Wavelet Approximations of Boundary Value Problems

Consider the problem

$$\begin{cases} Au = f & \text{in } \Omega \\ \\ Bu = 0 & \text{on } \partial\Omega \ . \end{cases}$$

where A, B are linear differential operators.

Assume that the problem can be formulated as:

Given $f \in V'$, find $u \in V$ such that

$$\langle Au, v \rangle = \langle f, v \rangle, \qquad \forall v \in V \ ,$$

where

- $V = V^{(r)}$ is a Hilbert space of regularity $r > 0$, continuously embedded in $L^2(\Omega)$, which may incorporate boundary conditions

- V' is the dual space of V

- $\langle \cdot \ , \ \cdot \rangle$ is the duality pairing between V' and V.

Assumption:

$A : V \to V'$ is an algebraic and topological isomorphism, so that

$$\|Av\|_{V'} \asymp \|v\|_V, \qquad \forall v \in V \ .$$

This holds, e.g., if A is continuous from V to V' and coercive on V, i.e., there exists $\alpha > 0$ such that

$$\langle Av, v \rangle \geq \alpha \|v\|_V^2, \qquad \forall v \in V.$$

Examples

1. *Dirichlet Problem for the Poisson Equation*

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \ . \end{cases}$$

Set $V = H_0^1(\Omega) = \{v \in H^1(\Omega) \ : \ v = 0 \text{ on } \partial\Omega\}$ equipped with the norm

$$\|v\|_{H_0^1(\Omega)} = \left(\int_\Omega |\nabla v|^2 \, dx \right)^{1/2} .$$

Then,

$$\langle Au, v\rangle := \int_\Omega \nabla u \cdot \nabla v \, dx, \qquad \langle f, v\rangle := \int_\Omega f \, v \ dx$$

with $f \in L^2(\Omega)$.

2. *Neumann Problem for the Helmholtz Equation*

$$\begin{cases} -\Delta u + u = f & \text{in } \Omega \\ \dfrac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega \ . \end{cases}$$

Set $V = H^1(\Omega)$ equipped with the norm

$$\|v\|_{H^1(\Omega)} = \left(\int_\Omega |\nabla v|^2 \, dx + \int_\Omega |v|^2 \, dx \right)^{1/2} .$$

Then,

$$\langle Au, v\rangle := \int_\Omega \nabla u \cdot \nabla v \, dx + \int_\Omega u \, v \ dx, \qquad \langle f, v\rangle := \int_\Omega f \, v \ dx \ .$$

Let us introduce a multilevel (wavelet) basis in V:

$$V = \text{span}\,\{\psi_\lambda \ : \ \lambda \in \mathcal{M}\}$$

and let us set

$$u = \sum_{\lambda\in\mathcal{M}} u_\lambda \psi_\lambda\,, \qquad \boldsymbol{u} = (u_\lambda)$$

$$a_{\mu\lambda} = \langle A\psi_\lambda, \psi_\mu\rangle, \qquad \boldsymbol{A} = (a_{\mu\lambda})$$

$$f_\mu = \langle f, \psi_\mu\rangle, \qquad \boldsymbol{f} = (f_\mu)\,.$$

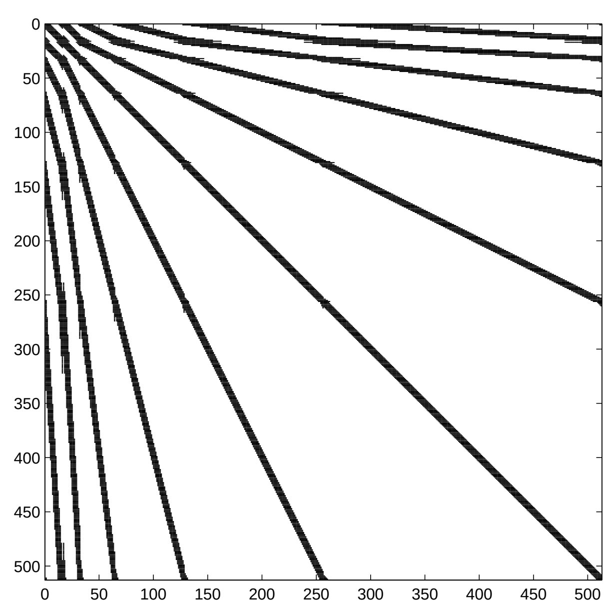
Then, the operator equation can be written as the (infinite order) algebraic system

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f}$$

Ordering the hierarchical basis functions by levels

$$(\psi_\lambda) = (\psi_{jk}) = ((\psi_{jk}), k \in \mathcal{K}_j), j \geq j_0)\,,$$

the matrix $\boldsymbol{A}$ has a typical non-banded, sparse block structure:



(compare with the banded matrix of a non-hierarchical basis)

Discrete representations of Operators

Let $V = V^{(r)}$ be a Hilbert space of regularity r and let $\Psi = \{\psi_\lambda\}_{\lambda \in \mathcal{M}}$ be a Riesz basis in V:

$$\|v\|_V \asymp \|\boldsymbol{v}\|_{\ell^2(\mathcal{M})}, \qquad \forall v = \sum_{\lambda \in \mathcal{M}} v_\lambda \psi_\lambda =: \boldsymbol{v}^T \Psi \in V.$$

Similarly, let $W = W^{(s)}$ be a Hilbert space of regularity s and let $\tilde{\Psi} = \{\tilde{\psi}_\lambda\}_{\lambda \in \mathcal{M}}$ be a Riesz basis in W.

Let $A : V \to W$ be an algebraic and topological isomorphism,

$$\|Av\|_W \asymp \|v\|_V, \qquad \forall v \in V\ .$$

For any $\lambda \in \mathcal{M}$, let

$$A\psi_\lambda =: \sum_{\mu \in \mathcal{M}} a_{\mu\lambda} \tilde{\psi}_\mu$$

be the expansion of $A\psi_\lambda \in W$ according to the Riesz basis $\tilde{\Psi}$.

Set

$$\boldsymbol{A} = (a_{\mu\lambda})_{\mu,\lambda \in \mathcal{M}}$$

Then,

$$v = \boldsymbol{v}^T \Psi \in V \qquad \text{iff} \qquad Av = (\boldsymbol{A}\boldsymbol{v})^T \tilde{\Psi} \in W$$

and

$$\boldsymbol{A} : \ell^2(\mathcal{M}) \to \ell^2(\mathcal{M})$$

is an algebraic and topological isomorphism.

Finite Dimensional (Galerkin) Approximation

Choose a finite subset $\Lambda \subset \mathcal{M}$. Set $N = \text{card}\,\Lambda$ and

$$V_\Lambda = \text{span}\,\{\psi_\lambda \ : \ \lambda \in \Lambda\}$$

Consider the Galerkin approximation

$$\begin{aligned} &Find\ u_\Lambda \in V_\Lambda\ such\ that \\ &\langle A u_\Lambda, v_\Lambda \rangle = \langle f, v_\Lambda \rangle, \qquad \forall v_\Lambda \in V_\Lambda\,, \end{aligned}$$

equivalent to the $N \times N$ algebraic system

$$\boldsymbol{A}_\Lambda \boldsymbol{u}_\Lambda = \boldsymbol{f}_\Lambda$$

with

$$\boldsymbol{A}_\Lambda = (a_{\mu\lambda})_{\mu,\lambda\in\Lambda}$$

$$\boldsymbol{u}_\Lambda = (u_{\Lambda,\lambda})_{\lambda\in\Lambda} \quad \text{if} \quad u_\Lambda = \sum_{\lambda\in\Lambda} u_{\Lambda,\lambda}\psi_\lambda$$

$$\boldsymbol{f}_\Lambda = (f_\mu)_{\mu\in\Lambda}\,.$$

The multilevel structure of the basis can be exploited in

- preconditioning
- compression
- adaptivity

Preconditioning

Basic facts from linear algebra:

Let $\boldsymbol{A}$ be a $N \times N$ non-singular matrix. Define

$$\mathsf{cond}_2(\boldsymbol{A}) := \|\boldsymbol{A}\|_2 \|\boldsymbol{A}^{-1}\|_2.$$

Assume that $\boldsymbol{A}$ is symmetric positive-definite, with eigenvalues $\lambda > 0$. Then

$$\mathsf{cond}_2(\boldsymbol{A}) = \frac{\lambda_{\max}}{\lambda_{\min}}.$$

Iterative solution of

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

by Richardson iterations

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \alpha(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^k), \qquad k = 0, 1, \ldots$$

for some $\alpha > 0$.

The error $\boldsymbol{e}^k = \boldsymbol{x}^k - \boldsymbol{x}$ satisfies

$$\boldsymbol{e}^{k+1} = (\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{e}^k,$$

whence

$$\|\boldsymbol{e}^k\|_2 \leq \rho^k \|\boldsymbol{e}^0\|_2$$

with

$$\rho := \rho(\boldsymbol{I} - \alpha\boldsymbol{A}) < 1 \qquad \text{provided} \quad \alpha < \frac{2}{\lambda_{\max}}.$$

The factor ρ is minimized for

$$\alpha_{\text{opt}} := \frac{2}{\lambda_{\text{max}} + \lambda_{\text{min}}},$$

for which

$$\rho_{\text{opt}} = \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{\lambda_{\text{max}} + \lambda_{\text{min}}} = \frac{\text{cond}_2(\boldsymbol{A}) - 1}{\text{cond}_2(\boldsymbol{A}) + 1}$$

If $\text{cond}_2(\boldsymbol{A}) \gg 1$, the convergence is slow.

Remedy: Preconditioning. Replace

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

by

$$\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{P}^T\boldsymbol{b}$$

i.e.,

$$\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}\boldsymbol{y} = \boldsymbol{P}^T\boldsymbol{b}, \qquad \boldsymbol{y} = \boldsymbol{P}^{-1}\boldsymbol{x},$$

where the matrix $\boldsymbol{P}$ is chosen in such a way that the symmetric matrix

$$\boldsymbol{B} := \boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$$

satisfies

$$\text{cond}_2(\boldsymbol{B}) \ll \text{cond}_2(\boldsymbol{A}).$$

Preconditioning is easily accomplished by wavelets.

[Jaffard '92, Dahmen - Kunoth '92, Masson '98]

Assume that

- A is a symmetric, continuous and coercive operator on V.

- The wavelet basis $\{\psi_\lambda\}_{\lambda\in\mathcal{M}}$ is normalized in $L^2(\Omega)$

Theorem. *$\boldsymbol{A}_\Lambda$ is spectrally equivalent to the diagonal matrix*

$$\boldsymbol{D}_\Lambda = \text{diag}\,(2^{2r|\lambda|})_{\lambda\in\Lambda}\,.$$

Indeed, for any $\boldsymbol{v}_\Lambda = (v_\lambda)_{\lambda\in\Lambda}$, set $v = \sum_{\lambda\in\Lambda} v_\lambda \psi_\lambda \in V$. Then

$$\begin{aligned}\boldsymbol{v}_\Lambda^T \boldsymbol{A}_\Lambda \boldsymbol{v}_\Lambda &= \langle Av, v\rangle \asymp \|v\|_V^2 \\ &\asymp \sum_{\lambda\in\Lambda} 2^{2r|\lambda|}|v_\lambda|^2 = \boldsymbol{v}_\Lambda^T \boldsymbol{D}_\Lambda \boldsymbol{v}_\Lambda.\end{aligned}$$

Corollary. *Set*

$$\boldsymbol{B}_\Lambda = \boldsymbol{D}_\Lambda^{-1/2} \boldsymbol{A}_\Lambda \boldsymbol{D}_\Lambda^{-1/2}$$

Then,

$$\text{cond}_2(\boldsymbol{B}_\Lambda) \asymp 1$$

for all $\Lambda \subset \mathcal{M}$.

Equivalently, normalize the wavelets in V rather than in $L^2(\Omega)$:

$$\psi_\lambda^* := 2^{-r|\lambda|}\psi_\lambda\,.$$

Then, $$\boldsymbol{B}_\Lambda = (\,\langle A\psi_\lambda^*, \psi_\mu^*\,\rangle)_{\mu,\lambda\in\Lambda}.$$

For any $\boldsymbol{v}_\Lambda = (v_\lambda)_{\lambda\in\Lambda}$, set $v^* = \sum_{\lambda\in\Lambda} v_\lambda \psi_\lambda^*$. One has

$$\|v^*\|_V \asymp \|\boldsymbol{v}_\Lambda\|_{\ell^2(\Lambda)} \qquad \text{and,} \qquad \boldsymbol{v}_\Lambda^T \boldsymbol{B}_\Lambda \boldsymbol{v}_\Lambda \asymp \|\boldsymbol{v}_\Lambda\|_{\ell^2(\Lambda)}^2\,.$$

Improvements:

- take $\boldsymbol{D}_\Lambda = \text{diag}\, \boldsymbol{A}_\Lambda$ [Masson '97]

- Exploit the possibility of compressing the inverse $\boldsymbol{A}_\Lambda^{-1}$ using as $\boldsymbol{D}_\Lambda$ a 'Sparse Approximate Inverse' of $\boldsymbol{A}_\Lambda$ [Masson '97, Chan - Tang - Wan '97]

Remark:
Multilevel preconditioning is achieved also in other frameworks, such as:

- multigrid methods [Brandt '77, Hackbusch '85]

- hierarchical finite elements [Yserentant '86, Bank '88]

- BPX-type preconditioners [Bramble - Pasciak - Xu '90]
  Set
  $$P_j v = \sum_{k \in \mathcal{K}_j} \frac{(v, \Phi_{jk})_{L^2(\Omega)}}{(1, \Phi_{jk})_{L^2(\Omega)}} \Phi_{jk}$$
  with $\{\Phi_{jk}\}_{k \in \mathcal{K}_j}$ is the nodal (Lagrange) basis related to a finite element mesh of level j

  Setting
  $$Q_j v = P_{j+1} v - P_j v$$
  one has
  $$\|v\|_{H^1(\Omega)} \sim \sum_{j \geq j_0} 2^{2j} \|Q_j v\|^2_{L^2(\Omega)}, \qquad \forall v \in H^1(\Omega).$$

Compression

[Beylkin - Coifman - Rokhlin '91, Dahmen - Prössdorf -Schneider '93, von Petersdorf - Schwab '95]

Many non-zero entries $a_{\mu\lambda}$ of the matrix $\boldsymbol{A}$ are indeed small, if level difference $|\,|\lambda| - |\mu|\,|$ is large, due to the moment condition.

Consider, for instance

$$a_{\mu\lambda} = \int_0^1 \psi'_{jk}(x)\psi'_{\ell h}(x)\,dx.$$

Then

$$|a_{\mu\lambda}| \lesssim 2^{j+\ell}2^{-\rho|j-\ell|}\chi_{\mathsf{supp}\psi_\lambda\cap\mathsf{supp}\psi_\mu}$$

where $\rho = \min(\tilde{L}+1, \sigma-1)$

Indeed, assuming e.g. $j \le \ell$,

$$\begin{aligned} a_{\mu\lambda} &= 2^{j+\ell}2^{(j-\ell)/2}\int_0^1 \psi'(2^{j-\ell}y-m)\psi'(y)\,dy \\ &= 2^{j+\ell}2^{(j-\ell)/2}\int_0^1 [\psi'(2^{j-\ell}y-m)-p(y)]\psi'(y)\,dy \end{aligned}$$

for some m and for any polynomial p of degree $\le \tilde{L}$.

By Whitney's Theorem (recall Jackson inequality) one has

$$\begin{aligned} \|\psi'(2^{j-\ell}\,\cdot\,-m)-p\|_{L^2(0,1)} &\lesssim |\psi'(2^{j-\ell}\,\cdot\,-m)|_{H^\rho(0,1)} \\ &\lesssim 2^{(j-\ell)/2}2^{(\rho-1)(j-\ell)}, \end{aligned}$$

whence the result.

For a large class of elliptic differential/integral/pseudo-differential operators, the following estimate holds

$$2^{-r(|\lambda|+|\mu|)}|\langle A\psi_\mu, \psi_\lambda\rangle| \lesssim \frac{2^{-\beta\,|\,|\lambda|-|\mu|\,|}}{[1+d(\lambda,\mu)]^\gamma}$$

with

$$d(\lambda,\mu) = 2^{\min(|\lambda|,|\mu|)}\text{dist}(\text{supp}\,\psi_\lambda, \text{supp}\,\psi_\mu)\ ,$$

and

$\beta > n/2$ depends on the smoothness of the wavelets,

$\gamma > n$ is related to $\tilde{L}$ and the order of the operator A.

These estimates allow an a-priori thresholding of the matrix elements, so that only 'non-negligible' elements are actually computed.

$\Rightarrow$ Savings in the costruction of the matrix and the matrix-vector multiply.

Remark: Recent results on efficient approximate evaluation of matrix coefficients exploiting tensor-products, refinability, moment conditions, quadratures.

See [Dahmen - Schneider - Xu '98], [Bertoluzza - C. - Urban '99], [Berrone - Urban '99], [Dahmen - Schneider '99]

Adaptivity

Assume that $A : V \to V'$ is an algebraic and topological isomorphism,

$$\|Av\|_{V'} \asymp \|v\|_V, \qquad \forall v \in V .$$

Let $\Psi = \{\psi_\lambda\}_{\lambda\in\mathcal{M}}$ be a Riesz basis in V, and let $\tilde{\Psi} = \{\tilde{\psi}_\lambda\}_{\lambda\in\mathcal{M}}$ be the dual biorthogonal basis in V'.

Exact Problem: *Given $f \in V'$, find $u \in V$ such that*

$$\langle Au, v\rangle = \langle f, v\rangle, \qquad \forall v \in V .$$

Galerkin Problem: Given $\Lambda \subset \mathcal{M}$ and $V_\Lambda = \text{span}\,\{\psi_\lambda \ : \ \lambda \in \Lambda\}$, *find $u_\Lambda \in V_\Lambda$ such that*

$$\langle Au_\Lambda, v_\Lambda\rangle = \langle f, v_\Lambda\rangle, \qquad \forall v_\Lambda \in V_\Lambda .$$

Goal (*fixed tolerance*): Given $\eta > 0$, find in the most efficient way the smallest index set $\Lambda \subset \mathcal{M}$ such that

$$\|u - u_\Lambda\|_V \sim \eta .$$

Goal (*fixed resources*): Given $N > 0$, find in the most efficient way an index set $\Lambda \subset \mathcal{M}$ with card $\Lambda = N$ such that

$$\|u - u_\Lambda\|_V \ \text{ is minimal.}$$

Precisely,

$$\|u - u_\Lambda\|_V \sim \inf_{v_N \in S_N} \|u - v_N\|_V$$

the cost of solving the problem is $\mathcal{O}(N)$.

This is a best N-term approximation problem, but u is not known!

Crucial step of any adaptive algorithm: Given the old set of active indices $\Lambda \subset \mathcal{M}$ and the corresponding discrete solution u_Λ, define a new set $\Lambda^* \subset \mathcal{M}$ such that

$$\|u - u_{\Lambda^*}\|_V < \|u - u_\Lambda\|_V .$$

The crucial step requires a *local error indicator:*

- inspect the wavelet coefficients of u_Λ (solution analysis)
- inspect the wavelet coefficients of the residual $r_\Lambda := f - Au_\Lambda$ (residual analysis).

Solution Analysis

[Liandrat - Tchamitchan '90, Maday - Perrier - Ravel '92, Bertoluzza - Maday - Ravel '94, C. - Cravero '97, Amat - Arandiga - Cohen - Donat '01]

Rationale: the wavelet coefficients of u are indicators of *local smoothness* "$\Rightarrow$" the wavelet coefficients of u are indicators of potential *local error*.

A model strategy to define the new Λ^* is as follows:

Fix $\varepsilon_{\max} > \varepsilon_{\min} > 0$. For any $\lambda \in \Lambda$, look at $(u_\Lambda)_\lambda$:

$$|(u_\Lambda)_\lambda| \begin{cases} < \varepsilon_{\min} & \Rightarrow \ \lambda \notin \Lambda^* \\ \in [\varepsilon_{\min}, \varepsilon_{\max}] & \Rightarrow \ \lambda \in \Lambda^* \\ > \varepsilon_{\max} & \Rightarrow \ \lambda' \in \Lambda^* \ \text{ for all } \lambda' \text{ 'near' } \lambda. \end{cases}$$

Pros

- Very easy and cheap to implement, particularly for evolution problems:

$$u_t + Au = f \qquad \Rightarrow \qquad u^{n+1} = u^n + \Delta t\, \Phi(u^n, \Delta t).$$

- Very effective when it works.

Cons

- Mostly heuristic
- Theory is lacking, or partial
- The method may miss significant structures in the solution.

Residual analysis

[Bertoluzza '95, Dahlke - Dahmen - Hochmuth - Schneider '97, Cohen - Masson '98, Cohen - Dahmen - DeVore '98, '00, Dahlke - Dahmen - Urban '01, Bertoluzza - Verani '01, C.- Urban '02]

Recall

$$\|v\|_V \asymp \|Av\|_{V'}, \qquad \forall v \in V \ .$$

This implies

$$\|u - u_\Lambda\|_V \asymp \|Au - Au_\Lambda\|_{V'} \asymp \|f - Au_\Lambda\|_{V'} \ .$$

The quantity

$$r_\Lambda := f - Au_\Lambda$$

is the residual generated by the approximation $\boldsymbol{u}_\Lambda$ to u.

• **Key point:** Using Riesz bases Ψ in V and $\tilde{\Psi}$ in V', everything can be expressed in terms of sequences in $\ell^2(\mathcal{M})$:

$$u = \boldsymbol{u}^T \Psi = \sum_{\lambda \in \mathcal{M}} v_\lambda \psi_\lambda, \qquad u_\Lambda = \boldsymbol{u}_\Lambda^T \Psi,$$

$$Au = (\boldsymbol{Au})^T \tilde{\Psi}, \qquad f = \boldsymbol{f}^T \tilde{\Psi}$$

with

$$\|u\|_V \asymp \|\boldsymbol{u}\|_{\ell^2(\mathcal{M})}, \qquad \|f\|_{V'} \asymp \|\boldsymbol{f}\|_{\ell^2(\mathcal{M})},$$

$$\boldsymbol{A} : \ell^2(\mathcal{M}) \to \ell^2(\mathcal{M}) \ \text{ isomorphism.}$$

Then,

$$\|\boldsymbol{u} - \boldsymbol{u}_\Lambda\|_{\ell^2(\mathcal{M})} \asymp \|\boldsymbol{f} - \boldsymbol{A}\boldsymbol{u}_\Lambda\|_{\ell^2(\mathcal{M})} \ .$$

From now on, we follow *Cohen - Dahmen - DeVore '00*. For simplicity, assume that A is self-adjoint and coercive. Then

$$\boldsymbol{A} \text{ is symmetric positive} - \text{definite},$$
$$\text{cond}_2(\boldsymbol{A}) \asymp 1.$$

Basic Algorithm. *Fix a target accuracy $\eta_T > 0$. Approximate u by Richardson iterations*

$$\begin{cases} \boldsymbol{u}^0 = \boldsymbol{0} \\ \boldsymbol{u}^{k+1} = \boldsymbol{u}^k + \alpha(\boldsymbol{f} - \boldsymbol{A}\boldsymbol{u}^k), \qquad k \geq 0, \end{cases}$$

until

$$\|\boldsymbol{u} - \boldsymbol{u}^k\|_{\ell^2(\mathcal{M})} \asymp \eta_T \,.$$

Recall: $\exists$ a range of $\alpha > 0$ for which

$$\rho := \|\boldsymbol{I} - \alpha\boldsymbol{A}\|_{\ell^2 \to \ell^2} < 1,$$

so that the algorithm converges. Let us fix one of such α.

Definition. *For any $\boldsymbol{v} \in \ell^2(\mathcal{M})$, set*

$$\Lambda = \text{supp}\,\boldsymbol{v} = \{\lambda \in \mathcal{M} : \; v_\lambda \neq 0\}.$$

The vector $\boldsymbol{v}$ is finite if

$$\text{card}\,\Lambda < +\infty.$$

A finite vector with support Λ will be denoted by $\boldsymbol{v}_\Lambda$.

Modified Algorithm.

$$\begin{cases} \boldsymbol{u}_\Lambda^0 = \boldsymbol{0} \\ \\ \boldsymbol{u}_\Lambda^{k+1} = (\boldsymbol{u}_\Lambda^k + \alpha(\boldsymbol{f}_\Lambda - (\boldsymbol{A}\boldsymbol{u}_\Lambda^k)_\Lambda))_\Lambda, \qquad k \geq 0. \end{cases}$$

Precisely, suppose that we are able to construct finite vectors

$$\boldsymbol{f}_\Lambda := \mathsf{RHS}[\boldsymbol{f}, \eta] \quad \text{such that} \quad \|\boldsymbol{f} - \boldsymbol{f}_\Lambda\|_{\ell^2(\mathcal{M})} \leq \eta\,,$$

$$\boldsymbol{w}_\Lambda := \mathsf{MULT}[\boldsymbol{A}, \boldsymbol{v}_\Lambda, \eta] \quad \text{such that} \quad \|\boldsymbol{A}\boldsymbol{v}_\Lambda - \boldsymbol{w}_\Lambda\|_{\ell^2(\mathcal{M})} \leq \eta\,,$$

$$\bar{\boldsymbol{v}}_\Lambda := \mathsf{COARSE}[\boldsymbol{v}_\Lambda, \eta] \quad \text{such that} \quad \|\boldsymbol{v}_\Lambda - \bar{\boldsymbol{v}}_\Lambda\|_{\ell^2(\mathcal{M})} \leq \eta$$

with minimal memory/complexity.

Adaptive Algorithm. Let K be a fixed integer.

i) Set $\boldsymbol{u}_\Lambda^0 := \boldsymbol{0}$ and $\eta_0 \sim \|\boldsymbol{f}\|_{\ell^2(\mathcal{M})}$

ii) for $\ell \geq 0$ until $\eta_\ell \leq \eta_T$ do:

a) Set $\boldsymbol{v}_\Lambda^0 := \boldsymbol{u}_\Lambda^\ell$

b) for $k = 0, \ldots, K$ do:

compute $\boldsymbol{f}_\Lambda^k = \mathsf{RHS}[\boldsymbol{f}, \rho^k \eta_\ell]$ and $\boldsymbol{w}_\Lambda^k := \mathsf{MULT}[\boldsymbol{A}, \boldsymbol{v}_\Lambda^k, \rho^k \eta_\ell]$

update $\boldsymbol{v}_\Lambda^{k+1} = \boldsymbol{v}_\Lambda^k + \alpha(\boldsymbol{f}_\Lambda^k - \boldsymbol{w}_\Lambda^k)$

c) set $\boldsymbol{u}_\Lambda^{\ell+1} =: \mathsf{COARSE}[\boldsymbol{v}_\Lambda^{K+1}, 2\eta_\ell/5]$ and $\eta_{\ell+1} := \eta_\ell/2$.

Proposition 1. *There exists $K > 0$ such that*

$$\|\boldsymbol{u} - \boldsymbol{v}_\Lambda^{K+1}\|_{\ell^2(\mathcal{M})} \leq \frac{1}{10}\eta_\ell.$$

Corollary 2. *The Adaptive Algorithm converges. Indeed, it is easily seen that*

$$\|\boldsymbol{u} - \boldsymbol{u}_\Lambda^{\ell+1}\|_{\ell^2(\mathcal{M})} \leq \eta_{\ell+1}.$$

Question: Is the Algorithm efficient?

Yes, because RHS, MULT and COARSE can be constructed taking into account the results of Nonlinear Approximation. The conclusion is as follows:

Theorem 3. *Let us assume that the exact solution $u \in V$ is such that $\boldsymbol{u} \in \ell^\tau_w(\mathcal{M})$. Then*

$$\mathrm{card}(\mathrm{supp}\, \boldsymbol{u}_\Lambda^{\ell+1}) \;\lesssim\; \eta_{\ell+1}^{-d/s} \|\boldsymbol{u}\|_{\ell^\tau_w(\mathcal{M})}^{d/s}$$

and

$$\#\, \text{floating point operations to compute } \boldsymbol{u}_\Lambda^{\ell+1} \;\lesssim\; \eta_{\ell+1}^{-d/s} \|\boldsymbol{u}\|_{\ell^\tau_w(\mathcal{M})}^{d/s}.$$

This means that the function $u^{\ell+1}$ whose vector of coefficients is $\boldsymbol{u}_\Lambda^{\ell+1}$ stays asymptotically close to the best N-term approximation of u. Furthermore, the operations needed to compute it is proportional to the number of its coefficients.