# Adaptive wavelet methods

Rob Stevenson
Korteweg-de Vries Institute for Mathematics
Faculty of Science
University of Amsterdam

partly joint work with Tammo Jan Dijkema (Utrecht),
Helmut Harbrecht (Bonn), Gantumur Tsogtgerel (San Diego),
Christoph Schwab (Zürich)

# Contents

- Nonlinear approximation from wavelet bases
- CDD2 and CDD1 adaptive wavelet schemes
- Adaptive tensor product approximation to overcome the curse of dimensionality
- Simultaneous space-time adaptive solution of parabolic problems

# Non-adaptive solution of PDE's

Poisson: $\begin{cases} -\triangle u = f & \text{on } \Omega \subset \mathbb{R}^n \\ u = 0 & \text{on } \partial\Omega \end{cases}$

Var. form.: Find $u \in H_0^1(\Omega)$ s.t.

$$a(u,v) := \int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv =: f(v) \quad (v \in H_0^1(\Omega)).$$

$|||\cdot||| := a(\cdot,\cdot)^{\frac{1}{2}} \eqsim \|\cdot\|_{H^1}$ on $H_0^1(\Omega)$ (ellipticity).

Galerkin: $V \subset H_0^1(\Omega)$, find $u_V \in V$ s.t.

$$a(u_V, v_V) = f(v_V) \quad (v_V \in V).$$

Cea's lemma: $|||u - u_V||| = \inf_{v_V \in V} |||u - v_V|||$

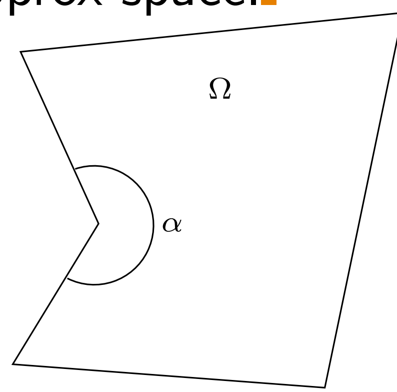
(Non-adaptive) FEM: $\tau_0 \subset \tau_1 \subset \cdots$ seq of subdivisions of $\Omega$ into $n$-simplices (say) based on uniform refinements (red-refinement),

$$V_j := H_0^1(\Omega) \cap \prod_{T \in \tau_j} P_{d-1}(T).$$

$$\frac{\inf_{v_j \in V_j} \|u - v_j\|_{H^1}}{|u|_{H^d}} \lesssim \text{“}h_j\text{”}^{d-1} \eqsim N_j^{-\frac{d-m}{n}}$$

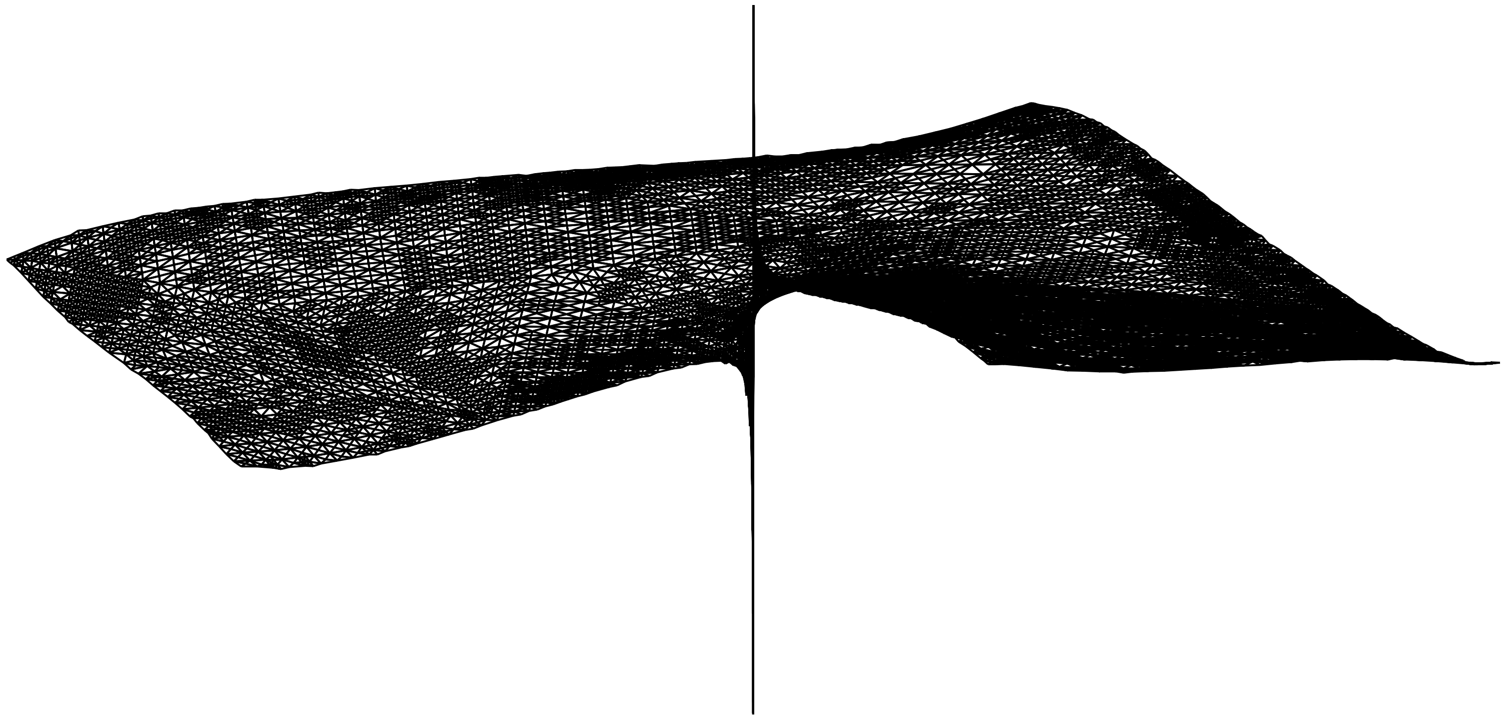
$d$ order of approx, $2m$ order of eq, $n$ space dim, $N_j$ dim approx space.

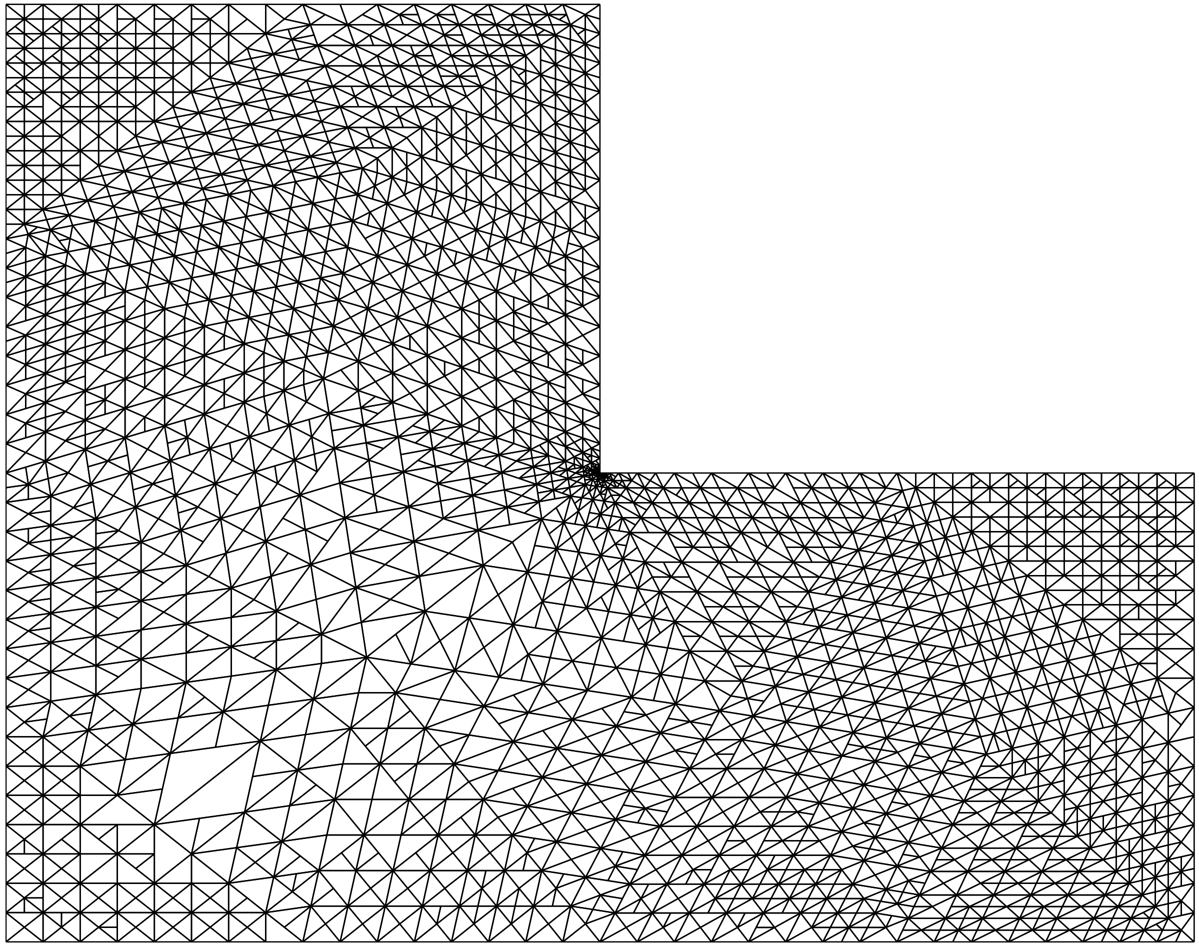
**Regularity theory:**
Poisson ($m = 1$), $\alpha$ max. int. angle of $\Omega \subset \mathbb{R}^2$ ($n = 2$). Then for smooth $f$, $u \in H^s(\Omega)$ if, and generally, only if, $s < 1 + \frac{\pi}{\alpha} \in (\frac{3}{2}, 4]$ ($1 + \frac{\pi}{\alpha} < 2$ for re-entrant corners).




Rate is $N^{-\frac{\min(d,s)-m}{n}}$.

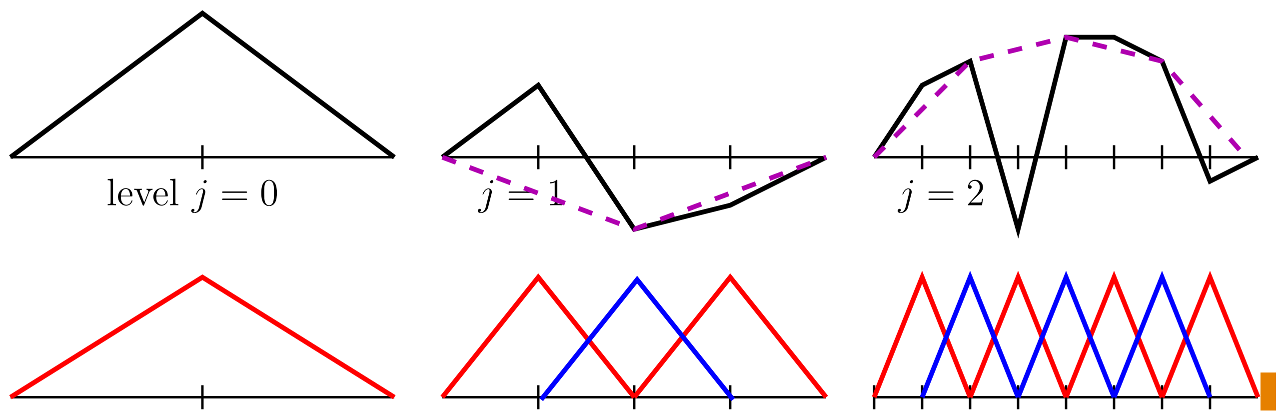
To recover best possible rates: adaptive finite element or adaptive wavelet methods.





# Multilevel bases

Consider seq. of cont. piecewise linears w.r.t. dyadically refined partitions.




$\cup_{j=0}^{\infty} \cup_i \{\psi_{j,i}\}$ Riesz basis for $H_0^1(\Omega)$ (even orthogonal), called hierarchical basis (later $(j,i) \to \lambda$ with $|\lambda| = j$)

Local refinement $\sim$ add basis functions only locally.

For finding seq over $N$ of (quasi-) best locally refined partitions into $N$ subintervals, sufficient to find seq over $N$ of (quasi-) best subsets with card. $N$ of infinite basis (actually latter is more general).

# Wavelets

Let $V_0 \subset V_1 \subset \cdots \subset L_2$, $\tilde{V}_0 \subset \tilde{V}_1 \subset \cdots \subset L_2(\Omega)$ s.t. for $d, \gamma, \tilde{d}, \tilde{\gamma} > 0$,

$$\inf_{v_j \in V_j} \|u - v_j\|_{L_2} \lesssim 2^{-jd}|u|_{H^d}, \qquad \|v_j\|_{H^s} \lesssim 2^{js}\|v_j\|_{L_2}\,(s \in [0, \gamma)),$$

$$\inf_{\tilde{v}_j \in \tilde{V}_j} \|u - \tilde{v}_j\|_{L_2} \lesssim 2^{-j\tilde{d}}|u|_{H^{\tilde{d}}}, \qquad \|\tilde{v}_j\|_{H^s} \lesssim 2^{js}\|\tilde{v}_j\|_{L_2}\,(s \in [0, \tilde{\gamma})),$$

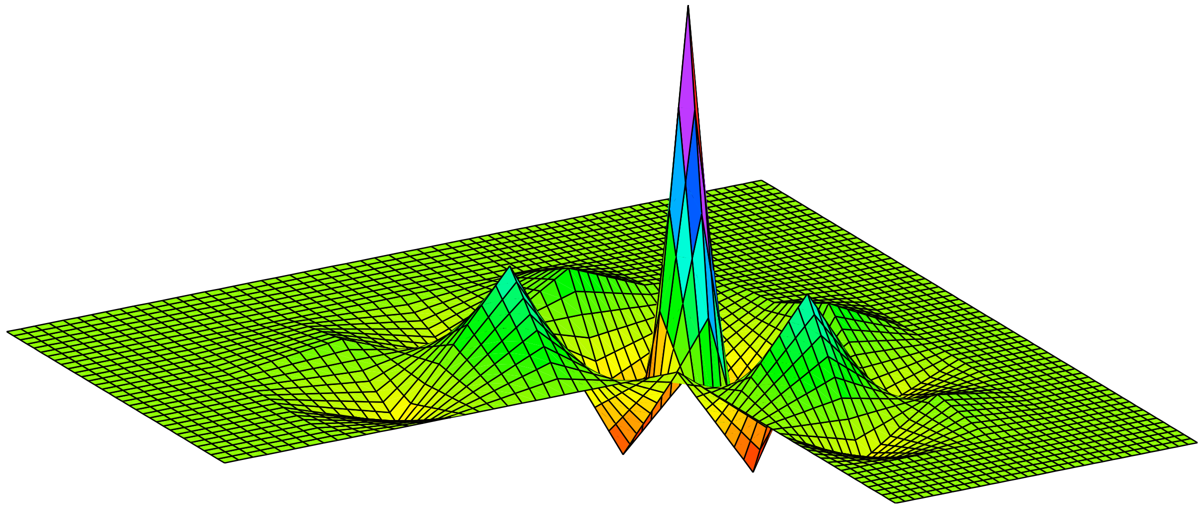
Let $\Phi_j$, $\tilde{\Phi}_j$ unif. $L_2$-Riesz b. for $V_j$, $\tilde{V}_j$ with $\langle \Phi_j, \tilde{\Phi}_j \rangle_{L_2} = I$ (biorth scal.)

Let $\Psi_j = \{\psi_\lambda : |\lambda| = j\}$ be unif. $L_2$-Riesz bases for $V_{j+1} \cap \tilde{V}_j^{\perp_{L_2}}$ ($\tilde{V}_{-1} := 0$). Then for $m \in (-\min\{\tilde{d}, \tilde{\gamma}\}, \min\{d, \gamma\})$, $\{2^{-m|\lambda|}\psi_\lambda : \lambda \in \nabla\}$ is Riesz basis for $H^m(\Omega)$, i.e., $u = \mathbf{u}^\top \Psi := \sum_{\lambda \in \nabla} \mathbf{u}_\lambda \psi_\lambda$ with ($\ell_2 = \ell_2(\nabla)$)

$$\boxed{\|u\|_{H^m} \eqsim \|(2^{|\lambda|m}\mathbf{u}_\lambda)_{\lambda \in \nabla}\|_{\ell_2}.}$$

If $\Xi_j \subset V_{j+1}$ (initial stable completion) s.t. $\Phi_j \cup \Xi_j$ unif. $L_2$-Riesz basis for $V_{j+1}$, then $\Psi_j$ can be constructed as

$$\boxed{\Psi_j := \Xi_j - \langle \Xi_j, \tilde{\Phi}_j \rangle_{L_2} \Phi_j.}$$



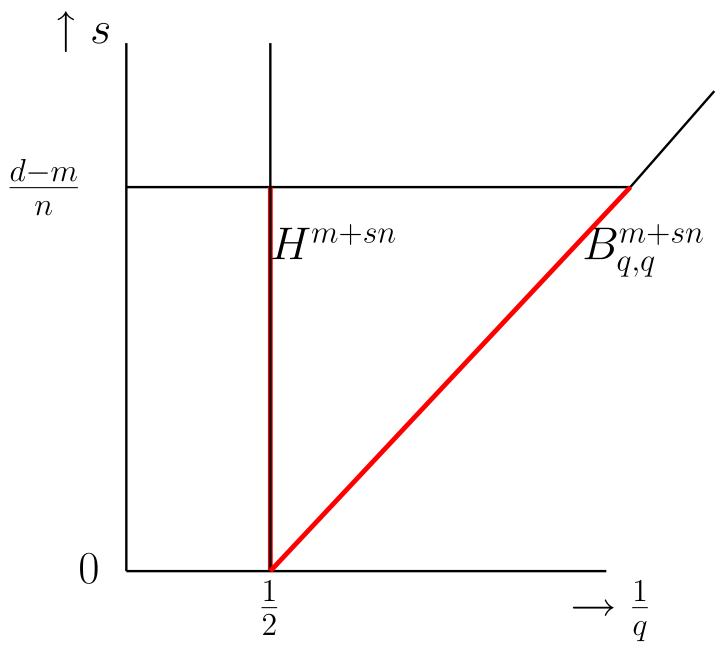
More generally, for a certain range of $s$, $p$ and $q$,

$$\|u\|_{B^s_{p,q}} \eqsim \|(2^{s+n(\frac{1}{2}-\frac{1}{p})j}\|(\mathbf{u}_\lambda)_{|\lambda|=j}\|_{\ell_p})_j\|_{\ell_q}.$$

in part, for $0 \leq s < \frac{d-m}{n}$,

$$\boxed{\|u\|_{B^{m+sn}_{q,q}} \eqsim \|(2^{|\lambda|m}\mathbf{u}_\lambda)_{\lambda\in\nabla}\|_{\ell_q}}, \text{ where } q = (\tfrac{1}{2}+s)^{-1}$$

([Dahmen '97], [Cohen '00])



# Approximation classes

Let $\mathcal{X}$ be Hilbert, $\Psi$ a Riesz basis, i.e., a basis and $u = \mathbf{u}^\top \Psi$ with $\|u\|_{\mathcal{X}} \eqsim \|\mathbf{u}\|_{\ell_2}$.

Let $\Sigma_N \subset \mathcal{X}$ set of functions that are linear combinations of $N$ basis functions.

$$\sigma_N(u) := \inf_{v \in \Sigma_N} \|u - v\|_{\mathcal{X}} \eqsim \inf_{\{\mathbf{v} \in \ell_2 : \# \operatorname{supp} \mathbf{v} \leq N\}} \|\mathbf{u} - \mathbf{v}\|_{\ell_2} =: \sigma_N(\mathbf{u}).$$

$\mathcal{A}_q^s(\mathcal{X}) = \mathcal{A}_q^s(\mathcal{X}, \Psi)$ set of all $u \in \mathcal{X}$ with

$$|u|_{\mathcal{A}_q^s} := \left\{ \begin{array}{ll} \left(\sum_N [N^s \sigma_N(u)]^q \frac{1}{N}\right)^{1/q} & 0 < q < \infty \\ \sup_N N^s \sigma_N(u) & q = \infty \end{array} \right\} < \infty.$$

$\mathcal{A}_q^s(\ell_2)$ set of all $\mathbf{u} \in \ell_2$ with ....

$$u \in \mathcal{A}_q^s(\mathcal{X}) \Longleftrightarrow \mathbf{u} \in \mathcal{A}_q^s(\ell_2)$$

with equivalent (quasi-)norms.

# A characterization of $\mathcal{A}^s_q$

$$\sigma_N(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_N\|_{\ell_2} = \sqrt{\sum_{k>N} |\mathbf{u}_{\lambda_k}|^2},$$

where $\mathbf{u}_N$ is a best $N$-term approximation for $\mathbf{u}$, and $(\mathbf{u}_{\lambda_k})_k$ is non-increasing rearrangement of $(\mathbf{u}_\lambda)_{\lambda\in\nabla}$.

For some $p \in (0,2)$, let $\mathbf{u} \in \ell_p$. Then

$$k|\mathbf{u}_{\lambda_k}|^p \leq \sum_{\ell=1}^{k} |\mathbf{u}_{\lambda_\ell}|^p \leq \|\mathbf{u}\|_{\ell_p}^p \quad \rightsquigarrow \quad \sigma_N(\mathbf{u}) \leq N^{-(\frac{1}{p}-\frac{1}{2})}\|\mathbf{u}\|_{\ell_p}.$$

For $\mathbf{v} \in \ell_2$ with $\operatorname{supp} \mathbf{v} \leq N$, $\|\mathbf{v}\|_{\ell_p} \leq N^{\frac{1}{p}-\frac{1}{2}}\|\mathbf{v}\|_{\ell_2}$.

J & B $\rightsquigarrow$ (DeVore '98]) For $s \in (0, \frac{1}{p} - \frac{1}{2})$ and $q \in (0, \infty]$

$$\mathcal{A}^s_q = (\ell_2, \ell_p)_{\frac{s}{\frac{1}{p}-\frac{1}{2}},q} = \ell_{\tau,q} \text{ where } \boxed{\tau = (\tfrac{1}{2} + s)^{-1}}$$

True for any $s \in (0, \infty)$ since $p < 2$ was arbitrary.

Special cases:

$$\mathcal{A}_\infty^s = \ell_{\tau,\infty} = \{\mathbf{u} \in \ell_2 : |\mathbf{u}|_{\ell_{\tau,\infty}} := \sup_k k^{1/\tau} |\mathbf{u}_{\lambda_k}| < \infty\}$$

$$\mathcal{A}_\tau^s = \ell_{\tau,\tau} = \ell_\tau \overset{0 \leq s < \frac{d-m}{n}}{=} B_{\tau,\tau}^{m+sn}(\Omega)$$

Last equality when $\mathcal{X} = H^m(\Omega)$, and $\Psi$ is $H^m(\Omega)$-normalized wavelet basis of order $d$.

$B_{\tau,\tau}^{m+sn}(\Omega)$ much larger than $H^{m+sn}(\Omega)$, membership of which is needed to get same rate with standard linear approximation.

**Regul. th.:** Poisson ($m = 1$), $n = 2$, polygonal domain; for suff. smooth $f$, sol $u \in B_{\tau,\tau}^{m+sn}(\Omega)$ for any $s$ ([Dahlke, DeVore '97]).

# Summary

Best $N$-term approximation from wavelet bases converges with rate $s$ under much milder regularity conditions than standard linear approximation from $\mathrm{span}\{\psi_\lambda : |\lambda| \leq \ell\}$.

## References


[1] A. Cohen. Wavelet methods in numerical analysis. In P.G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis. Vol. VII.*, pages 417–711. North-Holland, Amsterdam, 2000.

[2] S. Dahlke and R. DeVore. Besov regularity for elliptic boundary value problems. *Comm. Partial Differential Equations*, 22(1 & 2):1–16, 1997.

[3] W. Dahmen. Stability of multiscale transformations. *J. Fourier Anal. Appl.*, 4:341–362, 1996.

[4] R. DeVore. Nonlinear approximation. *Acta Numer.*, 7:51–150, 1998.

# Well posed lin. op eqs. and their formulation as well-posed bi-infinite MV eqs

Let $\mathcal{X}$, $\mathcal{Y}$ be sep. Hilbert spaces. Let $B \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$. Given $f \in \mathcal{Y}'$, we seek $u \in \mathcal{X}$ s.t.

$$Bu = f.$$

Ex.:

- $(Bw)(v) = \int_\Omega \nabla w \cdot \nabla v$, $\mathcal{X} = \mathcal{Y} = H_0^1(\Omega)$ (Poisson problem),
- $(B(\vec{w}, p))(\vec{v}, q) = \int_\Omega \nabla \vec{w} : \nabla \vec{v} - \int_\Omega p \operatorname{div} \vec{v} - \int_\Omega q \operatorname{div} \vec{w}$, $\mathcal{X} = \mathcal{Y} = H_0^1(\Omega)^n \times L_{2,0}(\Omega)$ for a domain $\Omega \subset \mathbb{R}^n$ (Stokes problem),
- $(Bw)(v) = \frac{1}{4\pi} \int_{\partial\Omega} \int_{\partial\Omega} \frac{(w(y) - w(x))(v(y) - v(x))}{|x-y|^3} dx dy$, $\Omega \subset \mathbb{R}^3$, $\mathcal{X} = \mathcal{Y} = H^{\frac{1}{2}}(\partial\Omega)/\mathbb{R}$ (hypersingular boundary integral equation).

Later ex. with $\mathcal{X} \neq \mathcal{Y}$.

Let $\Psi^{\mathcal{X}} = \{\psi_\lambda^{\mathcal{X}} : \lambda \in \nabla_{\mathcal{X}}\}$, $\Psi^{\mathcal{Y}} = \{\psi_\lambda^{\mathcal{Y}} : \lambda \in \nabla_{\mathcal{Y}}\}$ Riesz bases for $\mathcal{X}$, $\mathcal{Y}$. I.e., synthesis operator and its adjoint the analysis operator

$$s_{\Psi^{\mathcal{X}}} : \ell_2(\nabla_{\mathcal{X}}) \to \mathcal{X} : \mathbf{c} \mapsto \mathbf{c}^\top \Psi^{\mathcal{X}} := \sum_{\lambda \in \nabla_{\mathcal{X}}} c_\lambda \psi_\lambda^{\mathcal{X}}$$

$$s'_{\Psi^{\mathcal{X}}} : \mathcal{X}' \to \ell_2(\nabla_{\mathcal{X}}) : g \mapsto [g(\psi_\lambda^{\mathcal{X}})]_{\lambda \in \nabla_{\mathcal{X}}}.$$

boundedly invertible (anal. for $s_{\Psi^{\mathcal{Y}}}$).

$$Bu = f \Longleftrightarrow \underbrace{s'_{\Psi^{\mathcal{Y}}} B s_{\Psi^{\mathcal{X}}}}_{\mathbf{B}} \underbrace{s^{-1}_{\Psi^{\mathcal{X}}} u}_{\mathbf{u}} = \underbrace{s'_{\Psi^{\mathcal{Y}}} f}_{\mathbf{f}},$$

where

$$\mathbf{B} = [(B\psi_\mu^{\mathcal{X}})(\psi_\lambda^{\mathcal{Y}})]_{\lambda \in \nabla_{\mathcal{Y}}, \mu \in \nabla_{\mathcal{X}}} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{X}}), \ell_2(\nabla_{\mathcal{Y}}))$$

is boundedly invertible (usually $(Bu)(v) = b(u,v)$ so that $\mathbf{B}$ is infinite "stiffness" matrix),

$$\mathbf{f} = s'_{\Psi^{\mathcal{Y}}} f = [f(\psi_\lambda^{\mathcal{Y}})]_{\lambda \in \nabla_{\mathcal{Y}}} \in \ell_2(\nabla_{\mathcal{Y}})$$

(infinite "load" vector).

# Adaptive wavelet schemes

To solve $\mathbf{B}\mathbf{u} = \mathbf{f}$. $\quad \|\cdot\| = \|\cdot\|_{\ell_2}$ or $\|\cdot\|_{\ell_2 \to \ell_2}$.

Aim: Given $\varepsilon > 0$, find $\mathbf{u}_\varepsilon$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell_2} \leq \varepsilon$, with whenever $\mathbf{u} \in \mathcal{A}^s_\infty$ for some $s > 0$, both $\#\mathrm{supp}\,\mathbf{u}_\varepsilon$ and work $\lesssim \varepsilon^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s_\infty}$.

($\|\mathbf{u} - \mathbf{u}_N\| \leq N^{-s}|\mathbf{u}|_{\mathcal{A}^s_\infty}$ (gen. sharp), $N^{-s}|\mathbf{u}|_{\mathcal{A}^s_\infty} = \varepsilon \rightsquigarrow N = \varepsilon^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s_\infty}$)

For time being: $\mathbf{B} = \mathbf{B}^* > 0$.

## CDD2 scheme [Cohen, Dahmen, DeVore '02]:

Plan: Apply simple iter. scheme like damped Rich: Let $0 < \alpha < \|\mathbf{B}\|$,

$$\mathbf{u}^{(i+1)} := \mathbf{u}^{(i)} - \alpha(\mathbf{B}\mathbf{u}^{(i)} - \mathbf{f}).$$

$\rho := \|I - \alpha\mathbf{B}\| < 1$.

# Inexact iterations

$\mathbf{RHS_f}[\varepsilon] \to \mathbf{f}_\varepsilon$: Determines $\mathbf{f}_\varepsilon \in \ell_0$ with $\|\mathbf{f} - \mathbf{f}_\varepsilon\| \leq \varepsilon$.

$\mathbf{APPLY_B}[\mathbf{v}, \varepsilon] \to \mathbf{w}_\varepsilon$: For $\mathbf{v} \in \ell_0$, determ. $\mathbf{w}_\varepsilon \in \ell_0$ with $\|\mathbf{Bv} - \mathbf{w}_\varepsilon\| \leq \varepsilon$.

$\mathbf{SOLVE}[\varepsilon] \to \mathbf{u}_\varepsilon$:
$\mathbf{u}^{(0)} := 0$, $\varepsilon_0 := \|\mathbf{B}^{-1}\|\|\mathbf{f}\|$
$K$ smallest integer with $2\rho^K \leq \varepsilon/\varepsilon_0$
`for` $i := 1$ `to` $K$ `do`

$$\mathbf{u}^{(i+1)} := \mathbf{u}^{(i)} - \alpha(\mathbf{APPLY_B}[\mathbf{u}^{(i)}, \tfrac{\rho^i}{2\alpha K}\varepsilon_0] - \mathbf{RHS_f}[\tfrac{\rho^i}{2\alpha K}\varepsilon_0])$$

enddo
$\mathbf{u}_\varepsilon := \mathbf{u}^{(K)}$

**Th 1.** $\|\mathbf{u} - \mathbf{u}_\varepsilon\| \leq \varepsilon$.

*Proof.* $\mathbf{u} - \mathbf{u}^{(K)} = (\mathbf{I} - \alpha\mathbf{B})^K(\mathbf{u} - \mathbf{u}^{(0)}) + \sum_{i=1}^K (\mathbf{I} - \alpha\mathbf{B})^{K-i}\delta_i$ with $\|\delta_i\| \leq 2\alpha\frac{\rho^i}{2\alpha K}\varepsilon_0$. $\|\mathbf{u} - \mathbf{u}^{(K)}\| \leq \rho^K\varepsilon_0 + \sum_{i=1}^K \frac{\rho^K}{K}\varepsilon_0 = 2\rho^K\varepsilon_0 \leq \varepsilon$. □

# Cost

**Def 1.** $\mathbf{B}$ is **$s^*$-admissible** when for $s \in (0, s^*)$, $\mathbf{w}_\varepsilon := \mathbf{APPLY}_\mathbf{B}[\mathbf{v}, \varepsilon]$ satisfies $\#\operatorname{supp} \mathbf{w}_\varepsilon \lesssim \varepsilon^{-1/s} |\mathbf{v}|_{\mathcal{A}^s_\infty}^{1/s}$ and work $\lesssim \varepsilon^{-1/s} |\mathbf{v}|_{\mathcal{A}^s_\infty}^{1/s} + \#\operatorname{supp} \mathbf{v}$.

**Prop 1.** *If* $\mathbf{B}$ *is* $s^*$*-admis., then for* $s \in (0, s^*)$, $|\mathbf{w}_\varepsilon|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{v}|_{\mathcal{A}^s_\infty}$ *uniform in* $\varepsilon$, *and* $\mathbf{B} : \mathcal{A}^s_\infty \to \mathcal{A}^s_\infty$ *bounded.*

*Proof.* (second part). Let $C$ s.t. $\#\operatorname{supp} \mathbf{w}_\varepsilon \le C\varepsilon^{-1/s} |\mathbf{v}|_{\mathcal{A}^s_\infty}^{1/s}$. Let $\mathbf{v} \in \mathcal{A}^s_\infty$ and $N$ be given. For $\eta := C^s N^{-s} |\mathbf{v}_N|_{\mathcal{A}^s_\infty}$, let $\mathbf{w}_\eta := \mathbf{APPLY}_\mathbf{B}[\mathbf{v}_N, \eta]$, so that $\operatorname{supp} \mathbf{w}_\eta \le N$.

$$
\begin{aligned}
\|\mathbf{Bv} - (\mathbf{Bv})_N\| \le \|\mathbf{Bv} - \mathbf{w}_\eta\| &\le \|\mathbf{Bv}_N - \mathbf{w}_\eta\| + \|\mathbf{B}\| \|\mathbf{v} - \mathbf{v}_N\| \\
&\le C^s N^{-s} |\mathbf{v}_N|_{\mathcal{A}^s_\infty} + \|\mathbf{B}\| N^{-s} |\mathbf{v}|_{\mathcal{A}^s_\infty} \lesssim N^{-s} |\mathbf{v}|_{\mathcal{A}^s_\infty}. \quad \square
\end{aligned}
$$

# Cost

**Def 1.** $\mathbf{B}$ is $s^*$**-admissible** when for $s \in (0, s^*)$, $\mathbf{w}_\varepsilon := \mathbf{APPLY_B}[\mathbf{v}, \varepsilon]$ satisfies $\#\mathrm{supp}\,\mathbf{w}_\varepsilon \lesssim \varepsilon^{-1/s}|\mathbf{v}|^{1/s}_{\mathcal{A}^s_\infty}$ and work $\lesssim \varepsilon^{-1/s}|\mathbf{v}|^{1/s}_{\mathcal{A}^s_\infty} + \#\mathrm{supp}\,\mathbf{v}$.

**Prop 1.** *If* $\mathbf{B}$ *is* $s^*$*-admis., then for* $s \in (0, s^*)$, $|\mathbf{w}_\varepsilon|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{v}|_{\mathcal{A}^s_\infty}$ *uniform in* $\varepsilon$, *and* $\mathbf{B} : \mathcal{A}^s_\infty \to \mathcal{A}^s_\infty$ *bounded.*

**Assump 1.** $\mathbf{f}_\varepsilon := \mathbf{RHS_f}[\varepsilon]$ satisfies $\#\mathrm{supp}\,\mathbf{f}_\varepsilon \lesssim N$ (and work $\lesssim N$), for smallest $N$ with $\|\mathbf{f} - \mathbf{f}_N\| \leq \varepsilon$.

**Corol 1.** *For some* $s \in (0, s^*)$, *let* $\mathbf{u} \in \mathcal{A}^s_\infty$. *Then* $|\mathbf{f}_\varepsilon|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{f}|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{u}|_{\mathcal{A}^s_\infty}$ *and* $\mathrm{supp}\,\mathbf{f}_\varepsilon \lesssim \varepsilon^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s_\infty}$.

Application to **SOLVE**: For some $s \in (0, s^*)$, let $\mathbf{u} \in \mathcal{A}^s_\infty$.

$|\mathbf{u}^{(i+1)}|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{u}^{(i)}|_{\mathcal{A}^s_\infty} + |\mathbf{u}|_{\mathcal{A}^s_\infty}$, so by $\mathbf{u}^{(0)} = 0$, for **fixed** $i$, $|\mathbf{u}^{(i)}|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{u}|_{\mathcal{A}^s_\infty}$.

If this would hold **unif. in** $i$, then $\#\mathrm{supp}\,\mathbf{u}^{(i)} \lesssim \#\mathrm{supp}\,\mathbf{u}^{(i-1)} + \delta_i^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s_\infty}$ and work in $i$th iteration bounded by same expression.
Using linear decrease tolerances, then $\#\mathrm{supp}\,\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s_\infty}$ and total work $\lesssim \varepsilon^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s_\infty}$.

# Coarsening

$\mathbf{COARSE}[\mathbf{w}, \varepsilon] \to \mathbf{w}_\varepsilon$: Determines for $\mathbf{w} \in \ell_0$, the shortest (up to some absolute factor) $\mathbf{w}_\varepsilon \in \ell_0$ with $\|\mathbf{w} - \mathbf{w}_\varepsilon\| \leq \varepsilon$.

**Prop 2.** *Let $\mu > 1$ and $s > 0$. Then for any $\varepsilon > 0$, $\mathbf{u} \in \mathcal{A}^s_\infty$ and $\mathbf{w} \in \ell_0$ with $\|\mathbf{u} - \mathbf{w}\| \leq \varepsilon$, for $\mathbf{v} := \mathbf{COARSE}[\mathbf{w}, \mu\varepsilon]$, $\|\mathbf{u} - \mathbf{v}\| \leq (1 + \mu)\varepsilon$, $\#\operatorname{supp} \mathbf{v} \lesssim \varepsilon^{-1/s} |\mathbf{u}|_{\mathcal{A}^s_\infty}$ and $|\mathbf{v}|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{u}|_{\mathcal{A}^s_\infty}$.*

*Proof.* [Cohen '00] Let $N$ smallest with $\|\mathbf{u} - \mathbf{u}_N\| \leq (\mu - 1)\varepsilon$. Then $N \lesssim \varepsilon^{-1/s} |\mathbf{u}|^{1/s}_{\mathcal{A}^s_\infty}$. Furthermore $\|\mathbf{w} - \mathbf{u}_N\| \leq \|\mathbf{w} - \mathbf{u}\| + \|\mathbf{u} - \mathbf{u}_N\| \leq \mu\varepsilon$. So $\#\operatorname{supp} \mathbf{v} \lesssim N$. Last statement from foll. lemma: □

**Lem 1.** *For $\mathbf{u} \in \mathcal{A}^s_\infty$ and $\mathbf{w} \in \ell_0$, $|\mathbf{w}|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{u}|_{\mathcal{A}^s_\infty} + (\operatorname{supp} \mathbf{w})^s \|\mathbf{u} - \mathbf{w}\|$.*

*Proof.* With $N = \#\operatorname{supp} \mathbf{w}$,

$$|\mathbf{w}|_{\mathcal{A}^s_\infty} \lesssim |\mathbf{w} - \mathbf{u}_N|_{\mathcal{A}^s_\infty} + |\mathbf{u}_N|_{\mathcal{A}^s_\infty} \lesssim (2N)^s \|\mathbf{w} - \mathbf{u}_N\| + |\mathbf{u}|_{\mathcal{A}^s_\infty},$$

where we used $\#\operatorname{supp}(\mathbf{w} - \mathbf{u}_N) \leq 2N$. Now use

$$\|\mathbf{w} - \mathbf{u}_N\| \leq \|\mathbf{w} - \mathbf{u}\| + \|\mathbf{u} - \mathbf{u}_N\| \leq 2\|\mathbf{w} - \mathbf{u}\|. \quad \square$$

# CDD2 algorithm

$\mathbf{SOLVE}[\varepsilon] \to \mathbf{u}_\varepsilon$:
$\mathbf{u}^{(0)} := 0$, $\varepsilon_0 := \|\mathbf{B}^{-1}\|\|\mathbf{f}\|$
Let $\mu > 1$ and $K \in I\!N$ with $(\mu+1)2\rho^K < 1$, $L$ sm. with $[(\mu+1)2\rho^K]^L \leq \varepsilon/\varepsilon_0$
for $j := 1$ to $L$ do
  $\mathbf{v}^{(0)} := \mathbf{u}^{(j-1)}$
  for $i := 1$ to $K$ do
    $\mathbf{v}^{(i+1)} := \mathbf{v}^{(i)} - \alpha(\mathbf{APPLY_B}[\mathbf{v}^{(i)}, \frac{\rho^i}{2\alpha K}\varepsilon_j] - \mathbf{RHS_f}[\frac{\rho^i}{2\alpha K}\varepsilon_j])$
  enddo
  $\mathbf{u}^{(j)} := \mathbf{COARSE}[\mathbf{v}^{(K)}, \mu 2\rho^K \varepsilon_j]$
  $\varepsilon_{j+1} := (\mu+1)2\rho^K\varepsilon_j$
enddo
$\mathbf{u}_\varepsilon := \mathbf{u}^{(L)}$

**Th 2.** $\|\mathbf{u} - \mathbf{u}_\varepsilon\| \leq \varepsilon$. *If, for some* $s \in (0, s^*)$, $\mathbf{u} \in \mathcal{A}^s_\infty$, *then* $\#\mathrm{supp}\,\mathbf{u}_\varepsilon$ *and work* $\lesssim \varepsilon^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s}$.

## Verification of $s^*$-admissibility

**Def 2.** $\mathbf{B}$ is $s^*$-**computable**, when for any $N \in I\!N$, $\exists \mathbf{B}_N$ having in each column $\mathcal{O}(N)$ non-zeros, whose joint computation takes $\mathcal{O}(N)$ operations, s.t. for $s < s^*$, $\|\mathbf{B} - \mathbf{B}_N\| \lesssim N^{-s}$.

If $\mathbf{B}$ is $s^*$-computable, then for any $j$, $\exists \mathbf{B}^{(j)}$ having in each column $\mathcal{O}(\alpha_j 2^j)$ non-zeros, whose joint computation takes $\mathcal{O}(\alpha_j 2^j)$ operations, where $\sum_j \alpha_j < \infty$ s.t. for $s < s^*$, $\sum_j 2^{js}\|\mathbf{B} - \mathbf{B}^{(j)}\| < \infty$.

E.g. take $\alpha_j = (j+1)^{-2}$, $\mathbf{B}^{(j)} = \mathbf{B}_{\alpha_j 2^j}$, and use that for $s < \bar{s} < s^*$, $\sum_j 2^{js}(\alpha_j 2^j)^{-\bar{s}} < \infty$.

$\mathbf{APPLY_B}[\mathbf{v}, \varepsilon] \to \mathbf{w}_\varepsilon$:

Determine sm. $\ell$ with $\|\mathbf{B}\|\|\mathbf{v} - \mathbf{v}_{2^\ell}\| \leq \varepsilon/2$.

With $\mathbf{v}_{[k]} := \mathbf{v}_{2^k} - \mathbf{v}_{2^{k-1}}$, $\mathbf{v}_{[0]} := \mathbf{v}_{2^0}$,

determine sm. $j \geq \ell$ with $\displaystyle\sum_{k=0}^{\ell} \|\mathbf{B} - \mathbf{B}^{(j-k)}\|\|\mathbf{v}_{[k]}\| \leq \varepsilon/2$

$\mathbf{w}_\varepsilon := \mathbf{B}^{(j)}\mathbf{v}_{[0]} + \mathbf{B}^{(j-1)}\mathbf{v}_{[1]} + \cdots + \mathbf{B}^{(j-\ell)}\mathbf{v}_{[\ell]}$.

**Prop 3.** $\|\mathbf{B}\mathbf{v}-\mathbf{w}_\varepsilon\| \leq \varepsilon$, *and for* $s<s^*$, $\#\operatorname{supp}\mathbf{w}_\varepsilon \lesssim \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s_\infty}^{1/s}$ *and, apart from cost of sorting, work* $\lesssim \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s_\infty}^{1/s}+\#\operatorname{supp}\mathbf{v}$.

*Proof.* $\|\mathbf{B}\mathbf{v}-\mathbf{w}_\varepsilon\| \leq \|\mathbf{B}\|\|\mathbf{v}-\mathbf{v}_{2^\ell}\| + \sum_{k=0}^{\ell}\|\mathbf{B}-\mathbf{B}^{(j-k)}\|\|\mathbf{v}_{[k]}\| \leq \varepsilon.$

$\#\operatorname{supp}\mathbf{w}_\varepsilon$ and, apart from sorting, work $\lesssim \sum_{k=0}^{\ell}\alpha_{j-k}2^{j-k}2^k \lesssim 2^j$.

By def. of $j$,

$$
\begin{aligned}
\varepsilon/2 &< \sum_{k=0}^{\ell}\|\mathbf{B}-\mathbf{B}^{(j-1-k)}\|\|\mathbf{v}_{[k]}\| \\
&= 2^{(1-j)s}\sum_{k=0}^{\ell}2^{(j-1-k)s}\|\mathbf{B}-\mathbf{B}^{(j-1-k)}\|\;2^{ks}\|\mathbf{v}_{[k]}\| \lesssim 2^{(1-j)s}|\mathbf{v}|_{\mathcal{A}^s_\infty}
\end{aligned}
$$

or $2^j \lesssim \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s_\infty}^{1/s}$. $\square$

[Barinka '05],[Metselaar '02]: Log-factors due to sorting can be avoided. After normalization s.t. $\|\mathbf{v}\|_{\ell_\infty} = 1$, distribute its entries over buckets

$$(\tfrac{1}{2}, 1], (\tfrac{1}{4}, \tfrac{1}{2}], (\tfrac{1}{8}, \tfrac{1}{4}], \ldots$$

order $\mathbf{v}$ s.t. elements from bucket $k-1$ precede those from bucket $k$; generate $\mathbf{v}_{[0]}, \mathbf{v}_{[1]}, \ldots$ using this ordering. Number of buckets can be bounded.

**Th 3.** *An $s^*$-computable $\mathbf{B}$ is $s^*$-admissible.*

# A more efficient APPLY

• Instead of approximating $\mathbf{B}$ with the same accuracy on the $\mathbf{v}_{[k]}$'s (# is $2^k - 2^{k-1}$), use subdivision into buckets (values in range $[\|\mathbf{v}\|_{\ell_\infty} 2^{-k-1}, \|\mathbf{v}\|_{\ell_\infty} 2^{-k}]$). Saves work and is more natural.

• Instead of using an a priori distribution of the accuracies over the "chunks" compute

$$\mathbf{w}_\varepsilon := \sum_{k=0}^{\ell} \mathbf{B}_{2^{p_k}} \mathbf{v}_{[k]} \text{ where}$$

$$(p_0, \ldots, p_\ell) := \operatorname{argmin}\{\sum_{k=0}^{\ell} 2^{p_k} \# \operatorname{supp} \mathbf{v}_{[k]} : \sum_{k=0}^{\ell} \|\mathbf{B} - \mathbf{B}_{2^{p_k}}\| \|\mathbf{v}_{[k]}\| \leq \varepsilon/2\}$$

Never worse, but often much more efficient.

# Verification of $s^*$-computability

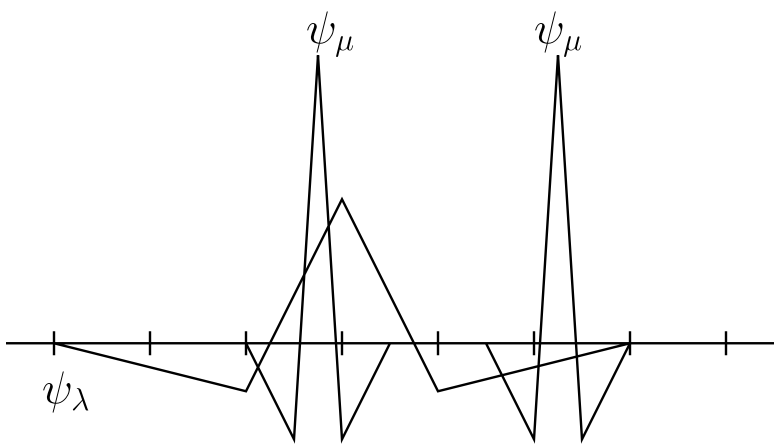
Let $B : \mathcal{X} \to \mathcal{X}'$ with $\mathcal{X} = H_0^m(\Omega)$, $\Omega \subset \mathbb{R}^n$, be defined by

$$(Bu)(v) = \int_\Omega \sum_{|\alpha|,|\beta|\leq m} a_{\alpha,\beta}\partial^\alpha u \partial^\beta v \text{ with } a_{\alpha,\beta} \text{ smooth.}$$

Wavelets of order $d$. Best possible rate $\frac{d-m}{n}$. **Required:** $s^* > \frac{d-m}{n}$.

**Prop 4.** *Consider locally supported* $(\operatorname{diam}(\operatorname{supp}\psi_\lambda) \lesssim 2^{-|\lambda|})$, *piecewise smooth, globally $C^r$ wavelets. Then for* $|\mu| \geq |\lambda|$,

$$|(B\psi_\lambda)(\psi_\mu)| \lesssim \begin{cases} 2^{(|\lambda|-|\mu|)(\frac{n}{2}+\tilde{d}+m)} & \psi_\mu \textit{vanishes on } \operatorname{sing\,supp}\psi_\lambda, \\ 2^{(|\lambda|-|\mu|)(\frac{n}{2}+r+1-m)} & \textit{otherwise.} \end{cases}$$




Ingredients proof: shift derivatives to $\psi_\lambda$ (int. by parts);
use van. moments $\psi_\mu$ and smoothness $\psi_\lambda$.

With a compression scheme based on these estimates, it was shown that

$$s^* = \min\Big(\frac{\tilde{d}+m}{n}, \frac{r+\frac{3}{2}-m}{n-1}\Big),$$

assuming each entry can be computed at unit cost. For spline wavs ($r = d-2$), $s^* > \frac{d-m}{n}$ **when** $\tilde{d} > d-2m$ **and** $\frac{d-m}{n} > \frac{1}{2}$**.**

There exists a quadrature scheme that keeps the error on the level of the compression error, where the average work over each column per entry is $\mathcal{O}(1)$.

Similar results for singular integral operators ([Gantumur, R.S. '06]).

*Rem* 1. Neumann bdr conds or missing vanishing moments of boundary wavelets lead to larger entries. For $n > 1$, this may give an $s^*$ that is too small.

# Non-SPD $\mathbf{B}$

Run adaptive wavelet scheme on $\mathbf{B}^*\mathbf{B}\mathbf{u} = \mathbf{B}^*\mathbf{f}$.

**Prop 5.** *If* $\mathbf{B} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{X}}), \ell_2(\nabla_{\mathcal{Y}}))$, $\mathbf{C} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{Y}}), \ell_2(\nabla_{\mathcal{Z}}))$ *are both* $s^*$*-admissible, then so is* $\mathbf{CB} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{X}}), \ell_2(\nabla_{\mathcal{Z}}))$. *A valid routine* $\mathbf{APPLY}_{\mathbf{CB}}$ *is*

$$[\mathbf{w}, \varepsilon] \mapsto \mathbf{APPLY}_{\mathbf{C}}\big[\mathbf{APPLY}_{\mathbf{B}}[\mathbf{w}, \varepsilon/(2\|\mathbf{C}\|)], \varepsilon/2\big].$$

*If for some* $s^* > s$, $\mathbf{C} \in \mathcal{L}(\ell_2(\nabla_{\mathcal{Y}}), \ell_2(\nabla_{\mathcal{Z}}))$ *is* $s^*$*-admissible, then for*

$$\mathbf{RHS}_{\mathbf{Cf}}[\varepsilon] := \mathbf{APPLY}_{\mathbf{C}}[\mathbf{RHS}_{\mathbf{f}}[\varepsilon/(2\|\mathbf{C}\|)], \varepsilon/2], \tag{1}$$

$\|\mathbf{Cf} - \mathbf{RHS}_{\mathbf{Cf}}[\varepsilon]\|_{\ell_2(\nabla_{\mathcal{Z}})} \leq \varepsilon$,

*and* $\#\mathrm{supp}\,\mathbf{RHS}_{\mathbf{Cf}}[\varepsilon]$ *and work* $\lesssim \varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell_2(\nabla_{\mathcal{X}}))}$.

# CDD1 algorithm [Cohen, Dahmen, DeVore 01]

$(\mathbf{B} = \mathbf{B}^* > 0)$

CDD2: Apply iter. meth. onto $\mathbf{B}\mathbf{u} = \mathbf{f}$

CDD1: Solve seq. of Galerkin problems, use residual $\mathbf{f} - \mathbf{B}\mathbf{u}^{(i)}$ as error indicator for stopping and for guiding the expansion of the set of 'active' wavelets.

Motivation for studying it: Quantitatively better. First indication: Generating subspaces is the most expensive part. If you have one, better to find best approximation from it.

Notations: $\Lambda \subset \nabla$, $\mathbf{I}_\Lambda : \ell_2(\Lambda) \to \ell_2(\nabla)$, $\mathbf{P}_\Lambda = \mathbf{I}_\Lambda^T : \ell_2(\nabla) \to \ell_2(\Lambda)$, $\mathbf{B}_\Lambda = \mathbf{P}_\Lambda \mathbf{B} \mathbf{I}_\Lambda$, $\mathbf{f}_\Lambda = \mathbf{P}_\Lambda \mathbf{f}$, $\mathbf{B}_\Lambda \mathbf{u}_\Lambda = \mathbf{f}_\Lambda$. $|||\cdot||| := \langle \mathbf{B}\cdot, \cdot\rangle^{\frac{1}{2}}$.

$$
\begin{aligned}
\|\mathbf{B}^{-1}\|^{-\frac{1}{2}}\|\cdot\| &\le |||\cdot||| &&\le \|\mathbf{B}\|^{\frac{1}{2}}\|\cdot\| && \text{on } \ell_2(\nabla) \\
\|\mathbf{B}^{-1}\|^{-\frac{1}{2}}|||\cdot||| &\le \|\mathbf{B}\cdot\| &&\le \|\mathbf{B}\|^{\frac{1}{2}}|||\cdot||| && \text{on } \ell_2(\nabla) \\
\|\mathbf{B}^{-1}\|^{-\frac{1}{2}}|||\mathbf{I}_\Lambda\cdot||| &\le \|\mathbf{B}_\Lambda\cdot\| &&\le \|\mathbf{B}\|^{\frac{1}{2}}|||\mathbf{I}_\Lambda\cdot||| && \text{on } \ell_2(\Lambda)
\end{aligned}
$$

**Prop 6** (CDD1). *Let* $\theta \in (0,1]$, $\Lambda \subset \Xi \subset \nabla$, *s.t.*

$$\|\mathbf{P}_\Xi(\mathbf{f} - \mathbf{B}\mathbf{u}_\Lambda)\| \geq \theta \|\mathbf{f} - \mathbf{B}\mathbf{u}_\Lambda\|. \tag{2}$$

*Then* $|||\mathbf{u} - \mathbf{u}_\Xi||| \leq [1 - \kappa(\mathbf{B})^{-1}\theta^2]^{\frac{1}{2}} |||\mathbf{u} - \mathbf{u}_\Lambda|||$.

*Proof.*

$$\begin{aligned} |||\mathbf{u}_\Xi - \mathbf{u}_\Lambda||| &\geq \|\mathbf{B}\|^{-\frac{1}{2}} \|\mathbf{B}(\mathbf{u}_\Xi - \mathbf{u}_\Lambda)\| \geq \|\mathbf{B}\|^{-\frac{1}{2}} \|\mathbf{P}_\Xi \mathbf{B}(\mathbf{u}_\Xi - \mathbf{u}_\Lambda)\| \\ &= \|\mathbf{B}\|^{-\frac{1}{2}} \|\mathbf{P}_\Xi(\mathbf{f} - \mathbf{B}\mathbf{u}_\Lambda)\| \geq \|\mathbf{B}\|^{-\frac{1}{2}} \theta \|\mathbf{f} - \mathbf{B}\mathbf{u}_\Lambda\| \\ &\geq \kappa(\mathbf{B})^{-\frac{1}{2}} \theta |||\mathbf{u} - \mathbf{u}_\Lambda|||. \end{aligned}$$

Now use $|||\mathbf{u} - \mathbf{u}_\Lambda|||^2 = |||\mathbf{u} - \mathbf{u}_\Xi|||^2 + |||\mathbf{u}_\Xi - \mathbf{u}_\Lambda|||^2$. □

CDD1 Algorithm: Using $\mathbf{RHS}_\mathbf{f}$ and $\mathbf{APPLY}_\mathbf{B}$ routines, solve Gal. systems and compute residuals inexactly, add coarsening. Optimality for $\mathbf{u} \in \mathcal{A}^s_\infty$ when $s \in (0, s^*)$.

# Avoidance of coarsening

**Prop 7** (Gantumur, Harbrecht, R.S. '07). *If in* (2), $\theta < \kappa(\mathbf{B})^{-\frac{1}{2}}$ *and* $\Xi$ *is the* smallest *set satisfying* (2), *then* $\#(\Xi\backslash\Lambda) \leq N$ *for sm.* $N$ *s.t.*

$$|||\mathbf{u} - \mathbf{u}_N||| \leq [1 - \theta^2\kappa(\mathbf{B})]^{\frac{1}{2}}|||\mathbf{u} - \mathbf{u}_\Lambda|||.$$

*Proof.* Let $\Sigma := \Lambda \cup \operatorname{supp} \mathbf{u}_N$. Then $|||\mathbf{u} - \mathbf{u}_\Sigma||| \leq |||\mathbf{u} - \mathbf{u}_N|||$, and so $|||\mathbf{u}_\Sigma - \mathbf{u}_\Lambda||| \geq \theta\kappa(\mathbf{B})^{\frac{1}{2}}|||\mathbf{u} - \mathbf{u}_\Lambda|||$. This gives

$$\begin{aligned}\|\mathbf{P}_\Sigma(\mathbf{f} - \mathbf{B}\mathbf{u}_\Lambda)\| = \|\mathbf{B}_\Sigma(\mathbf{u}_\Sigma - \mathbf{u}_\Lambda)\| &\geq \|\mathbf{B}^{-1}\|^{-\frac{1}{2}}|||\mathbf{u}_\Sigma - \mathbf{u}_\Lambda||| \\ &\geq \|\mathbf{B}^{-1}\|^{-\frac{1}{2}}\theta\kappa(\mathbf{B})^{\frac{1}{2}}|||\mathbf{u} - \mathbf{u}_\Lambda||| \geq \theta\|\mathbf{f} - \mathbf{B}\mathbf{u}_\Lambda\|.\end{aligned}$$

By assumption on $\Xi$, $\#(\Xi\backslash\Lambda) \leq \#(\Sigma\backslash\Lambda) \leq N$. $\square$

**Corol 2.** *If* $\mathbf{u} \in \mathcal{A}^s_\infty$, *then* $\#(\Xi\backslash\Lambda) \lesssim |||\mathbf{u} - \mathbf{u}_\Lambda|||^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty}$.

Majorized linear convergence + upper bound on sizes of expansions gives quasi-optimal support sizes: Let

$$\Lambda_1 \subset \Lambda_2 \subset \cdots,$$
$$\mathbf{u}_{\Lambda_1}, \mathbf{u}_{\Lambda_2}, \ldots,$$

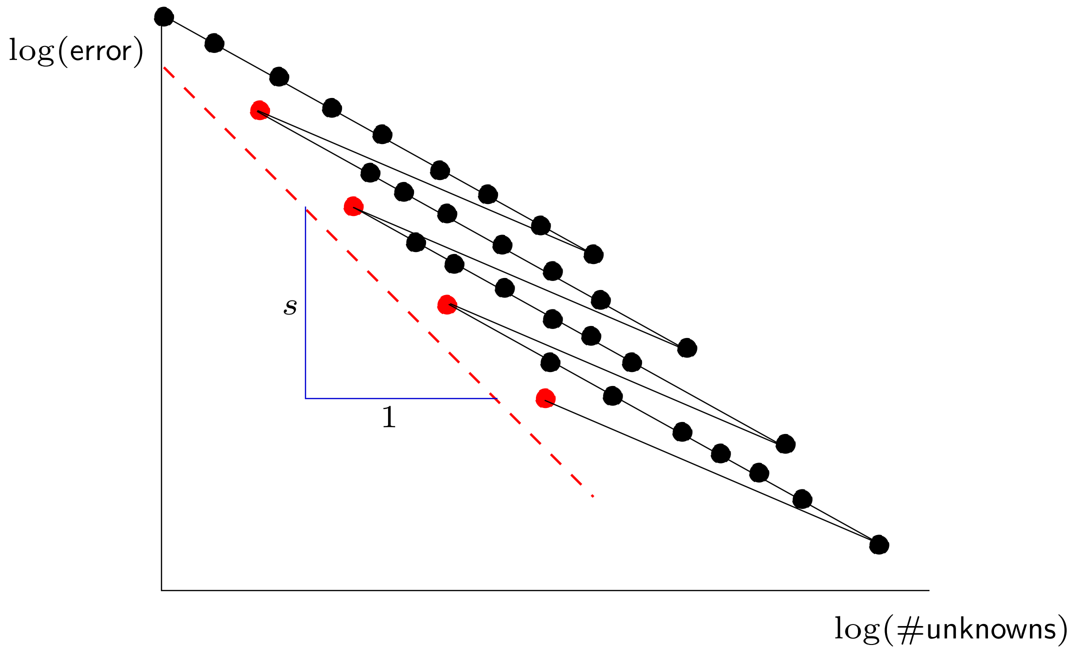
be produced by adapt wav-Gal scheme stopped when $\|\mathbf{f} - \mathbf{B}\mathbf{u}_{\Lambda_\ell}\| \leq \varepsilon$. Then

$$\begin{aligned}\#\Lambda_\ell = \sum_{k=1}^{\ell} \#(\Lambda_k \backslash \Lambda_{k-1}) &\lesssim \sum_{k=1}^{\ell} |||\mathbf{u} - \mathbf{u}_{\Lambda_{k-1}}|||^{-1/s} \|\mathbf{u}\|_{\mathcal{A}^s}^{1/s} \\ &\lesssim |||\mathbf{u} - \mathbf{u}_{\Lambda_{\ell-1}}|||^{-1/s} \|\mathbf{u}\|_{\mathcal{A}^s}^{1/s} \lesssim \varepsilon^{-1/s} \|\mathbf{u}\|_{\mathcal{A}^s}^{1/s}.\end{aligned}$$

**SOLVE**$[\nu_{-1},\varepsilon] \to \mathbf{w}_k$:
*% Let* $\alpha,\omega,\gamma,\theta$ *be constants with* $\omega \in (0,\alpha)$, $\frac{\alpha+\omega}{1-\omega} < \kappa(\mathbf{A})^{-\frac{1}{2}}$, $\theta > 0$ *and*
*%* $\gamma \in \left(0, \frac{1}{6}\kappa(\mathbf{A})^{-\frac{1}{2}}\frac{\alpha-\omega}{1+\omega}\right)$. *The parameter* $\nu_{-1}$ *is an estimate for the norm*
*% of the initial residual* $\mathbf{f}$.

$k := 0$, $\mathbf{w}_k := 0$, $\mathbf{\Lambda}_k := \emptyset$
`do` $\zeta := \theta\nu_{k-1}$
  `do` $\zeta := \zeta/2$, $\mathbf{r}_k := \mathbf{RHS}_{\mathbf{f}}[\boldsymbol{\nabla}, \zeta/2] - \mathbf{APPLY}_{\mathbf{B}}[\boldsymbol{\nabla}, \mathbf{w}_k, \zeta/2]$
    `if` $\nu_k := \|\mathbf{r}_k\| + \zeta \leq \varepsilon$ `then stop endif`
  `until` $\zeta \leq \omega\|\mathbf{r}_k\|$
  $[\cdot, \Pi] := \mathbf{BUCKETSORT}[\mathbf{r}_k|_{\boldsymbol{\nabla}\setminus\mathbf{\Lambda}_k}, \sqrt{1-\alpha^2}\,\|\mathbf{r}_k\|]$
  *determine the smallest* $K \in \mathbb{N}_0$ *with*
    $\|\mathbf{r}_k|_{\boldsymbol{\nabla}\setminus\mathbf{\Lambda}_k} - \mathbf{r}_k|_{\Pi(\{1,\ldots,K\})}\| \leq \sqrt{1-\alpha^2}\,\|\mathbf{r}_k\|$
  $\mathbf{\Lambda}_{k+1} := \mathbf{\Lambda}_k \cup \Pi(\{1,\ldots,K\})$
  $\mathbf{g}_{k+1} := \mathbf{RHS}_{\mathbf{f}}[\mathbf{\Lambda}_{k+1}, \gamma\nu_k]$
  $\mathbf{w}_{k+1} := \mathbf{GALSOLVE}[\mathbf{\Lambda}_{k+1}, \mathbf{g}_{k+1}, \mathbf{w}_k, (1+\gamma)\nu_k, \gamma\nu_k]$
  $k := k+1$
`enddo`

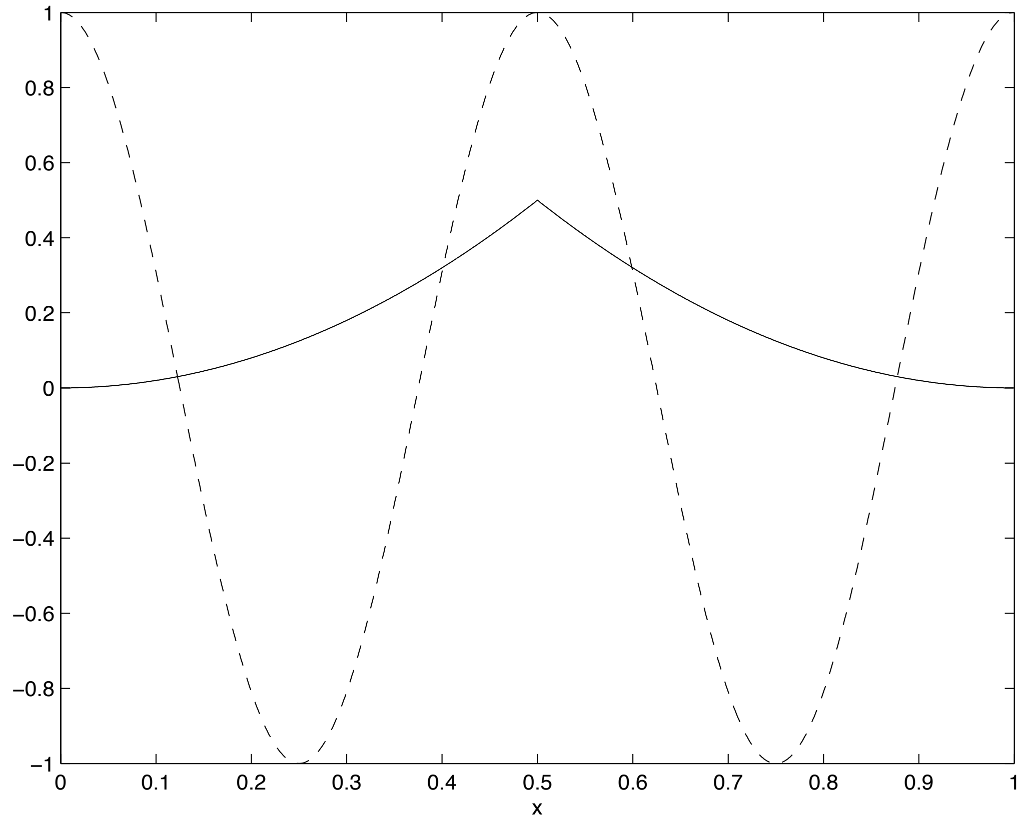
**Th 4.** $\|\mathbf{u} - \mathbf{w}_k\| \leq \varepsilon$. *If, for some* $s \in (0, s^*)$, $\mathbf{u} \in \mathcal{A}^s_\infty$, *then* $\#\mathrm{supp}\,\mathbf{w}_k$ *and work* $\lesssim \varepsilon^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s}$.



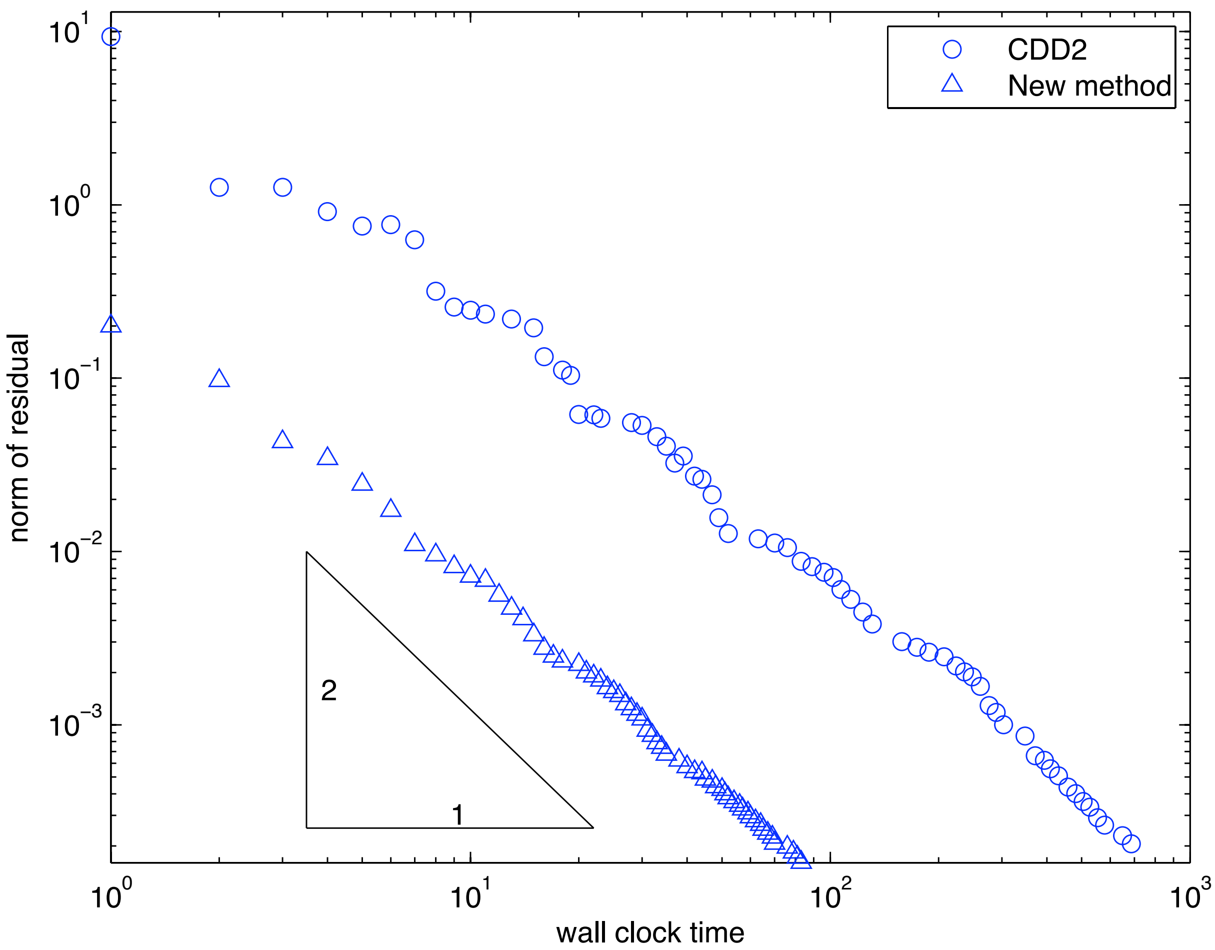
$\log(\text{error})$
$s$
1
$\log(\#\text{unknowns})$

# Numerical illustration

$$-u'' + u = f \quad \text{on } \mathbb{R}/\mathbb{Z}, \text{ with } u(x) = \cos(4\pi x) + \begin{cases} 2x^2, & \text{if } x \in [0, 1/2), \\ 2(1-x)^2, & \text{if } x \in [1/2, 1], \end{cases}$$



B-splines of order 3.



# Summary

For a wide class of operators, adaptive wavelet methods realize rate of convergence of best $N$-term approximation in linear complexity. With CDD1 method coarsening can be avoided.

## References


[1] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations – Convergence rates. *Math. Comp*, 70:27–75, 2001.

[2] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods II - Beyond the elliptic case. *Found. Comput. Math.*, 2(3):203–245, 2002.

[3] W. Dahmen, H. Harbrecht, and R. Schneider. Adaptive methods for boundary integral equations - complexity and convergence estimates. *Math. Comp.*, 76:1243–1274, 2007.

[4] T. Gantumur, H. Harbrecht, and R. S. An optimal adaptive wavelet method without coarsening of the iterands. *Math. Comp.*, 77:615–629, 2007.

[5] T. Gantumur and R. S. Computation of differential operators in wavelet coordinates. *Math. Comp.*, 75:697–709, 2006.

[6] T. Gantumur and R. S. Computation of singular integral operators in wavelet coordinates. *Computing*, 76:77–107, 2006.

[7] R. S. On the compressibility of operators in wavelet coordinates. *SIAM J. Math. Anal.*, 35(5):1110–1132, 2004.

# Tensor product bases

Let $\square = (0,1)^n$. Let $\Psi$ be a Riesz basis for $L_2(0,1)$ of order $d > m$ such that $\{2^{-|\lambda|m}\psi_\lambda : \lambda \in \nabla\}$ is Riesz basis for $H^m(0,1)$. Thanks to

$$H^m(\square) = H^m \otimes L_2 \otimes \cdots \otimes L_2 \cap \; \cdots \; \cap L_2 \otimes \cdots \otimes L_2 \otimes H^m,$$

after normalization,

$$\boldsymbol{\Psi} = \Psi \otimes \cdots \otimes \Psi$$

is Riesz basis $H^m(\square)$.

$$\boldsymbol{\psi}_{\boldsymbol{\lambda}}(x) = \prod_{i=1}^{n} \psi_{\lambda_i}(x_i), \quad |\boldsymbol{\lambda}| = (|\lambda_1|, \ldots, |\lambda_n|) \in \mathbb{N}^n.$$

# Sparse grids [Zenger '90], [Bungartz, Griebel '04]

Sparse grid space: $\mathrm{span}\{\psi_{\boldsymbol{\lambda}} : |||\boldsymbol{\lambda}|||_1 \leq k\}$. With such a space of dim $N$, error is $\mathcal{O}(N^{-(d-m)}(\log N)^{\cdots})$ **(no curse)** when

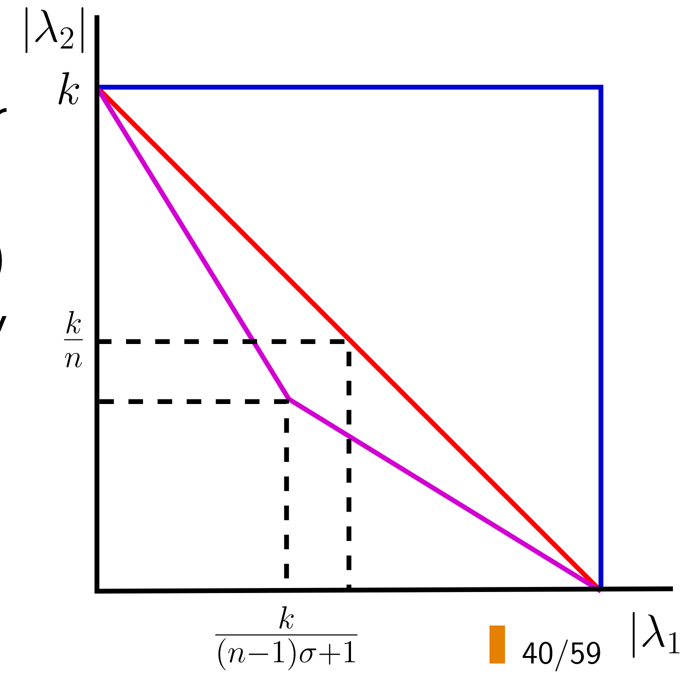
$$\partial_i^m \partial_1^{d-m} \partial_2^{d-m} \cdots \partial_n^{d-m} u \in L_2(\square) \quad (1 \leq i \leq n).$$

*Rem* 2. With standard (isotropic) approx of order $\bar{d}$, rate is $\frac{\bar{d}-m}{n}$ when $u \in H^{\bar{d}}(\square)$.

$$\tfrac{\bar{d}-m}{n} = d - m \iff \bar{d} = n(d-m) + m$$

So instead that some (mixed) partial derivatives of order $n(d-m)+m$ have to be in $L_2(\square)$, all partial derivatives of this order should be in $L_2(\square)$.

**Optimized sparse grids:** [Griebel, Knapek '00] $\mathrm{span}\{\psi_{\boldsymbol{\lambda}} : \sigma|||\boldsymbol{\lambda}|||_1 - (\sigma - 1)|||\boldsymbol{\lambda}|||_\infty \leq k\}$ for some $\sigma \in (1, \frac{d}{d-m})$ .
With such a space of dim $N$, error $\mathcal{O}(N^{-(d-m)})$ when $\partial_1^d \ldots \partial_n^d u \in L_2(\square)$ (can be slightly reduced).



# Best $N$-term approximation in tensor product bases

Despite their reduction, usually reg. conds for lin. approx. not satisfied by sols (elliptic) PDEs (unless rhs $f$ is smooth and van. near corners, edges).

Since $\mathbf{\Psi}$ is Riesz, $u \in \mathcal{A}_q^s(H^m(\square), \mathbf{\Psi}) \Longleftrightarrow \mathbf{u} \in \mathcal{A}_q^s(\ell_2)$.

[Nitsche '06]: For $s \in (0, d-m)$, $\tau = (\frac{1}{2} + s)^{-1}$,

$$u \in \bigcap_{j=1}^{n} \bigotimes_{i=1}^{n}{}_{\tau} B_{\tau,\tau}^{s+\delta_{ij}m}(0,1) \Longleftrightarrow \mathbf{u} \in \mathcal{A}_\tau^s$$

(thinking of $s = d - m$ (not covered), same partial derivatives as for rate $s$ (mod log-factors) with sparse grids, but now boundedness in $L_\tau$ instead of in $L_2$)

[Nitsche '05]: Solutions of elliptic PDEs (for $\Omega \neq \square$, as function transported to cube) have infinite smoothness in this scale.

For $n \geq 3$, not true for scale that governs best $N$-term approximation with isotropic wavelets. Generally, anisotropic refinements towards the boundary are needed.

# Adaptive tensor product wavelet methods

To realize rate of best $N$-term approx: AWM.

**Prop 8** (Schwab, R.S. '07)**.** *Consider* $(Bu)(v) = \int_{\Box} \sum_{|\alpha|,|\beta| \leq m} a_{\alpha,\beta} \partial^\alpha u \partial^\beta v$ *with* $a_{\alpha,\beta}$ *smooth. With locally supported, piecewise polynomial univariate wavelets that have* $\tilde{d} > d - m$ *vanishing moments,* $b(\boldsymbol{\Psi}, \boldsymbol{\Psi})$ *is* $s^*$*-computable with* $s^* > d - m$*.*

# High dim problems: Depend. of “hidden” constants on $n$

$\kappa(\Sigma, H) := \sup / \inf$ of $\dfrac{\|\mathbf{v}^\top \Sigma\|_H^2}{\sum_\lambda |\mathbf{v}_\lambda|^2 \|\sigma_\lambda\|_H^2}$.

$H = H_0^1(\Box)$ equipped with $|\cdot|_{H^1(\Box)}$, $\mathbf{\Psi} = \otimes_{i=1}^n \Psi$ with $\Psi$ Riesz for $L_2(0,1)$ and, prop. sc., for $H_0^1(0,1)$. Then

$$\kappa(\Psi, L_2(0,1))^{n-1} \lesssim \kappa(\mathbf{\Psi}, H_0^1(\Box)) \leq \kappa(\Psi, H_0^1(0,1))\kappa(\Psi, L_2(0,1))^{n-1}.$$

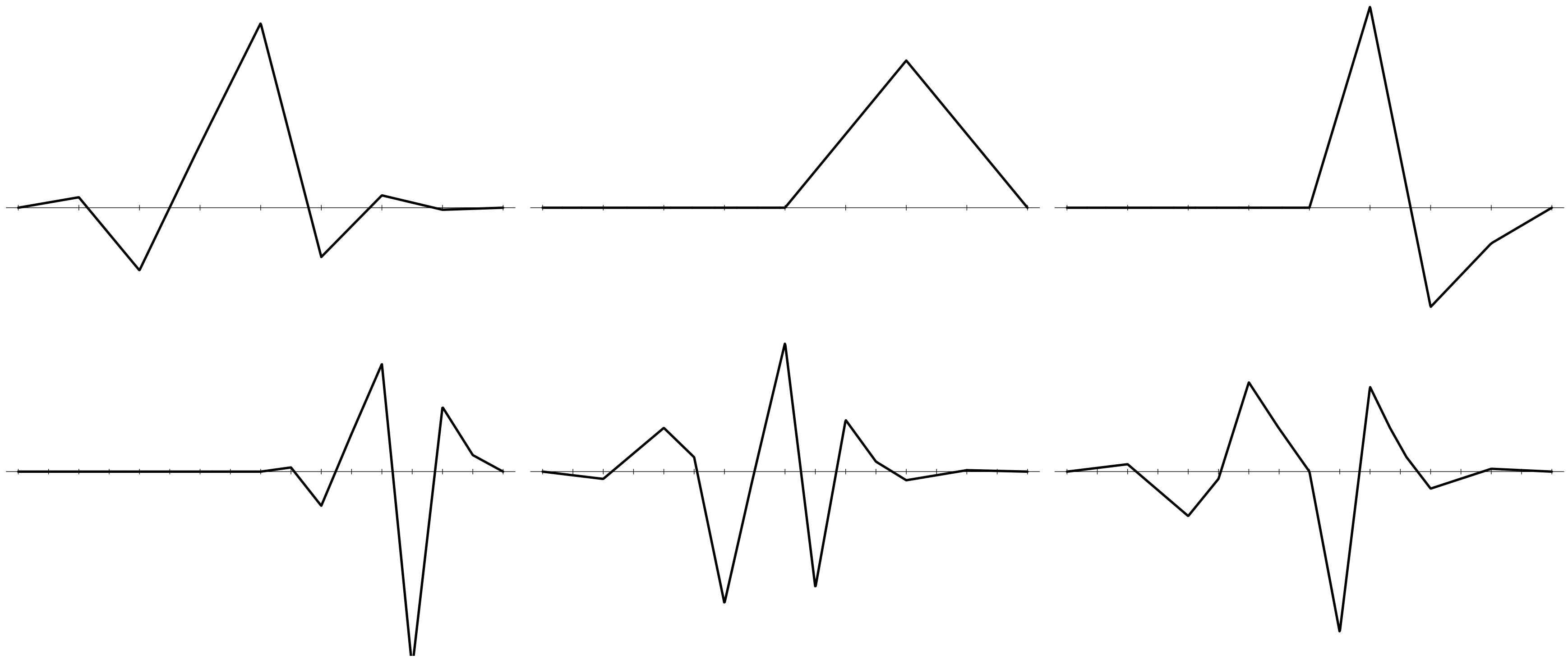
Quantitatively, both

$$\mathcal{A}_\infty^s(H_0^1(\Box), \mathbf{\Psi})) \sim \mathcal{A}_\infty^s(\ell_2) \sim \sup_j (\#\mathrm{supp}\, \mathbf{u}^{(j)})^s \|\mathbf{u} - \mathbf{u}^{(j)}\|,$$

with $(\mathbf{u}^{(j)})_j$ seq. prod. by AWM, depend critically but exclusively on $\kappa$.

So for high dim problems, only option is to use $L_2(0,1)$ orth. wavelets. For compression, we need piecewise smooth wavs.

[Donovan, Geronimo, Hardin ’96]: Piecewise pol. $L_2$-orth. multiwavs. $d = 2$:

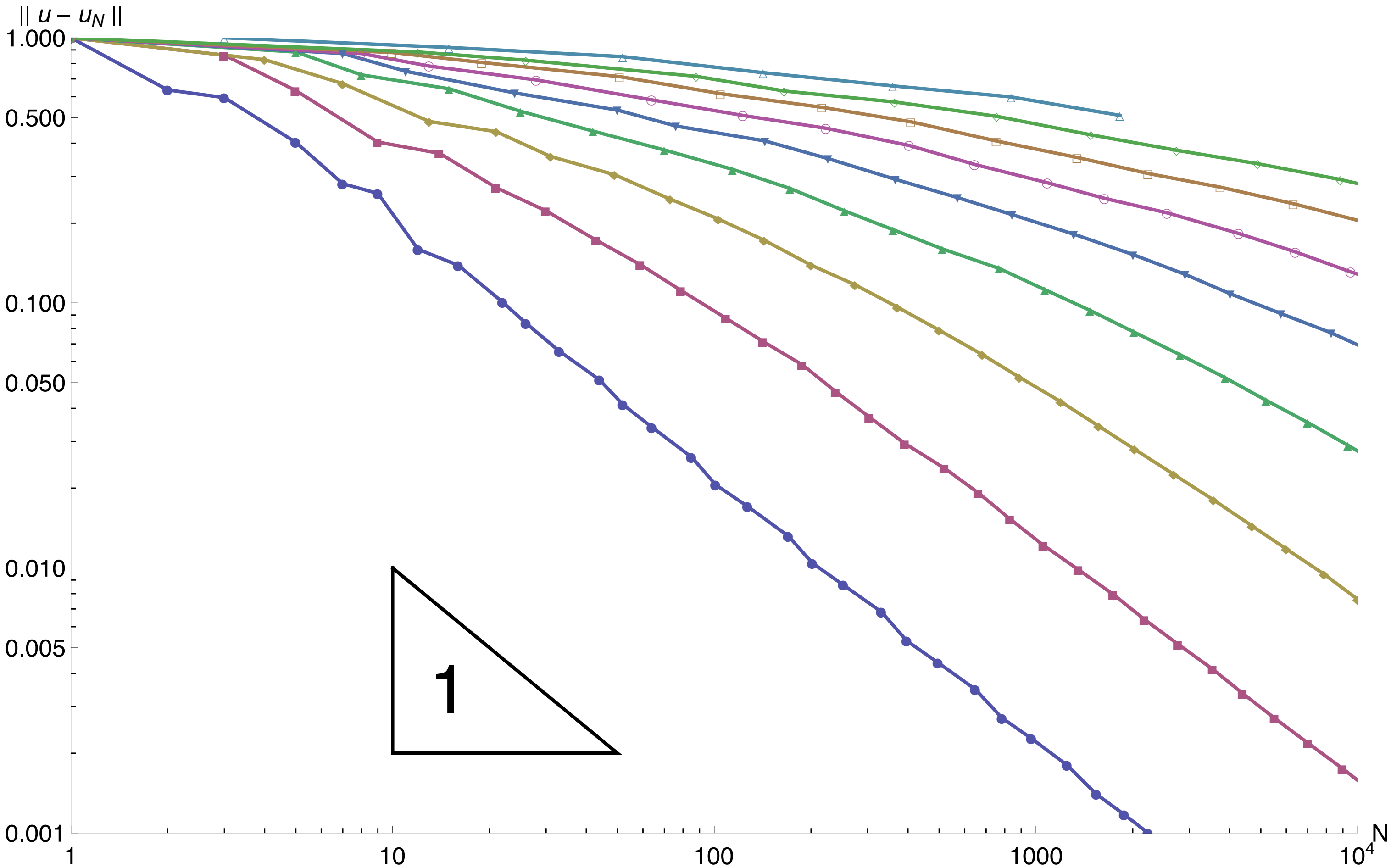


Work count AWM depends on $s^*$-computability constants:

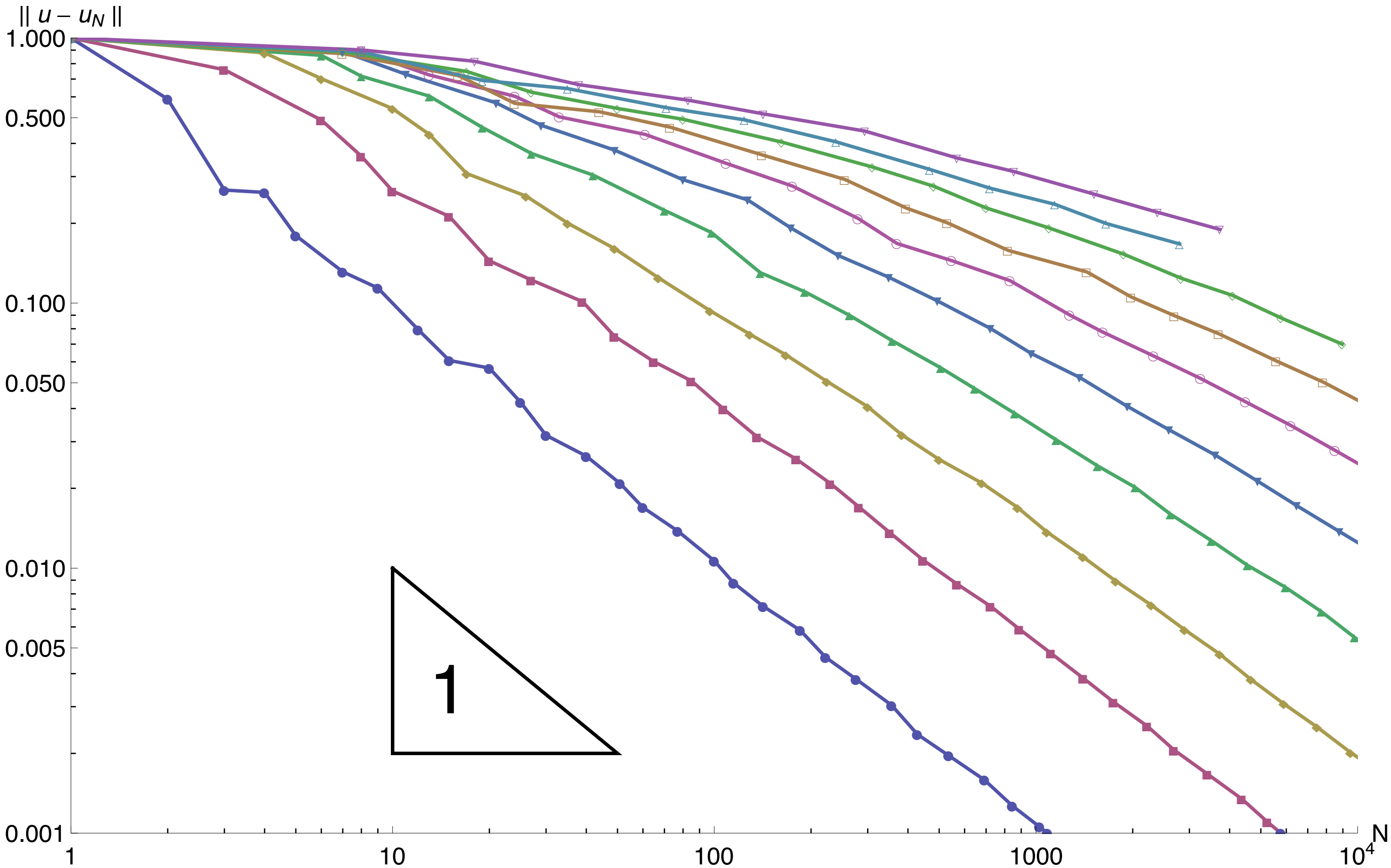
**Th 5** (Dijkema, Schwab, R.S. '07). *For Poisson on* $\square$*, given* $\varepsilon > 0$*, adaptive tensor product wavelet method using DGH wavelets produces approximation within tolerance* $\varepsilon$*, with, whenever* $u \in \mathcal{A}^s_\infty$*, support length* $\leq C\varepsilon^{-1/s}\|u\|^{1/s}_{\mathcal{A}^s_\infty}$ *and work* $\leq Dn\varepsilon^{-1/s}\|u\|^{1/s}_{\mathcal{A}^s_\infty}$*, with* $C$*,* $D$ *being absolute constants.*

# High dim problems: Numerical results

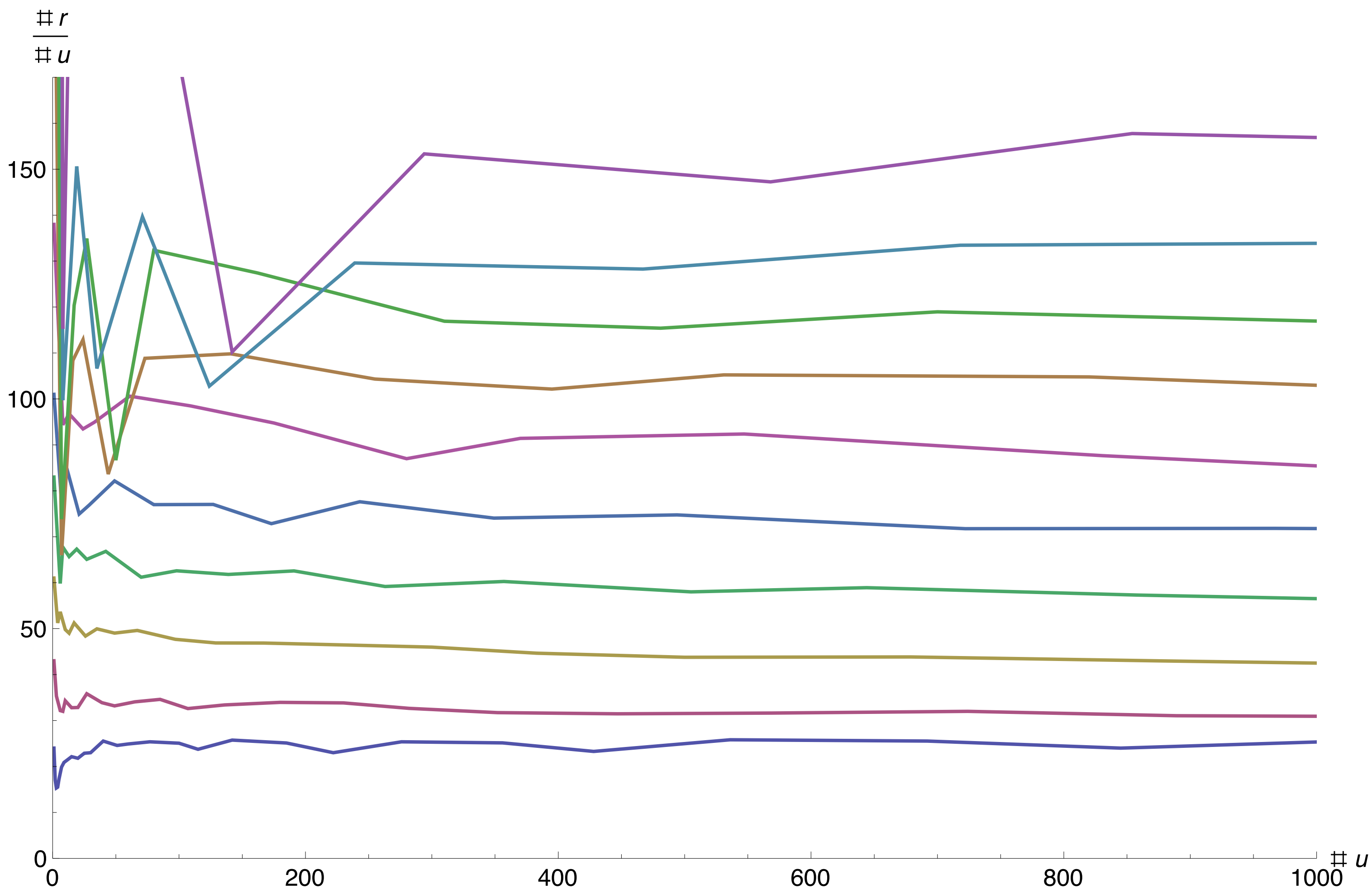
Poisson on $(0,1)^n$, $f = 1$, hom Dirichlet, $n = 1, \ldots, 9$.



Poisson on $(0,1)^n$, $f = 1$, hom Dirichlet only at left bdr (tensorized), $n = 1, \ldots, 10$



Work



# Summary

On product domains, the curse of dimensionality can be circumvented by adaptive tensor product wavelet schemes.

## References

[1] H.J. Bungartz and M. Griebel. Sparse grids. *Acta Numer.*, 13:147–269, 2004.

[2] T.J. Dijkema, Ch. Schwab, and R. S. An adaptive wavelet method for solving high-dimensional elliptic PDEs. Technical report, January/February 2008. In preparation.

[3] G.C. Donovan, J.S. Geronimo, and D.P. Hardin. Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets. *SIAM J. Math. Anal.*, 27(6):1791–1815, 1996.

[4] M. Griebel and S. Knapek. Optimized tensor-product approximation spaces. *Constr. Approx.*, 16(4):525–540, 2000.

[5] P.-A. Nitsche. Sparse approximation of singularity functions. *Constr. Approx.*, 21(1):63–81, 2005.

[6] P.-A. Nitsche. Best $N$-term approximation spaces for tensor product wavelet bases. *Constr. Approx.*, 41(1):49–70, 2006.

[7] Ch. Schwab and R. S. Adaptive wavelet algorithms for elliptic PDEs on product domains. *Math. Comp.*, 77:71–92, 2008.

[8] Ch. Zenger. Sparse grids. In *Parallel algorithms for partial differential equations (Kiel, 1990)*, volume 31 of *Notes Numer. Fluid Mech.*, pages 241–251. Vieweg, Braunschweig, 1991.

# Parabolic problems

Let $V \hookrightarrow H \hookrightarrow V'$, $I := (0, T)$. Given $g \in L_2(I; V') = \{f : I \to V' : \int_I \|f(t)\|^2_{V'} dt < \infty\}$ and $h \in H$, consider parabolic problem

$$\frac{du}{dt}(t) + A(t)u(t) = g(t) \quad \text{in } V', \quad u(0) = h \text{ in } H,$$

where $a(t; \eta, \zeta) := (A(t)(\eta))(\zeta)$ satisfies for a.e. $t \in I$,

$$|a(t; \eta, \zeta)| \leq M_a \|\eta\|_V \|\zeta\|_V \quad (\eta, \zeta \in V) \quad (\textit{boundedness}),$$
$$\Re a(t; \eta, \eta) + \lambda \|\eta\|^2_H \geq \alpha \|\eta\|^2_V \quad (\eta \in V) \quad (\textit{coercivity}).$$

Ex. $A(t)$ differential or integrodifferential operator of order $2m > 0$.

Classical approaches: Methods of lines or Rothe's method.
Not easy to get optimal distribution of work over time; even harder to get optimality simultaneously over space and time.

# Variational formulation

Let

$$\begin{aligned}\mathcal{X} &= L_2(I;V) \cap H^1(I;V') \eqsim (L_2(I) \otimes V) \cap (H^1(I) \otimes V') \\ \mathcal{Y} &= L_2(I;V) \times H \qquad\qquad \eqsim (L_2(I) \otimes V) \times H\end{aligned}$$

Find $u \in \mathcal{X}$ s.t.

$$b(u,v) = f(v) \qquad (v = (v_1, v_2) \in \mathcal{Y})$$

where

$$\begin{aligned}b(w,(v_1,v_2)) &:= \int_I \langle \tfrac{dw}{dt}(t), v_1(t)\rangle_H + a(t; w(t), v_1(t))dt + \langle w(0), v_2\rangle_H, \\ f(v_1,v_2) &:= \int_I \langle g(t), v_1(t)\rangle_H dt + \langle h, v_2\rangle_H.\end{aligned}$$

**Th 6** (Dautray & Lions '92, Wloka '82)**.** *$B : \mathcal{X} \to \mathcal{Y}'$ defined by $(Bw)(v) = b(v,w)$ is boundedly invertible.*

# Wavelet bases

Let, properly scaled, $\Theta \subset H^1(I)$ or $\Sigma \subset V$ be temporal or spatial wavelet bases for $L_2(I)$ and $H^1(I)$, or for $V'$, $H$ and $V$, resp.

Then

$$\left\{ \frac{\theta_\lambda \otimes \sigma_\mu}{\sqrt{\|\theta_\lambda\|^2_{L_2(I)}\|\sigma_\mu\|^2_V + \|\theta_\lambda\|^2_{H^1(I)}\|\sigma_\mu\|^2_{V'}}} : (\lambda, \mu) \in \nabla_{\mathcal{X}} := \nabla_t \times \nabla_x \right\},$$

$$\left\{ \left( \frac{\theta_\lambda \otimes \sigma_\mu}{\|\theta_\lambda\|_{L_2(I)}\|\sigma_\mu\|_V}, 0 \right) : (\lambda, \mu) \in \nabla_t \times \nabla_x \right\} \cup \left\{ (0, \frac{\sigma_\mu}{\|\sigma_\mu\|_H}) : \mu \in \nabla_x \right\}$$

are Riesz bases for $\mathcal{X}$ or $\mathcal{Y}$, resp.

$$Bu = f \Longleftrightarrow \mathbf{Bu} = \mathbf{f}.$$

AWMs applicable onto $\mathbf{B}^*\mathbf{Bu} = \mathbf{B}^*\mathbf{f}$. To verify $s^*$-admissability of $\mathbf{B}$ and $\mathbf{B}^*$ for $s^*$ larger than any $s$ for which $\mathbf{u} \in \mathcal{A}^s_\infty$ can be expected.

## Rates of best $N$-term approximation

Let temporal wavs $\Theta$ be of order $d_t$; $H = L_2(\Omega)$, $V = H^m(\Omega)$ (or $H_0^m(\Omega)$), $\Omega \subset \mathbb{R}^n$.

2 cases for spatial wavs $\Sigma$:

A) isotropic wav of order $d_x$

B) $\Omega = \square = (0,1)^n$, $\Sigma = \otimes_{i=1}^n \Sigma_i$, $\Sigma_i$ univariate wavs of order $d_x$

Using (optimized) sparse grids: In any case for smooth $u$,

$$\mathbf{u} \in \begin{cases} \mathcal{A}_\infty^{\min(d_t-1,\frac{d_x-m}{n})} & \text{case A)} \\ \mathcal{A}_\infty^{\min(d_t-1,d_x-m)} & \text{case B)} \end{cases}$$

So no penalty for working in $n+1$ dims, and in case B) no curse.

*Rem* 3. Charac. of $\mathcal{A}_\infty^s$ for $0 < s < \begin{cases} \min(d_t-1, \frac{d_x-m}{n}) & \text{case A)} \\ \min(d_t-1, d_x-m) & \text{case B)} \end{cases}$ in terms of (tensor products of ) Besov spaces seems possible.

## $s^*$-admissibility of $B$ in wavelet coordinates

With $[\Phi]_Z$ denoting $\Phi$ normalized in $\|\cdot\|_Z$, $\mathbf{B} =$

$$\begin{bmatrix} \langle[\Theta]'_{H^1},[\Theta]_{L_2}\rangle_{L_2}\otimes\langle[\Sigma]_{V'},[\Sigma]_V\rangle_H\mathbf{D}_1+\int_I a(t,[\Theta]_{L_2}\otimes[\Sigma]_V,[\Theta]_{L_2}\otimes[\Sigma]_V)dt\mathbf{D}_2 \\ \langle\Sigma,\Sigma\rangle_H\mathbf{R} \end{bmatrix}$$

where $\mathbf{D}_1$, $\mathbf{D}_2$ are diagonal matrices with $|\text{entries}| \leq 1$,
$\mathbf{R} \in \mathcal{L}(\ell_2(\nabla_t \times \nabla_x), \ell_2(\nabla_x))$ is given by

$$\mathbf{R}_{\mu,(\lambda,\nu)} = \begin{cases} \dfrac{\theta_\lambda(0)}{\sqrt{\|\theta_\lambda\|^2_{L_2}\|\sigma_\mu\|^2_V + \|\theta_\lambda\|^2_{H^1}\|\sigma_\mu\|^2_{V'}}} & \text{when } \mu = \nu, \\ 0 & \text{otherwise.} \end{cases}$$

*Rem* 4. If $a(t;\eta,\zeta) = a(\eta,\zeta)$, then

$$\int_I a(t,[\Theta]_{L_2}\otimes[\Sigma]_V,[\Theta]_{L_2}\otimes[\Sigma]_V)dt = \langle[\Theta]_{L_2},[\Theta]_{L_2}\rangle_{L_2}\otimes a([\Sigma]_V,[\Sigma]_V).$$

**Lem 2.** *Let* $\mathbf{A}$, $\mathbf{C}$ *be* $s^*$*-computable, then so is* $\mathbf{A} \otimes \mathbf{C}$.

**Th 7.** *Let* $H = L_2(\Omega)$, $V = H_0^m(\Omega)$, $a(\cdot,\cdot,\cdot)$ *diff op with suff sm coeffs. Consider locally supported piecewise smooth temporal and spatial wavs. Then* $\mathbf{B}$ *and* $\mathbf{B}^*$ *are* $s^*$*-admissible with*

$$s^* > \begin{cases} \min(d_t - 1, \frac{d_x - m}{n}) & \textit{case } \mathrm{A}) \\ \min(d_t - 1, d_x - m) & \textit{case } \mathrm{B}) \end{cases}$$

*when* $\tilde{d}_t > d_t - 1$, *and in case* A), $d_x = 3$, $n \in \{1, 2, 3\}$, $\tilde{d}_x > d_x$, *and in case* B), $\tilde{d}_t > d_t - 1$, $\tilde{d}_x > d_x - m$.

*Rem* 5. Restrictions on params in case A) due to fact that, properly sc., $\Sigma$ has to be basis for $V'$, i.e., dual wavs have to be basis for $V$. Consequently, for $V = H_0^m(\Omega)$, primal bdr wavs have no v.m.'s.

Similar results for singular integral operators (in case A)).

# Parabolic problems in high space dims

Consider $\Omega = \square$; $a(t;\eta,\zeta) = \int_\square \nabla\eta \cdot \nabla\zeta$; case B), i.e., $\Sigma = \otimes_{i=1}^n \Sigma_i$, with $\Sigma_i$ an $L_2(0,1)$-orthogonal basis. Then $\|\mathbf{B}\|\|\mathbf{B}^{-1}\|$ bounded unif. in $n$.

$$\mathbf{B} = \begin{bmatrix} (\langle[\Theta]'_{H^1}, [\Theta]_{L_2}\rangle_{L_2} \otimes \mathrm{Id}_x)\mathbf{D}_1 + \langle[\Theta]_{L_2}, [\Theta]_{L_2}\rangle_{L_2} \otimes a([\Sigma]_V, [\Sigma]_V)\mathbf{D}_2 \\ \mathbf{R} \end{bmatrix}.$$

Let the wavs be locally supported, piecewise pol (DGH spatial wavs). Then for any $s^*$, $\mathbf{B}$ and $\mathbf{B}^*$ are $s^*$-admissible with "hidden constant" growing at most linearly with $n$.

**Th 8** (Schwab, R.S. '07). *Given $\varepsilon > 0$, adaptive wavelet method produces approximation within tolerance $\varepsilon$, with, whenever $u \in \mathcal{A}^s_\infty$, support length $\leq C\varepsilon^{-1/s}\|u\|^{1/s}_{\mathcal{A}^s_\infty}$ and work $\leq D n^2 \varepsilon^{-1/s}\|u\|^{1/s}_{\mathcal{A}^s_\infty}$, with $C$, $D$ being absolute constants.*

# Summary

Parabolic problems can be formulated as well-posed operator equations simultaneously in space and time.

Thanks to product structure of space-time cylinder, tensor product can be applied of bases in space and time, and so no reduction in rate, or equivalently, payment in complexity, due to the additional time dimension.

Adaptive wavelet methods give best possible rates.

For high dimensional spatial product domains, the “curse” can be circumvented by taking spatial tensor product wavelets.

## References

[1] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 5*. Springer-Verlag, Berlin, 1992. Evolution problems I.

[2] M. Griebel and D. Oeltz. A Sparse Grid Space-Time Discretization Scheme for Parabolic Problems. *Computing*, 2007. DOI 10.1007/s00607-007-0241-3.

[3] Ch. Schwab and R.S. A space-time adaptive wavelet method for parabolic evolution problems. Technical report, January 2008. Submitted

[4] T. von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. *M2AN Math. Model. Numer. Anal.*, 38(1):93–127, 2004.

[5] J. Wloka. *Partielle Differentialgleichungen*. B. G. Teubner, Stuttgart, 1982. Sobolevräume und Randwertaufgaben.