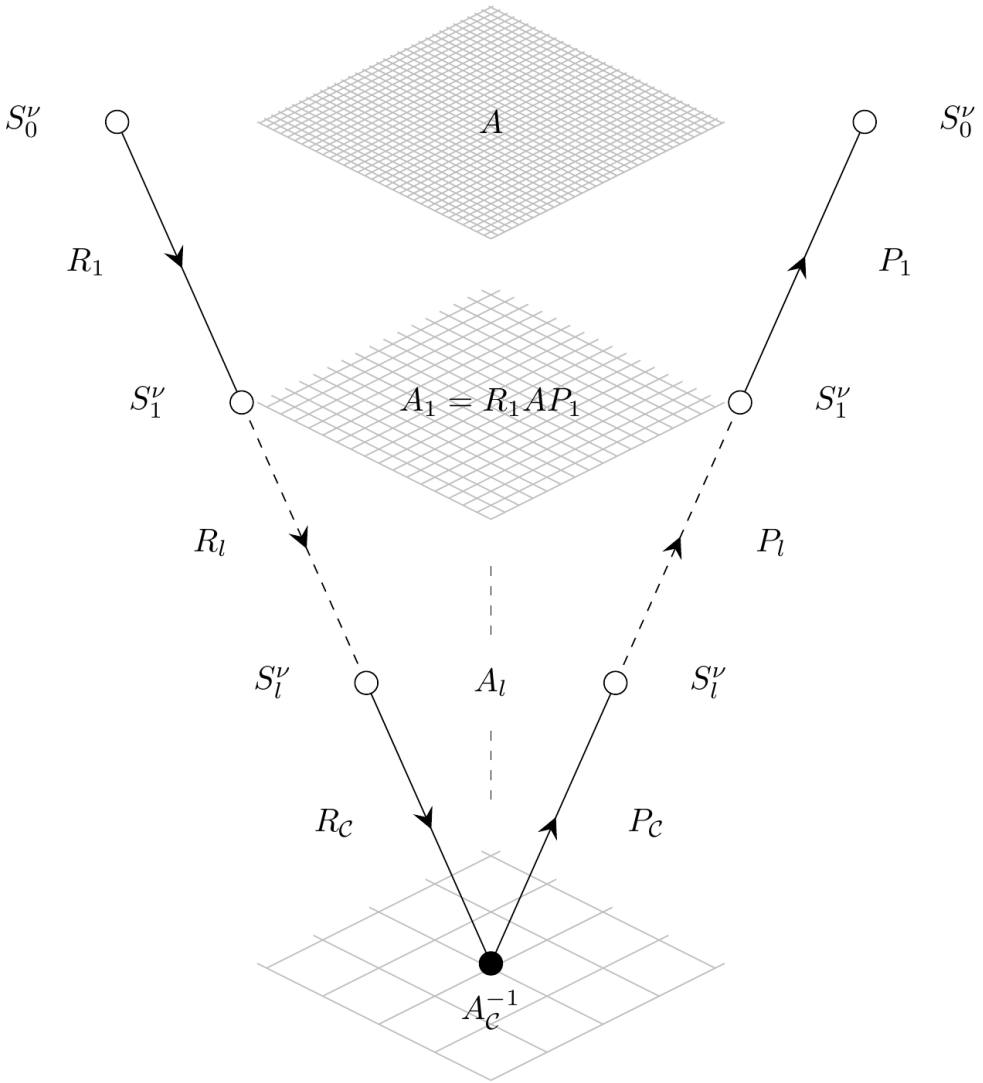


Figure 1: V-cycle of a multigrid method

In this method, the computation of the solution x is accelerated thanks to a hierarchy of coarse problems.

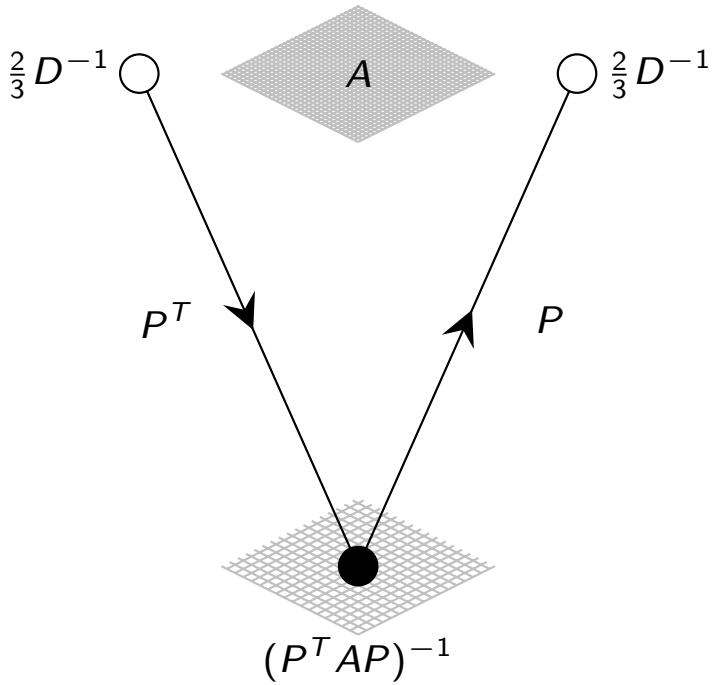
We introduce :

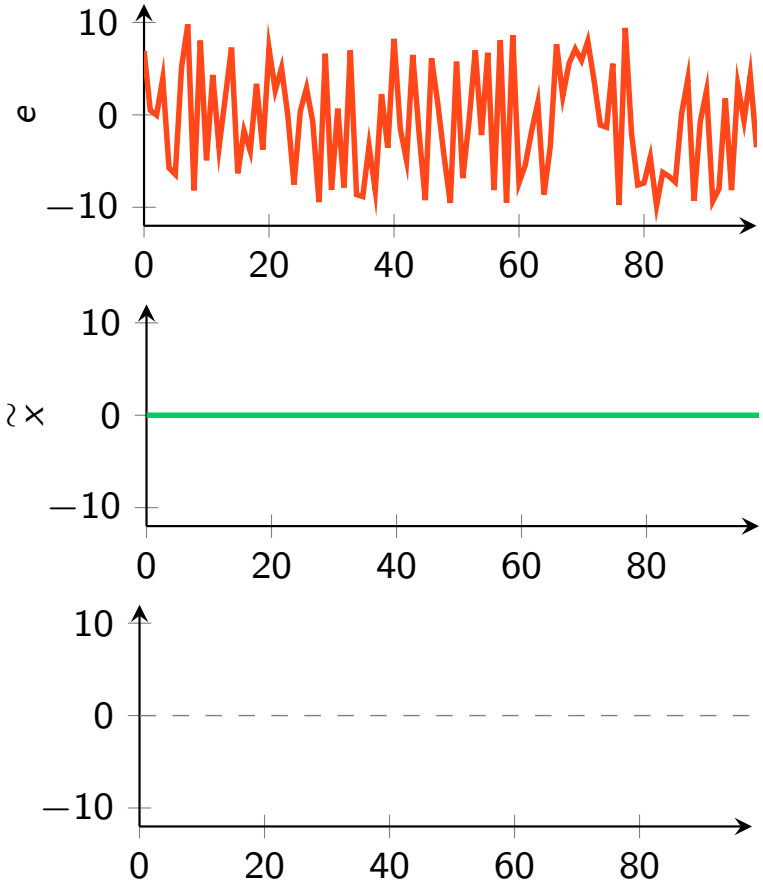
- A_l : Matrix of the level l
- S_0^ν : ν smoother iterations on the level l
- P_l : Interpolation operator of size $n_{l-1} \times n_l$
- R_l : Restriction operator of size $n_l \times n_{l-1}$

In most application

$$R = P^H. \tag{1}$$

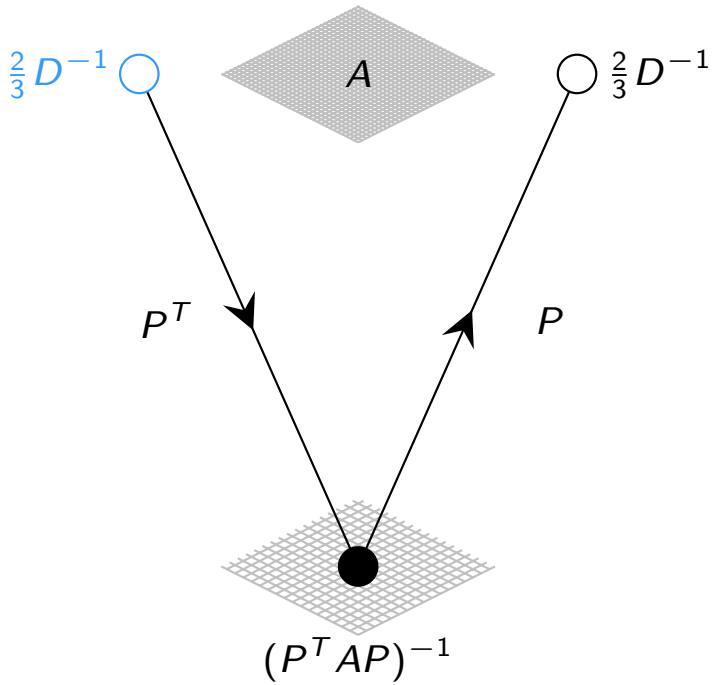
Let's illustrate one iteration of a two-level cycle on a 1D Laplace Problem. Here, A is SPD.



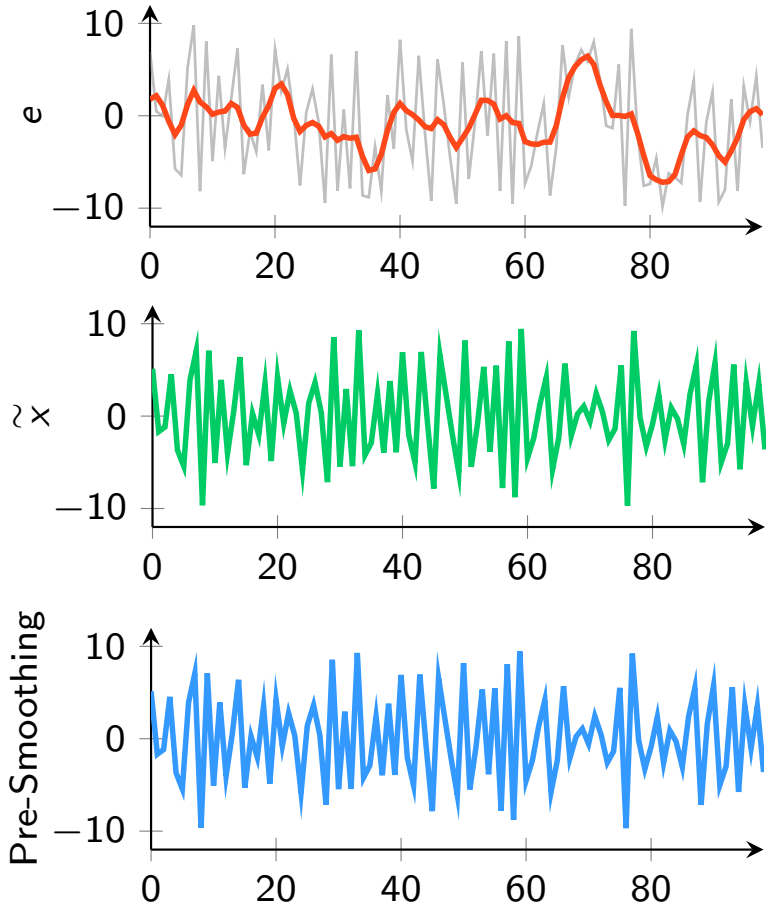


- Here, x is chosen randomly, and $b = Ax$.
- $\widetilde{x}$ approximation of x
- $e = x - \widetilde{x}$ is the remaining information to capture

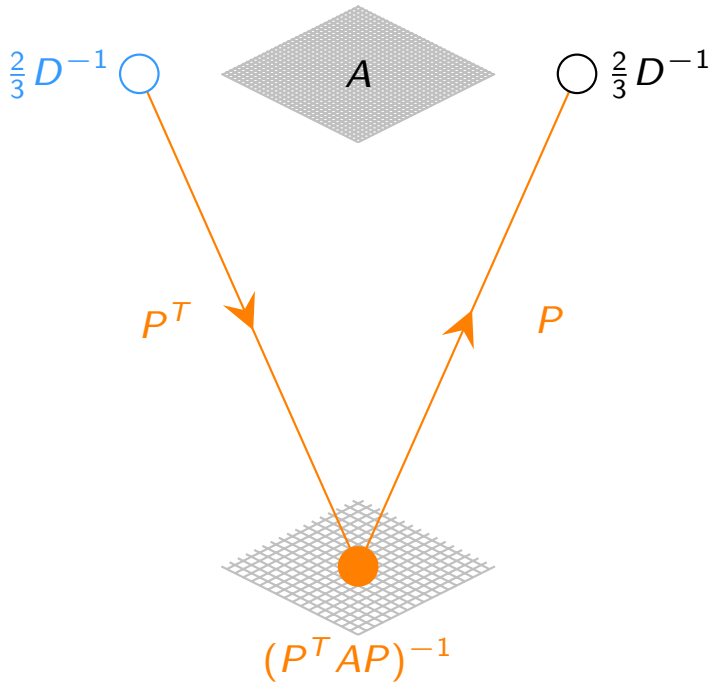
Let's illustrate one iteration of a two-level cycle on a 1D Laplace Problem. Here, A is SPD.



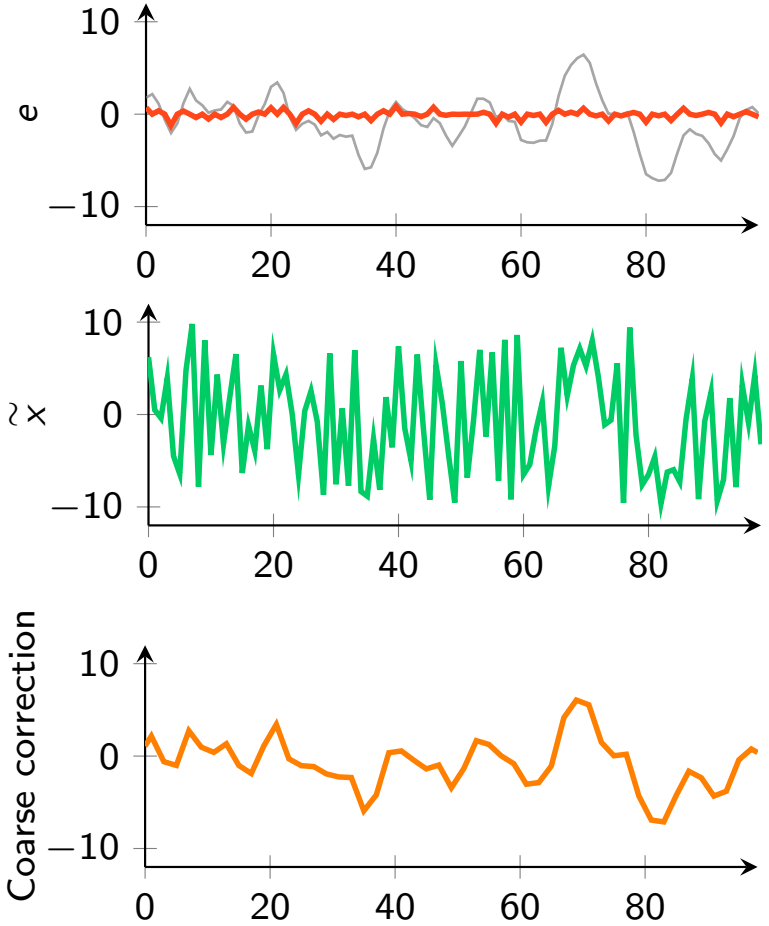
- $\tilde{x}$ approximation of x, $e = x - \tilde{x}$
- The smoother captured oscillatory information
- e is geometrically smooth



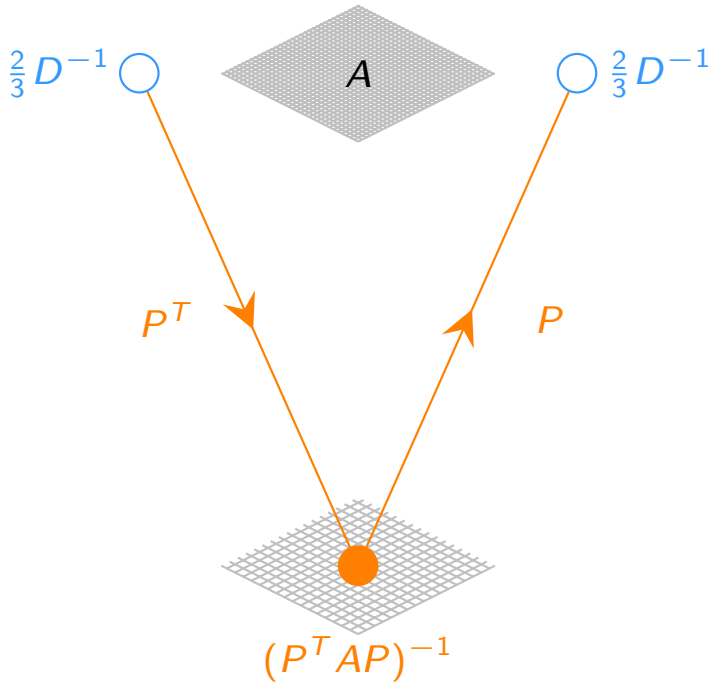
Let's illustrate one iteration of a two-level cycle on a 1D Laplace Problem. Here, A is SPD.



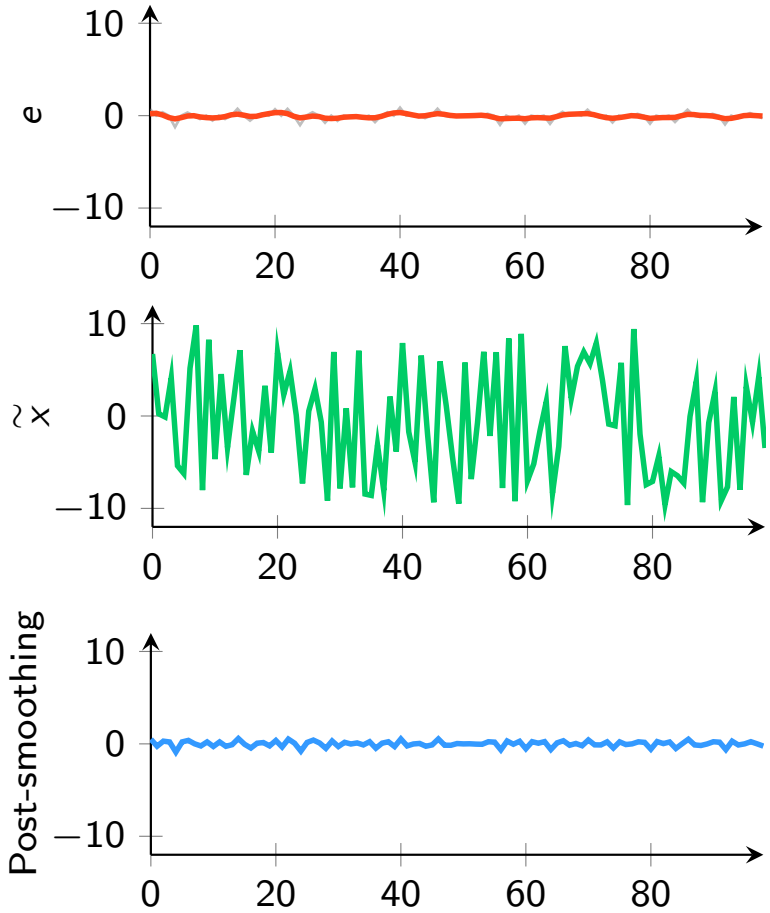
- $\tilde{x}$ approximation of x, $e = x - \tilde{x}$
- The coarse correction captured the smooth part
- e contains remaining high frequency information



Let's illustrate one iteration of a two-level cycle on a 1D Laplace Problem. Here, A is SPD.



- $\widetilde{x}$ approximation of x, $e = x - \widetilde{x}$
- The smoother captured remaining oscillations
- $\widetilde{x}$ is converging toward x !



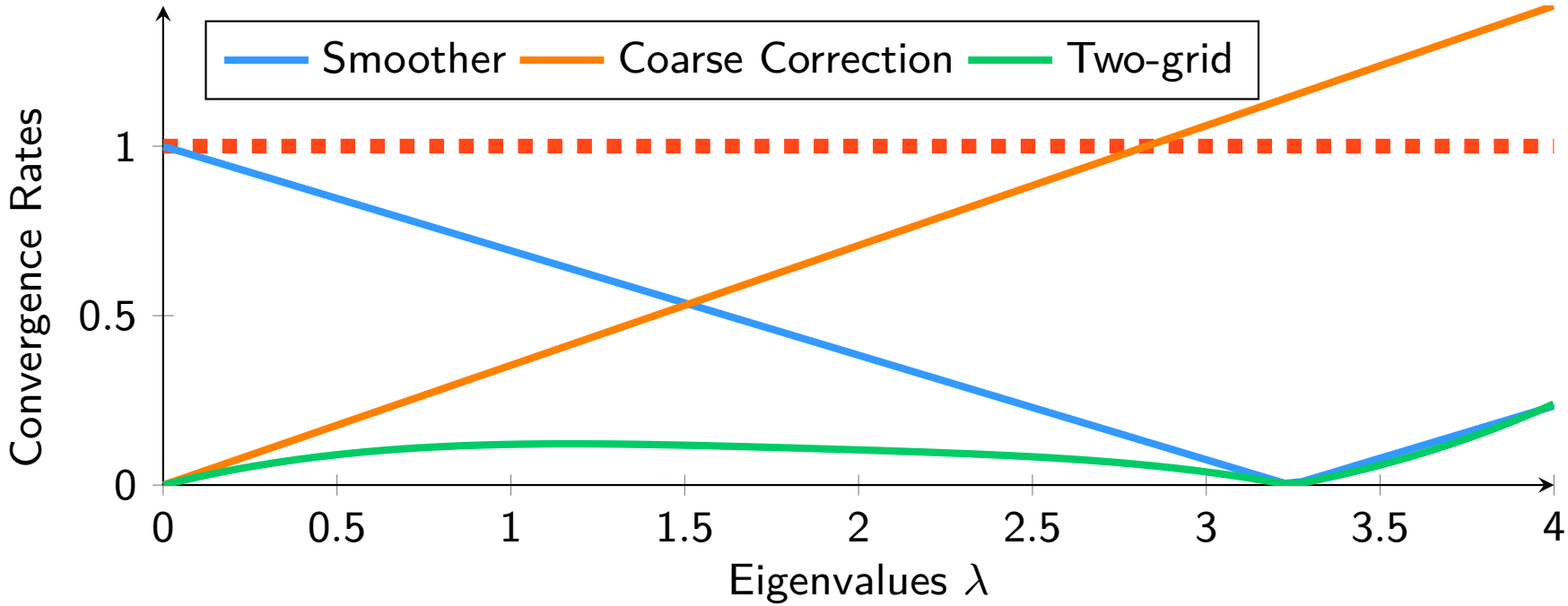


Figure 2: Convergence rates of the different operators

- The **lowest** eigenvalues are damped by the coarse correction
- The **highest** eigenvalues are damped by the smoother

$$\text{(Laplace Problem)} \quad \Leftrightarrow \quad \begin{cases} -\Delta u = f \text{ on } \Omega \\ u_{|\partial\Omega} = 0 \end{cases} \tag{2}$$

Applying a 2^{nd} order finite difference scheme on a uniform discretization of Ω in 1D, and using a local Fourier analysis, it yields

$$\forall j = 1, \dots, n \quad , \quad \lambda_j(A) = 2h^{-2}(1 - \cos(j\pi h)) \quad , \quad v_j(A) = [\sin(lj\pi h)]_{l=1}^{n}$$

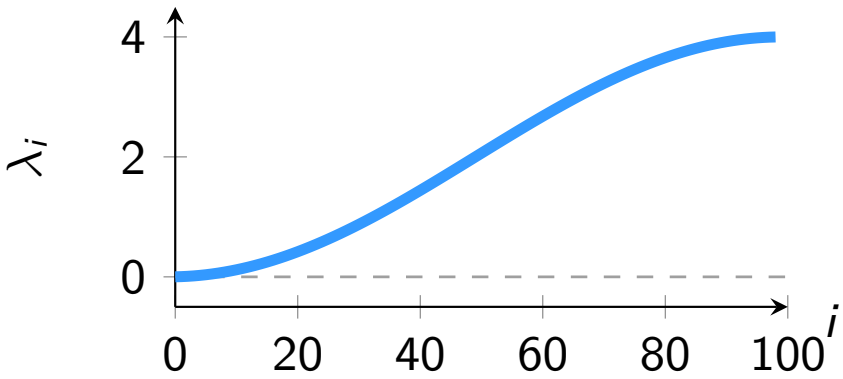


Figure 3: Eigenvalues

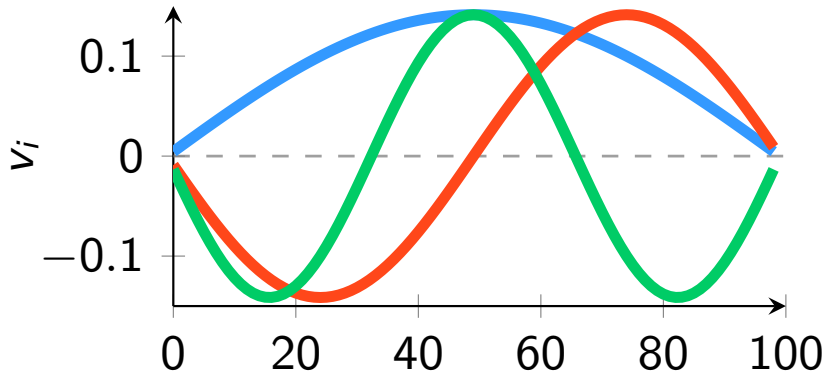


Figure 4: Few of the lowest eigenvectors

The complementarity principle is easy to satisfy

- $\lambda_j(A) > 0 \Rightarrow$ Usual relaxation methods are efficient for capturing HF eigenvectors
- The NKS is geometrically smooth $\Rightarrow$ Interpolation rules are easy to build in this case

$$\text{(Helmholtz Problem)} \quad \Leftrightarrow \quad \begin{cases} -\Delta u - k^2 u = f \text{ sur } \Omega = [0,1] \\ u_{|\partial\Omega} = 0 \end{cases} \tag{3}$$

Applying a 2nd order finite difference scheme on a uniform discretization of Ω in 1D, and using a local Fourier analysis, it yields

$$\forall j = 1, \ldots, n, \ \lambda_j(A) = 2h^{-2}(1 - \cos(j\pi h)) - k^2 \ , \quad v_j(A) = [\sin(lj\pi h)]_{l=1}^{n}$$

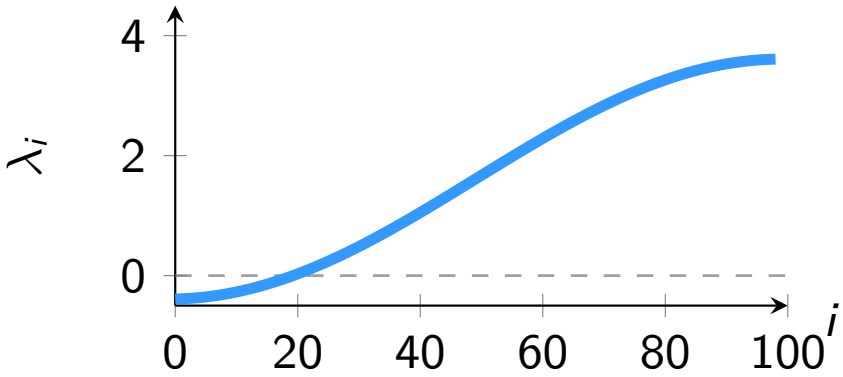


Figure 5: Eigenvalues

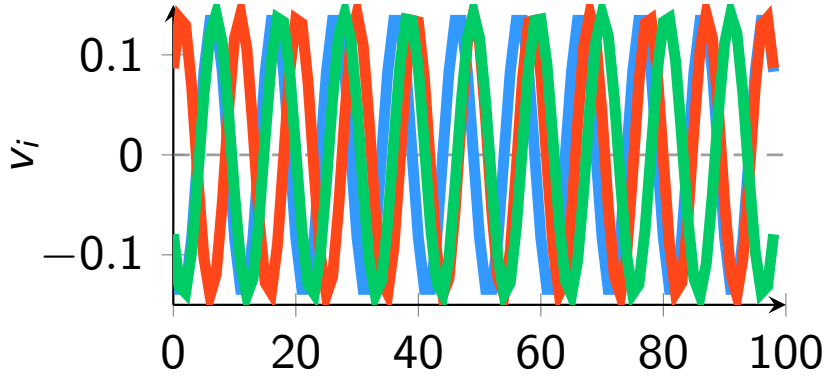


Figure 6: Few of the lowest eigenvectors

Eigenvalues are shifted, the complementarity principle is now difficult to satisfy

- ~~$\lambda_j(A) > 0$~~ $\Rightarrow$ Requires adapted smoothers : Krylov iterations, normal equations
- The NKS is now oscillatory $\Rightarrow$ Requires new interpolation rules

Table of contents

We seek a multilevel method able to solve a 2D Helmholtz Problem defined by

$$\text{(Helmholtz Problem)} \quad \Leftrightarrow \quad \begin{cases} -\Delta u - k^2 u = f \text{ on } \Omega \\ \partial_n u - iku = 0 \text{ on } \partial\Omega \end{cases} \tag{4}$$

Let the error propagation matrix for the coarse correction of a two-level cycle be

$$E_C = I - P(P^H A P)^{-1} P^H A. \tag{5}$$

Let the error propagation matrix for the smoother be

$$E_S = I - S^{-1} A \tag{6}$$

where S^{-1} is an approximation of A^{-1}.

Furthermore, we define :

- V_0 : Eigenvectors associated to the lowest eigenvalues in absolute values
- V_+ : Eigenvectors associated to the highest eigenvalues in absolute values

Target : Find a smoother able to damp V_+, without touching V_0.

Problem : V_+ eigenvectors are either associated with **negative** or **positive** eigenvalues.

Alternative to usual multigrid smoothers :

- *Krylov methods* are good smoothers in the *indefinite* case but :
  - They minimize $||r||_2$ regardless of the eigenvalues
  - They are non-linear because of their right-hand side dependence
- *Chebyshev Polynomial Smoother* built on *normal equations* will be considered
  - Normal equations are helpful to damp both negative and positive eigenvalues
  - The Chebyshev framework is practical to find a minimum polynomial within an interval
  - This smoother has the following error propagation formula

$$q(A^2) := I - p(A^2)A^2, \tag{7}$$

giving for $Av_0 = \lambda_0 v_0 \approx 0$

$$q(A^2)v_0 = (1 - p(\lambda_0^2)\lambda_0^2)v_0 \approx v_0 \tag{8}$$

$\rightarrow$ We seek a polynomial smoother p such that q is minimum in a given interval.

1. Choose an appropriate interval $\mathcal{I} = [x_{\min}, x_{\max}]$ where q must be minimum
2. Compute First Kind Chebyshev roots as best interpolation points within $\mathcal{I}$

$$x_i := \frac{x_{\max} + x_{\min}}{2} + \frac{x_{\max} - x_{\min}}{2} \cos\left(\frac{(2i-1)\pi}{2d}\right) \tag{9}$$

3. Construct the polynomial using the Lagrangian formula

$$q(x_i) = 0 \;\Leftrightarrow\; p(x_i) = \frac{1}{x_i}\,,\; p(x) = \sum_{j=0}^{d} \frac{1}{x_j} \prod_{i=0, i\neq j}^{d} \frac{x - x_i}{x_j - x_i} \tag{10}$$

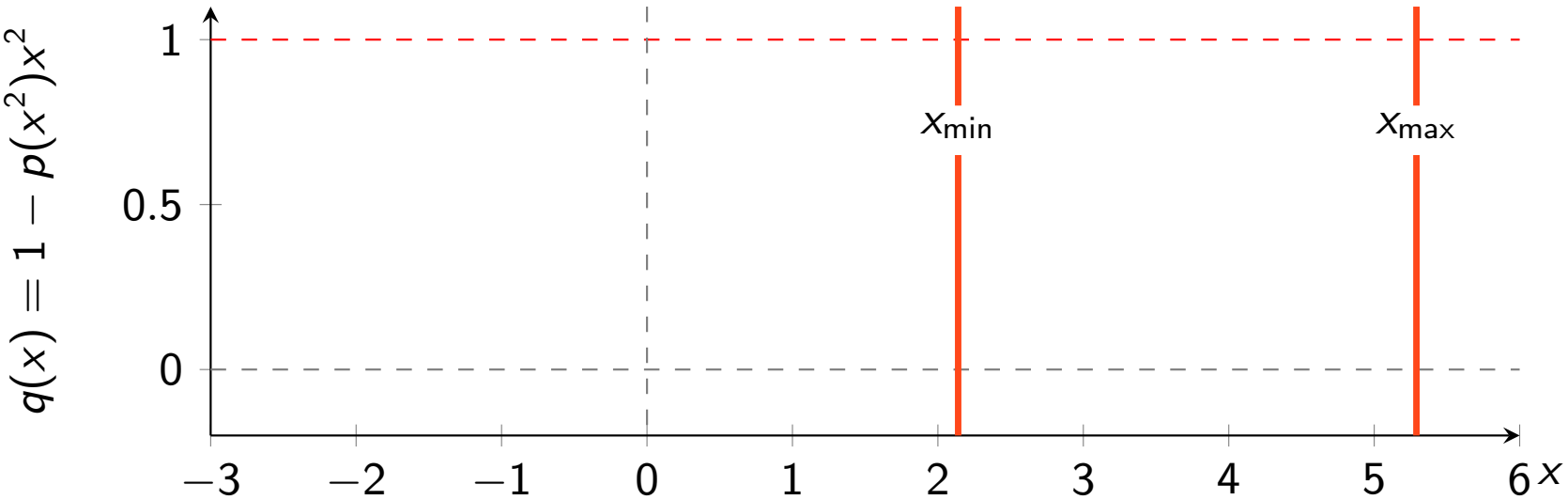


Figure 7: Spectrum of the polynomial smoother error propagation matrix

1. Choose an appropriate interval $\mathcal{I} = [x_{\min}, x_{\max}]$ where q must be minimum
2. Compute First Kind Chebyshev roots as best interpolation points within $\mathcal{I}$

$$x_i := \frac{x_{\max} + x_{\min}}{2} + \frac{x_{\max} - x_{\min}}{2} \cos\left(\frac{(2i-1)\pi}{2d}\right) \tag{11}$$

3. Construct the polynomial using the Lagrangian formula

$$q(x_i) = 0 \;\Leftrightarrow\; p(x_i) = \frac{1}{x_i}\,,\; p(x) = \sum_{j=0}^{d} \frac{1}{x_j} \prod_{i=0, i\neq j}^{d} \frac{x - x_i}{x_j - x_i} \tag{12}$$

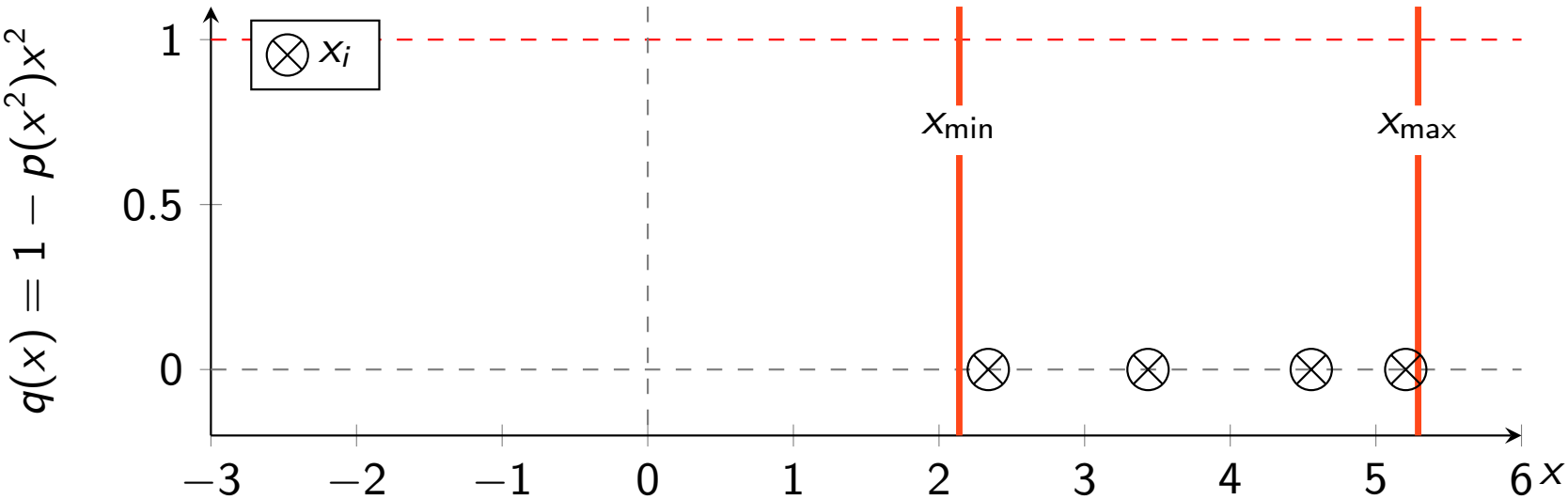


Figure 8: Spectrum of the polynomial smoother error propagation matrix

1. Choose an appropriate interval $\mathcal{I} = [x_{\min}, x_{\max}]$ where q must be minimum
2. Compute First Kind Chebyshev roots as best interpolation points within $\mathcal{I}$

$$x_i := \frac{x_{\max} + x_{\min}}{2} + \frac{x_{\max} - x_{\min}}{2} \cos\left(\frac{(2i-1)\pi}{2d}\right) \tag{13}$$

3. Construct the polynomial using the Lagrangian formula

$$q(x_i) = 0 \;\Leftrightarrow\; p(x_i) = \frac{1}{x_i}\,,\; p(x) = \sum_{j=0}^{d} \frac{1}{x_j} \prod_{i=0, i\neq j}^{d} \frac{x - x_i}{x_j - x_i} \tag{14}$$

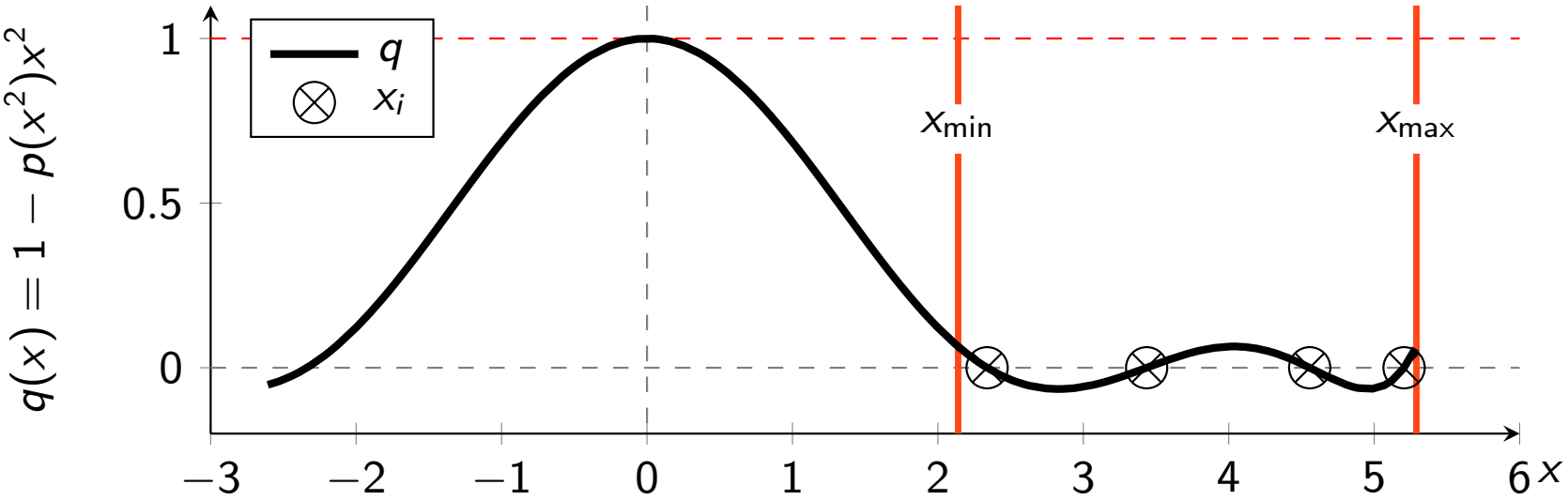


Figure 9: Spectrum of the polynomial smoother error propagation matrix

Table of contents

Target : Construct an interpolator P containing V_0 in its range

From the theory, we can define, under SPD assumption of A, an ideal interpolator P_* from coarse ($\mathcal{C}$) and fine ($\mathcal{F}$) variable selection operators

$$\underset{(n_\mathcal{C} \times n)}{R} : \Omega \mapsto \mathcal{C} \text{ and } \underset{(n_\mathcal{F} \times n)}{S^H} : \Omega \mapsto \mathcal{F}, \tag{15}$$

such that

$$n_\mathcal{C} + n_\mathcal{F} = n \text{ and } RS = 0. \tag{16}$$

This operator is defined by

$$P_* = (I - S(S^H A S)^{-1} S^H A) R^H \tag{17}$$

and minimizes the quantity

$$\mu_X = \min_P \max_{e \neq 0} \frac{\langle X(I - PR)e, (I - PR)e \rangle}{\langle Ae, e \rangle}. \tag{18}$$

where X is a matrix defining the smoothing space.

Target : Construct an interpolator P containing V_0 in its range

Exemple 1 : Let the coarse and fine selection operators be defined by

$$R = [\, 0 \;\; I_{n_{\mathcal{C}}} \,] \text{ and } S^T = [\, I_{n_{\mathcal{F}}} \;\; 0 \,]. \tag{19}$$

After reorganizing A by coarse/fine blocks, such that

$$A = \begin{bmatrix} S^H AS & S^H AR^H \\ RAS & RAR^H \end{bmatrix} = \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix},$$

the ideal interpolator can be written

$$P_* = \begin{bmatrix} -A_{ff}^{-1} A_{fc} \\ I_{cc} \end{bmatrix} \tag{20}$$

and the coarse matrix

$$A_{\mathcal{C}} = P_*^H AP_* = A_{cc} - A_{cf} A_{ff}^{-1} A_{fc} \tag{21}$$

Here we can prove

$$E_* E_{\mathcal{F}} = (I - P_*(P_*^H AP_*)^{-1} P_*^H A)(I - S(S^H AS)^{-1} S^H A) = 0 \tag{22}$$

Target : Construct an interpolator P containing V_0 in its range

Example 2 : Let the coarse and fine selection operators be defined by

$$R^H = V_0 \text{ and } S = V_+, \tag{23}$$

with V_0/V_+ eigenvectors associated with lowest/highest eigenvalues. It yields

$$S^H A R^H = V_+^H \begin{bmatrix} V_0 & V_+ \end{bmatrix} \begin{bmatrix} \text{Diag}(\lambda_0) & 0 \\ 0 & \text{Diag}(\lambda_+) \end{bmatrix} \begin{bmatrix} V_0^H \\ V_+^H \end{bmatrix} V_0 = 0,$$

so the ideal interpolator is defined by

$$P_* = (I - S(S^H A S)^{-1} S^H A) R^H = R^H = V_0 \tag{24}$$

and gives the coarse matrix

$$A_{\mathcal{C}} = \text{Diag}(\lambda_0) \tag{25}$$

Here again

$$E_* E_{\mathcal{F}} = (I - P_*(P_*^H A P_*)^{-1} P_*^H A)(I - S(S^H A S)^{-1} S^H A) = 0 \tag{26}$$

Target : Construct an interpolator P containing V_0 in its range

Remark : The ideal framework requires A to be SPD.

$\hookrightarrow$ **However the reduction viewpoint is still valid :** Removing orthogonal information that the smoother captures will correct the coarse selection operator

Question : Are coarse selection operators in both previous examples good for Helmholtz?

- $R^H = [\, 0 \;\; I_{n_C} \,]^T$ is not relevant for Helmholtz
- $R^H = V_0$ is not practical (too expensive to compute and we need sparsity)

We can construct a better coarse selection operator R^H for Helmholtz.

Idea : Construct fine interpolation rules $R_{\mathcal{F}}$ of R^H with a least-squares minimization strategy from smoothed random vectors approximating the near-kernel space V_0

First, we generate a set K of κ smoothed random vectors that approximates V_0

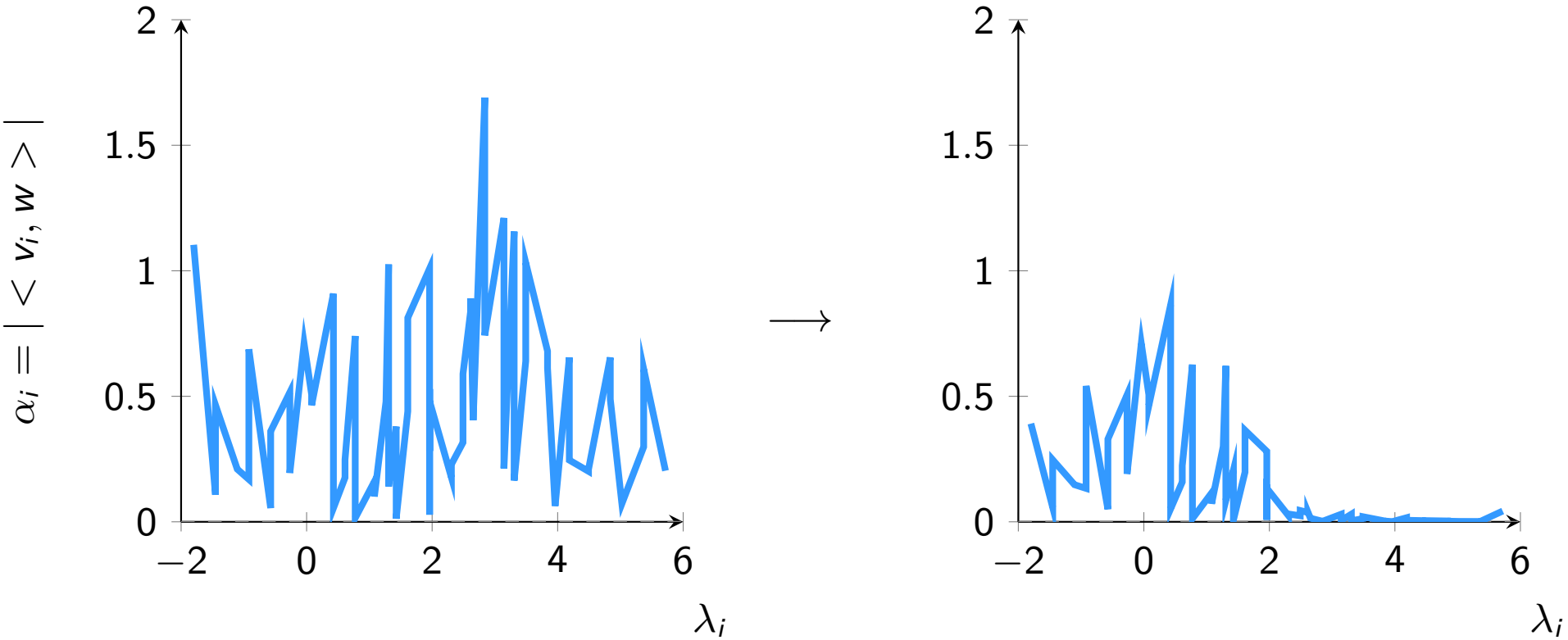


Figure 10: Eigen decomposition of a random vector (left) vs. smoothed random vector (right)

First, we generate a set K of κ smoothed random vectors that approximates V_0

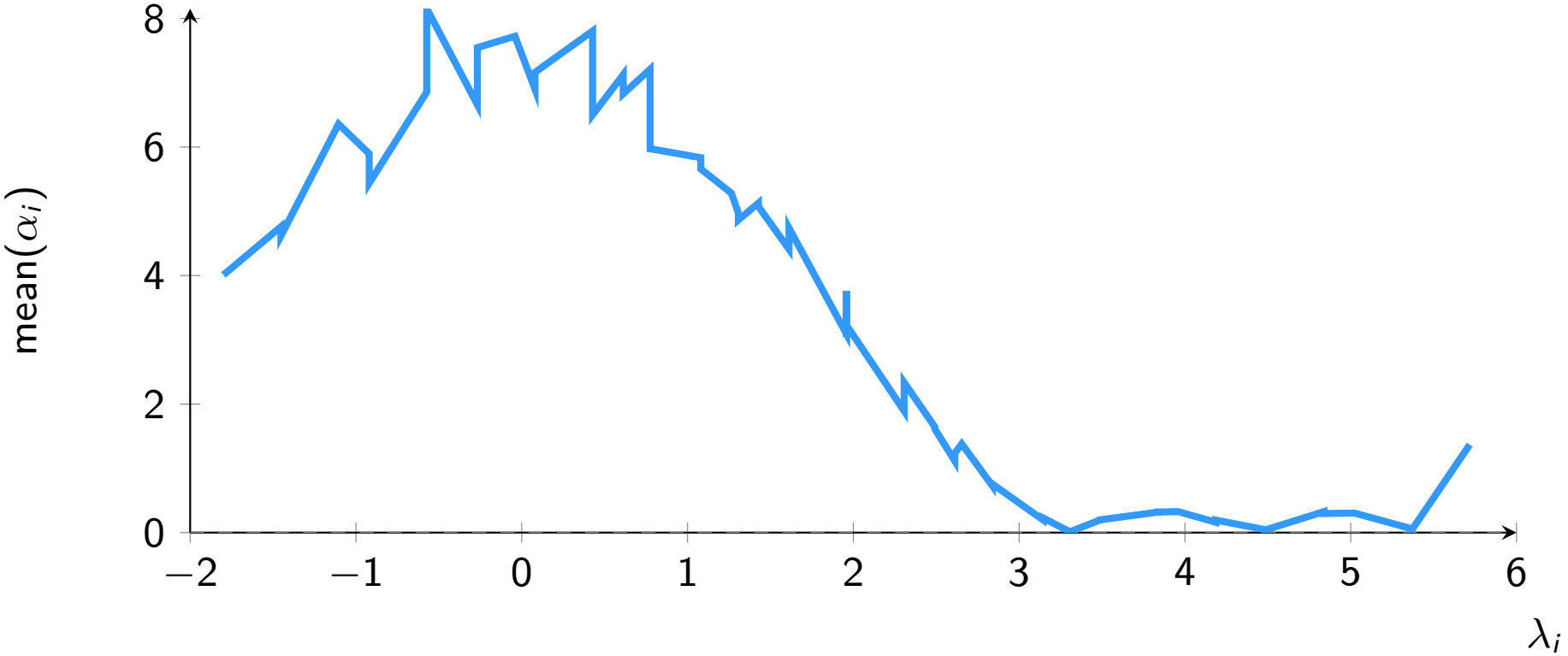


Figure 11: Average eigen decomposition of the total set K

The Least-Squares Minimization strategy consists of constructing each **fine interpolation rule** r_i by minimizing the squared difference between **fine values of the near-kernel candidate vectors** and the **interpolation from their connected coarse variables** $\mathcal{C}_i$.

$$\forall i \in \mathcal{F} \ , \ r_i = \arg\min_r \sum_{l=1}^{\kappa} w_l \left(K_{i,l} - r \cdot K_{\mathcal{C}_i,l}\right)^2 := \arg\min_r \mathcal{L}_i(r) \tag{27}$$

Finding the minimum of the convex loss function $\mathcal{L}_i$ is equivalent to solving

$$\nabla \mathcal{L}_i(r_i) = 0. \tag{28}$$

Equation (28) can be rewritten element-wise

$$\frac{\partial \mathcal{L}_i(r_i)}{\partial r_{ij}} = \sum_{l=1}^{\kappa} 2w_l(K_{i,l} - r_i \cdot K_{\mathcal{C}_i,l})K_{\mathcal{C}_{ij},l} = 0 \ , \ \forall j \in [1, \text{card}(C_i)]. \tag{29}$$

Finally, (29) leads to a system of linear equations to solve for each fine variable i

$$r_i K_{\mathcal{C}_i} W K_{\mathcal{C}_i}^H = K_i W K_{\mathcal{C}_i}^H \tag{30}$$

Using the ideal framework and the LSM strategy, let the coarse and fine selection operators

$$\hat{R}^H = [\ R_{\mathcal{F}}\ \ I_{\mathcal{C}}\ ]^T\ ,\ \hat{S} = [\ I_{\mathcal{F}}\ \ -R_{\mathcal{F}}^H\ ]^T, \tag{31}$$

where $\hat{R}^H$ is the least squares operator with $R_{\mathcal{F}}$ its fine variable interpolation block. We define $A_{\mathcal{F}} := \hat{S}^H A \hat{S}$. Following the ideal interpolator definition

$$\hat{P} := (I - \hat{S} A_{\mathcal{F}}^{-1} \hat{S}^H A)\hat{R}^H \tag{32}$$

However, $\hat{P}$ requires the inverse of $A_{\mathcal{F}}$! Instead

$$\hat{P} \approx \hat{R}^H - \hat{S} X_{\mathcal{K}}^{-1} \hat{S}^H A \hat{R}^H, \tag{33}$$

where $X_{\mathcal{K}}^{-1}$ is the best polynomial approximating $A_{\mathcal{F}}^{-1}$ within the Krylov subspace $\mathcal{K}$. In practice, $\hat{P}$ is approximated column-wise under sparsity constraints

$$\mathcal{K}_{\mathcal{P}_i}^m = \left\{ Z_i b_i\ ,\ Z_i A_{\mathcal{F}} Z_i^T Z_i b_i\ ,\ \dots\ ,\ (Z_i A_{\mathcal{F}} Z_i^T)^{m-1} Z_i b_i \right\}, \tag{34}$$

where $b_i := \hat{S}^H A \hat{R}^H_{:,i}$, and $Z_i : \mathbb{C}^n \to \mathbb{C}^{\mathrm{Card}(\mathcal{P}_i)}$ is the i^{th} column associated sparsity constraint operator, that restricts any full vector to the non-zero pattern $\mathcal{P}_i$.

- 5-points stencil Cartesian discretization with a.b.c ($\partial_n u - iku = 0$ on $\partial\Omega$)
- 10 points per wavelength ($h = \lambda/10 \Leftrightarrow kh = 2\pi/10 \approx 0.625$)
- Multigrid parameters :
  - Normal equations polynomial smoother of degree $d = 3$
  - Number of test vectors tends to grow by a factor of 2 between each level
  - Krylov sub-size $m = 3$ in the construction of $\hat{P}$

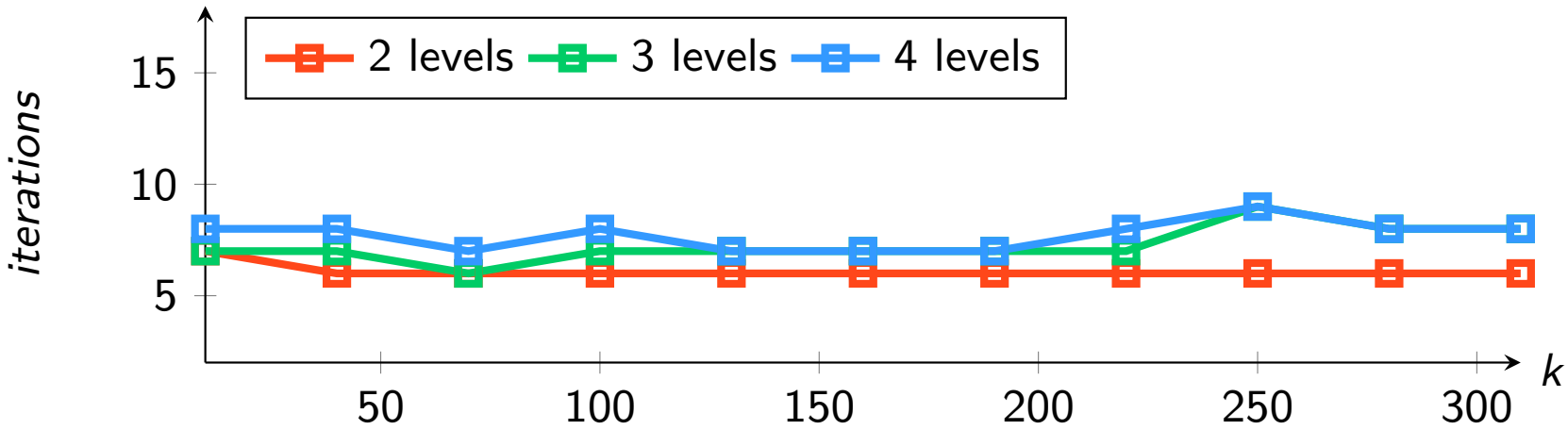


Figure 12: Number of iterations following the wavenumber k

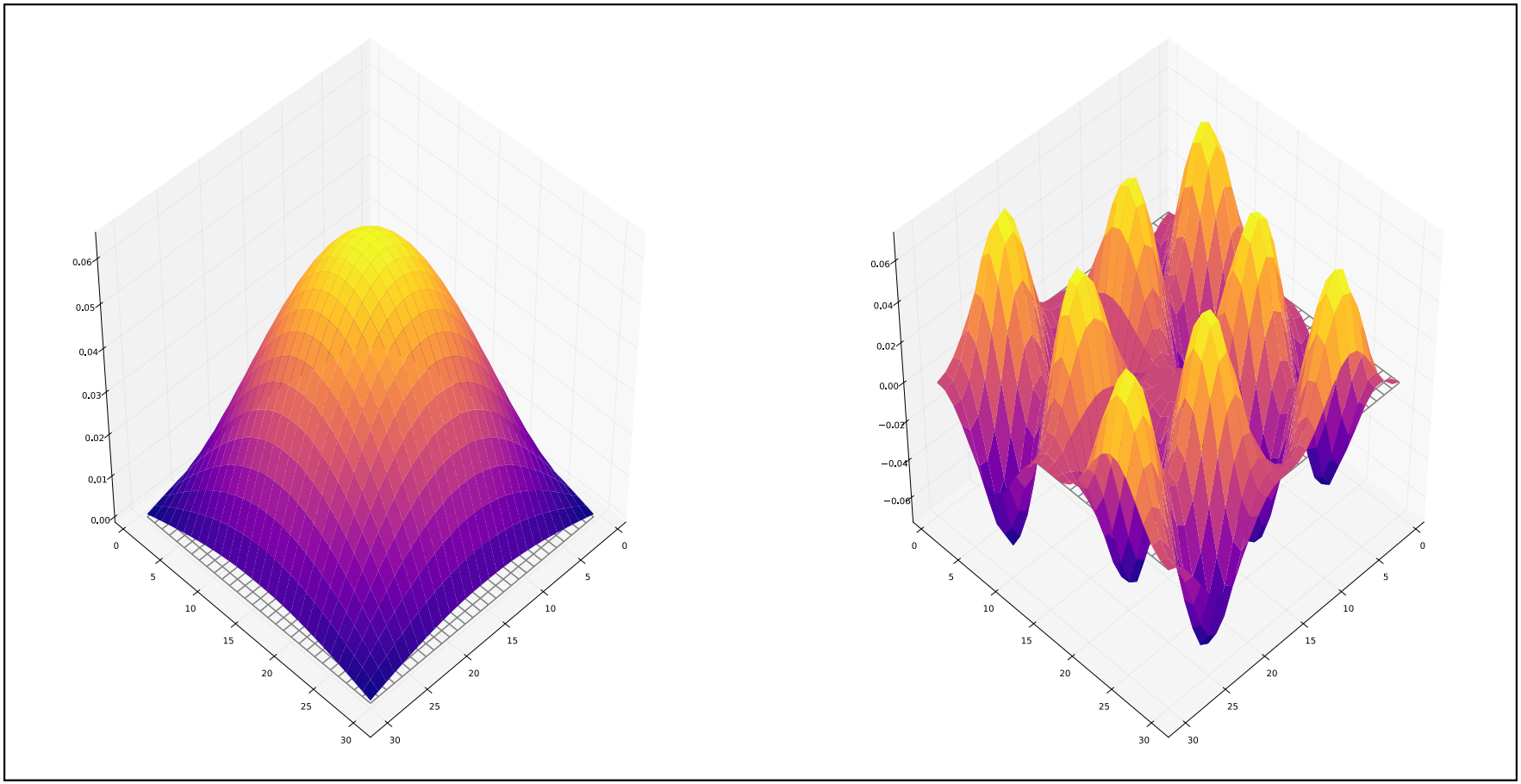


Figure 13: Solution x of a Laplace Problem vs. a Helmholtz Problem

Helmholtz is difficult because :

- Negative components appear and require an adapted smoother
- The near-kernel space is oscillatory and needs adapted interpolation rules

$\Rightarrow$ We proposed an algebraic multigrid method that reaches those requirements up to a certain limit.

Further researches will focus on :

1. Increasing the depth of the multigrid cycle
2. Sparsity has to be improved
3. Construct a better coarse selection operator that minimizes $\sum_i^{\kappa} ||K_{:,i} - R^H R K_{:,i}||_2$
4. Work on an adapted framework for Helmholtz