# 25+ Years of Wavelets for PDEs (Partial Differential Equations)

Angela Kunoth

University of Cologne, Germany

MAIA, September 2016

# 25+ Years of Wavelets for PDEs (Partial Differential Equations)

Angela Kunoth

University of Cologne, Germany

MAIA, September 2016

Wavelets in signal and image processing . . .

- Signal or image: explicitly given object described by $N$ data points

# 25+ Years of Wavelets for PDEs (Partial Differential Equations)

Angela Kunoth

University of Cologne, Germany

MAIA, September 2016

Wavelets in signal and image processing . . .

- Signal or image: explicitly given object described by $N$ data points
- Goal: data compression without loosing essential information
- Method: single-(fine-)scale $\longleftrightarrow$ multi-scale representation of object
- Change of representation by Fast Wavelet Transform in $\mathcal{O}(N)$ operations (based on locally supported functions)

  $\rightsquigarrow$ Discard small coefficients in multi-scale representation

  $\rightsquigarrow$ Data compression
- Landmark: Daubechies' construction of $L_2(\mathbb{R})$ orthonormal wavelets with compact support [1988]

## Image Compression — (Old) Examples

Original (768×768 pixels, 589.824 bytes) JPEG compression (12.9:1, 45.853 bytes)
Wavelet compression: JPEG 2000 (12.9:1, 45.621 bytes) [Brislawn, FBI, Los Alamos Laboratory, 1996]

# Image Compression — (Old) Examples

Original (768×768 pixels, 589.824 bytes)    JPEG compression (12.9:1, 45.853 bytes)
Wavelet compression: JPEG 2000 (12.9:1, 45.621 bytes)    [Brislawn, FBI, Los Alamos Laboratory, 1996]

Original (left), compression 100:1 (MT-WICE (Wavelet Based Image Compression), Mevis, right)

Compression 80:1 (MT-WICE left)    JPEG (right)

## Wavelets: Multiscale Basis with Additional Properties

Image $v \in V_N \subset L_2(\mathbb{R}^2)$ (or $L_2((0,1)^2)$ with dim $V_N = N < \infty$

## Wavelets: Multiscale Basis with Additional Properties

Image $v \in V_N \subset L_2(\mathbb{R}^2)$ (or $L_2((0,1)^2)$ with dim $V_N = N < \infty$

... more general:
consider objects in (infinite-dimensional) Hilbert space $H$ on domain $\Omega \subset \mathbb{R}^d$ with $\|\cdot\|_H$

$\boldsymbol{\Psi} := \{\psi_\lambda : \lambda \in \mathbb{I}\} \subset H$ wavelets, $\mathbb{I}$ (infinite) index set, $\lambda$ index: resolution $|\lambda|$, location $k$...

# Wavelets: Multiscale Basis with Additional Properties

Image $v \in V_N \subset L_2(\mathbb{R}^2)$ (or $L_2((0,1)^2)$ with dim $V_N = N < \infty$

... more general:
consider objects in (infinite-dimensional) Hilbert space $H$ on domain $\Omega \subset \mathbb{R}^d$ with $\| \cdot \|_H$

$\boldsymbol{\Psi} := \{\psi_\lambda : \lambda \in \mathbb{I}\} \subset H$ wavelets, $\mathbb{I}$ (infinite) index set, $\lambda$ index: resolution $|\lambda|$, location $k$...

(NE) Norm equivalence: $\boldsymbol{\Psi}$ Riesz basis for $H$:

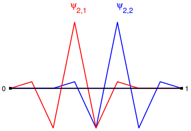
$$v \in H: \quad v = \mathbf{v}^T \boldsymbol{\Psi} := \sum_{\lambda \in \mathbb{I}} v_\lambda \psi_\lambda \quad \text{such that} \quad \|v\|_H \sim \|\mathbf{v}\|_{\ell_2(\mathbb{I})}$$

(L) Locality $\qquad \mathrm{diam}\,(\mathrm{supp}\,\psi_\lambda) \sim 2^{-|\lambda|} \qquad |\lambda|$ resolution

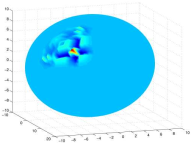
$\psi_\lambda$ centered around $2^{-|\lambda|}k$

(CP) Cancellation property (vanishing moments)

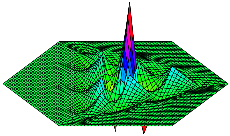
$$\langle v, \psi_\lambda \rangle \lesssim 2^{-|\lambda|(\frac{d}{2}+\tilde{m})} \|v^{(\tilde{m})}\|_{L_\infty(\mathrm{supp}\;\psi_\lambda)} \quad \text{for some } \tilde{m}$$




[Dahmen, Kunoth, Urban '99]



[Dahmen, Schneider '99], [Kunoth, Sahner '06]



[Harbrecht, Schneider '00]

Constructions of (biorthogonal spline-)wavelets on bounded domains (based on [Cohen, Daubechies, Feauveau '92])

Applications to PDEs:
(Local) wavelet transforms to detect shocks/discontinuities for hyperbolic conservation laws
. . . spectral viscosity methods . . .

Applications to PDEs:
(Local) wavelet transforms to detect shocks/discontinuties for hyperbolic conservation laws

...spectral viscosity methods ...

Wavelets as multiscale bases for numerically solving PDEs? A typical example ...

Applications to PDEs:
(Local) wavelet transforms to detect shocks/discontinuties for hyperbolic conservation laws
...spectral viscosity methods ...

Wavelets as multiscale bases for numerically solving PDEs? A typical example ...

## Elliptic PDE (Elliptic Partial Differential Equation)

Elliptic PDE of order $2r$ on domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$

| | | | | |
|---|---|---|---|---|
| $r = 1$ (Laplace operator): | $-\Delta y$ | $= f$ | in $\Omega$, | $y\|_{\partial\Omega} = 0$ |
| $r = 2$ (biharmonic operator): | $\Delta^2 y$ | $= f$ | in $\Omega$, | $y\|_{\partial\Omega} = \mathbf{n} \cdot \nabla y\|_{\partial\Omega} = 0$ |

Variational formulation $\rightsquigarrow$ Weak operator form: for given $f \in H^{-r}(\Omega)$, find $y \in H_0^r(\Omega)$ such that

$$\boxed{Ay = f \quad \text{in} \quad H^{-r}(\Omega)}$$

Elliptic operator $A$ defined by $\langle Av, w\rangle := a(v, w)$ symmetric, continuous

and coercive on $H_0^r(\Omega)$: $\boxed{\|Av\|_{H^{-r}(\Omega)} \sim \|v\|_{H^r(\Omega)} \quad \text{(mapping poperty (MP))}}$

Example: $r = 1$ (Laplace operator $\rightsquigarrow$ Dirichlet problem)

$$a(v, w) := \int_\Omega \nabla v(x) \cdot \nabla w(x)\, dx$$

# Numerical Solution on Finite-Dimensional Space — A View from Finite Elements

Discretization on uniform grid: $V_h \subset H_0^r(\Omega)$ $\dim V_h < \infty$ $\rightsquigarrow$ $\boxed{A_h\, y_h = f_h \qquad (*)}$

$0 < h < 1$ grid spacing

Goal: Realize discretization error accuracy $\varepsilon$
with minimal amount of work $\mathcal{O}(N(\varepsilon))$ in amount of unknowns $N(\varepsilon)$

## Numerical Solution on Finite-Dimensional Space — A View from Finite Elements

Discretization on uniform grid: $V_h \subset H_0^r(\Omega)$ $\quad$ $\dim V_h < \infty$ $\quad \leadsto \quad \boxed{A_h\, y_h = f_h \qquad (*)}$

$0 < h < 1$ grid spacing

Goal: Realize discretization error accuracy $\varepsilon$
with minimal amount of work $\mathcal{O}(N(\varepsilon))$ in amount of unknowns $N(\varepsilon)$

Obstructions for fast numerical solution:

- Typical: representation of $V_h$ using basis with compact support (finite element basis)

  $\leadsto$ large sparse linear system of equations $(*)$ $\leadsto$ iterative solver
- Convergence speed of iterative solver depends on $\mathrm{cond}_2(A_h)$
- Standard discretizations with finite elements $\leadsto$ $\mathrm{cond}_2(A_h) \sim h^{-2r}$

  $0 < h < 1$ grid spacing
- High desired accuracy, resolution of singularities in data and/or geometry $\leadsto$ small $h$

  $\leadsto$ larger problem $\leadsto$ worse condition number

# A-priori Estimates for Finite Elements

Quality measure: Approximation in norm $\|y - y_h\|_{L_2(\Omega)} \leq \varepsilon$

A–priori error estimates: $\Omega \subset \mathbb{R}^d$ $\dim V_h = N \sim h^{-d}$ uniform grid

$$\|y - y_h\|_{L_2(\Omega)} \lesssim h^s \, \|y\|_{H^s(\Omega)} \qquad y_h \in V_h \qquad 0 \leq s \leq p+1$$

$$\Longleftrightarrow \qquad \|y - y_N\|_{L_2(\Omega)} \lesssim N^{-s/d} \, \|y\|_{H^s(\Omega)}$$

$N$ degrees of freedom $\longleftrightarrow$ accuracy $\mathcal{O}(N^{-(p+1)/d})$

Approximation rate determined by

(i) (piecewiese polynomials of degree $p \rightsquigarrow$) approximation order $p+1$ of $V_h$

(ii) space dimension $d$

(iii) amount of smoothness of $y$ in $L_2$

Target:

Realize discretization error accuracy $\varepsilon \sim h^{p+1} \sim 2^{-(p+1)J}$ for fine grid with spacing $h \sim 2^{-J}$

Problem complexity: For $h \sim 2^{-J}$ a total of $N \sim 2^{Jd}$ unknowns

Optimal complexity for iterative solver: Minimal amount of work is $\mathcal{O}(N)$

# A-priori Estimates for Finite Elements

Quality measure: Approximation in norm $\|y - y_h\|_{L_2(\Omega)} \leq \varepsilon$

A–priori error estimates: $\Omega \subset \mathbb{R}^d$ $\dim V_h = N \sim h^{-d}$ uniform grid

$$\|y - y_h\|_{L_2(\Omega)} \lesssim h^s \|y\|_{H^s(\Omega)} \qquad y_h \in V_h \qquad 0 \leq s \leq p+1$$

$$\iff \quad \|y - y_N\|_{L_2(\Omega)} \lesssim N^{-s/d} \|y\|_{H^s(\Omega)}$$

$N$ degrees of freedom $\longleftrightarrow$ accuracy $\mathcal{O}(N^{-(p+1)/d})$

Approximation rate determined by

(i) (piecewiese polynomials of degree $p \rightsquigarrow$) approximation order $p+1$ of $V_h$

(ii) space dimension $d$

(iii) amount of smoothness of $y$ in $L_2$

Target:

Realize discretization error accuracy $\varepsilon \sim h^{p+1} \sim 2^{-(p+1)J}$ for fine grid with spacing $h \sim 2^{-J}$

Problem complexity: For $h \sim 2^{-J}$ a total of $N \sim 2^{Jd}$ unknowns

Optimal complexity for iterative solver: Minimal amount of work is $\mathcal{O}(N)$

Multilevel setting:
$V_h \longleftrightarrow V_J$ $\quad h \sim 2^{-J}$ $\quad J$ finest resolution level

Multiresolution $V_{j_0} \subset V_{j_0+1} \subset \ldots \subset V_j \subset \ldots \subset V_J \subset H_0^r(\Omega)$

# Multilevel Preconditioners

Asymptotically optimal preconditioner $C_J$ such that $\text{cond}_2(C_J A_J) \sim 1$ and setup and application of $C_J$ in optimal linear complexity $\mathcal{O}(N)$

Schwarz iterative schemes based on subspace corrections

$\rightsquigarrow$ Multilevel schemes yielding optimal preconditioners:

- Multiplicative schemes $\rightsquigarrow$ multigrid methods — Brandt, Braess, Bramble, Hackbusch ...
- Additive schemes ($\rightsquigarrow$ BPX preconditioner [Bramble, Pasciak, Xu '90])
  Wavelet discretizations/Preconditioner based on Fast Wavelet Transform
  [Jaffard '92], [Dahmen, Kunoth '92], [Oswald '92]

  Relevant idea from Approximation Theory: Multilevel characterization of function spaces;
  Isomorphism $\|Av\|_{H^{-r}(\Omega)} \sim \|v\|_{H^r(\Omega)}$ combined with norm equivalences (NE)

Ingredients for reaching goal to reach discretization accuracy in optimal complexity:

(i) Multilevel preconditioner $C_h$
multigrid methods, BPX preconditioner, wavelet discretizations $\rightsquigarrow$ $\text{cond}_2(C_J A_J) \sim 1$

(ii) Nested iteration
optimal condition of system matrix $C_j A_j$ for each $j \rightsquigarrow$ fixed amount of iterations on each level to reach discretization error accuracy on that level;
spaces nested and $N_j \sim 2^{dj}$ and geometric series argument

## Multilevel Preconditioners

Asymptotically optimal preconditioner $C_J$ such that $\text{cond}_2(C_J A_J) \sim 1$ and setup and application of $C_J$ in optimal linear complexity $\mathcal{O}(N)$

Schwarz iterative schemes based on subspace corrections

$\rightsquigarrow$ Multilevel schemes yielding optimal preconditioners:

- Multiplicative schemes $\rightsquigarrow$ multigrid methods — Brandt, Braess, Bramble, Hackbusch ...
- Additive schemes ($\rightsquigarrow$ BPX preconditioner [Bramble, Pasciak, Xu '90])
  Wavelet discretizations/Preconditioner based on Fast Wavelet Transform
  [Jaffard '92], [Dahmen, Kunoth '92], [Oswald '92]

  Relevant idea from Approximation Theory: Multilevel characterization of function spaces;
  Isomorphism $\|Av\|_{H^{-r}(\Omega)} \sim \|v\|_{H^r(\Omega)}$ combined with norm equivalences (NE)

Ingredients for reaching goal to reach discretization accuracy in optimal complexity:

(i) Multilevel preconditioner $C_h$
multigrid methods, BPX preconditioner, wavelet discretizations $\rightsquigarrow$ $\text{cond}_2(C_J A_J) \sim 1$

(ii) Nested iteration
optimal condition of system matrix $C_j A_j$ for each $j \rightsquigarrow$ fixed amount of iterations on each level to reach discretization error accuracy on that level;
spaces nested and $N_j \sim 2^{dj}$ and geometric series argument

Beyond point of view of finite elements:
wavelets can represent infinite-dimensional and implicitly given objects ...
$\rightsquigarrow$ (elliptic) PDEs well-conditioned in (properly scaled) wavelet bases
and allow for adaptivity for non-smooth solutions ...

## Paradigm of Adaptive Wavelet Method for One Stationary (Elliptic) PDE

[Cohen, Dahmen, DeVore '01/'02]

(i) Well–posed variational problem: given $f \in \mathcal{Q}'$, $B : \mathcal{Y} \to \mathcal{Q}'$, find $y \in \mathcal{Y}$ such that $\boxed{By = f}$

$$\boxed{\text{(MP)} \qquad \|Bw\|_{\mathcal{Q}'} \sim \|w\|_{\mathcal{Y}} \qquad \text{for all } w \in \mathcal{Y} \qquad \text{mapping property}}$$

(ii) $\Psi^{\mathcal{Y}}, \Psi^{\mathcal{Q}}$ wavelet bases for $\mathcal{Y}, \mathcal{Q}$ :

$$\boxed{\text{(NE)} \qquad \|\mathbf{w}^T\Psi^{\mathcal{Y}}\|_{\mathcal{Y}} \sim \|\mathbf{w}\|_{\ell_2} \qquad \text{for all } \mathbf{w} = (w_\lambda)_{\lambda\in\mathbb{I}} \in \ell_2}$$

$$\mathbf{Bw} := (\langle \psi_\lambda^{\mathcal{Y}}, Bw\rangle)_{\lambda\in\mathbb{I}} \qquad \mathbf{f} := (\langle \psi_\lambda^{\mathcal{Y}}, f\rangle)_{\lambda\in\mathbb{I}}$$

$\rightsquigarrow$ **Theorem** $\quad By = f \iff \mathbf{By} = \mathbf{f}$ well-posed in $\ell_2$ $\qquad (\mathbf{B} : \ell_2 \to \ell_2)$

$$\text{(MP)} + \text{(NE)} \iff \qquad \|\mathbf{Bw}\|_{\ell_2} \sim \|\mathbf{w}\|_{\ell_2} \quad \text{for all } \mathbf{w} \in \ell_2$$

(iii) Practical solution schemes for $\mathbf{By} = \mathbf{f}$:

(A) Perturbed Richardson iteration (for symmetric $\mathbf{B}$):

(A.1) $\mathbf{y}^{n+1} = \mathbf{y}^n + (\mathbf{f} - \mathbf{By}^n) \quad n = 0, 1, 2, \ldots \qquad \|\mathbf{y}^{n+1} - \mathbf{y}\|_{\ell_2} \leq \rho\, \|\mathbf{y}^n - \mathbf{y}\|_{\ell_2} \quad \rho < 1$

(A.2) Approximate realization: adaptive evaluation of $\mathbf{By}^n$ in $\textsc{Solve}\,[\varepsilon, \mathbf{B}, \mathbf{f}] \to \mathbf{y}_\varepsilon$

(A.3) Coarsening (thresholding) of the iterands (for complexity)

(B) Adaptive wavelet Galerkin method and bulk chasing strategy

## Extension to a Single Parabolic Evolution PDE in Space-Time Variational Form

[Ladyshenskaya et al. 1967], [Wloka '82], [Dautray, Lions '92], [Schwab, Stevenson '09] ...

(i) Variational space-time form of (PDE)

$$\begin{aligned} y'(t) + A(t)\,y(t) &= f(t) \qquad \text{a.e. } t \in I \\ y(0) &= y_0 \end{aligned}$$

solution space: Lebesgue-Bochner space $\mathcal{Y} := (L_2(I) \otimes Y) \cap (H^1(I) \otimes Y')$

with norm $\|w\|_{\mathcal{Y}}^2 := \|w\|_{L_2(I)\otimes Y}^2 + \|w'\|_{H^1(I)\otimes Y'}^2$

test space $\quad \mathcal{Q} := L_2(I; Y) \times L_2(\Omega) \quad$ with norm $\quad \|v\|_{\mathcal{Q}}^2 := \|v_1\|_{L_2(I)\otimes Y}^2 + \|v_2\|_{L_2(\Omega)}^2$

bilinear form $b(\cdot,\cdot) : \mathcal{Y} \times \mathcal{Q} \to \mathbb{R}$

$$b(y,(v_1,v_2)) := \int_I \left[\langle y'(t,\cdot), v_1(t,\cdot)\rangle + \langle A(t)y(t,\cdot), v_1(t,\cdot)\rangle\right] dt + \langle y(0,\cdot), v_2\rangle =: \langle By, v\rangle$$

right hand side

$$\langle f, v\rangle := \int_I \langle f(t,\cdot), v_1(t,\cdot)\rangle\, dt + \langle y_0, v_2\rangle$$

(PDE) $\quad \rightsquigarrow$ given $f \in \mathcal{Q}'$, find $y \in \mathcal{Y}$: $\quad \boxed{By = f}$

| | | | | |
|---|---|---|---|---|
| Theorem (MP) | $\|Bw\|_{\mathcal{Q}'} \sim \|w\|_{\mathcal{Y}}$ | for all $w \in \mathcal{Y}$ | mapping property | |

(ii) $\Psi^{\mathcal{Y}}, \Psi^{\mathcal{Q}}$ wavelet bases for $\mathcal{Y}, \mathcal{Q}$ $\quad \rightsquigarrow \quad$ $\mathbf{By} := (\langle \psi_\lambda^{\mathcal{Q}}, By\rangle)_{\lambda\in\mathbb{I}} \quad \mathbf{f} := (\langle \psi_\lambda^{\mathcal{Q}}, f\rangle)_{\lambda\in\mathbb{I}}$

**Theorem** $\quad By = f \iff$ $\mathbf{By} = \mathbf{f}$ $\quad \mathbf{B} : \ell_2 \to \ell_2$ and $\quad \mathbf{By} = \mathbf{f}$ well-posed in $\ell_2$

(MP) + (NE) $\Longrightarrow$ $\quad \|\mathbf{Bv}\|_{\ell_2} \sim \|\mathbf{v}\|_{\ell_2}, \quad \mathbf{v} \in \ell_2 \qquad$ **B** unsymmetric

## Complexity Analysis

Based on benchmark:
decay rate $s$ for (wavelet-)best $N$ term approximation $\mathcal{A}^s := \{\mathbf{v} \in \ell_2 : \|\mathbf{v} - \mathbf{v}_N\| \lesssim N^{-s}\}$

Work/accuracy balance of best $N$ term approximation:

$$\boxed{\text{Target accuracy } \varepsilon \ (\sim N^{-s}) \quad \longleftrightarrow \quad \text{Work } \varepsilon^{-1/s} \ (\sim N)}$$

## Convergence and Complexity

For solution routine (A): (Idealized) iteration (for symmetric $\mathbf{B}$)
$\mathbf{v}^{n+1} = \mathbf{v}^n + (\mathbf{f} - \mathbf{B}\mathbf{v}^n)$ update via $\textsc{Res}\,[\eta, \mathbf{B}, \mathbf{f}, \mathbf{v}] \to \mathbf{r}_\eta$ $\rightsquigarrow$ $\textsc{Solve}\,[\varepsilon, \mathbf{B}, \mathbf{f}] \to \mathbf{v}_\varepsilon$

**Benchmark Theorem** [Cohen, Dahmen, DeVore '01/'02]

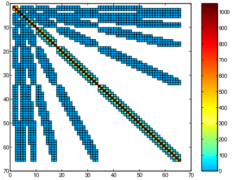
Vanishing moments (CP) for wavelets $\Longrightarrow$ $\mathbf{B}$ is $s^*$–compressible

$\Longrightarrow$ for variational problem satisfying (MP) scheme $\textsc{Solve}$ can be designed with properties:

(I) For every target accuracy $\varepsilon > 0$ $\textsc{Solve}$ produces after finitely many steps approximate solution $\mathbf{v}_\varepsilon$ such that $\|\mathbf{v} - \mathbf{v}_\varepsilon\| \le \varepsilon$

(II) Exact solution $\mathbf{v} \in \mathcal{A}^s \Longrightarrow \operatorname{supp} \mathbf{v}_\varepsilon$, $\#$ flops $\sim \varepsilon^{-1/s} \sim N$

# Core Ingredient of SOLVE: Compressible Operators

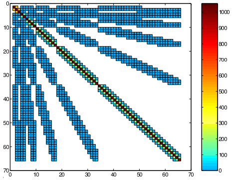



(CP) $\rightsquigarrow$ $\mathbf{B}$ is $s^*$–compressible:
for every $s \in (0, s^*)$ there exists $\mathbf{B}_j$ with $\leq \alpha_j 2^j$ nonzero entries per row and column s.th. for $j \in \mathbb{N}_0$

$$\|\mathbf{B} - \mathbf{B}_j\| \leq \alpha_j 2^{-sj}, \quad \sum_{j \in \mathbb{N}_0} \alpha_j < \infty$$

($\mathbf{B}$ 'close to' sparse matrix)

## Core Ingredient of SOLVE: Compressible Operators



(CP) $\rightsquigarrow$ **B** is $s^*$–compressible:
for every $s \in (0, s^*)$ there exists $\mathbf{B}_j$ with $\leq \alpha_j 2^j$ nonzero entries per row and column s.th. for $j \in \mathbb{N}_0$

$$\|\mathbf{B} - \mathbf{B}_j\| \leq \alpha_j 2^{-sj}, \quad \sum_{j \in \mathbb{N}_0} \alpha_j < \infty$$

(**B** 'close to' sparse matrix)

## Application of (Non)Linear Operators in Wavelet Bases

Theory: [Dahmen, Schneider, Xu '00], [Cohen, Dahmen, DeVore '03] ...
Implementation with isotropic tensor-product wavelets: $d = 2$: [Vorloeper '10] general $d$: [Stapel '11], [Mollet, Pabel '12], [Pabel '15]

Input: finitely supported vector $\mathbf{v} = (v_\mu)_{\mu \in \Lambda}$ $\Lambda \subset \mathbb{I}$ finite

Output: approximation of $\mathbf{Bv}$ with infinite-dimensional operator $\mathbf{B} : \ell_2(\mathbb{I}) \to \ell_2(\mathbb{I})$

$B : \mathcal{Y} \to \mathcal{Q}' \rightsquigarrow$ expand $Bv \in \mathcal{Q}'$ in dual wavelet basis for $\mathcal{Q}'$ and $v$ in primal wavelet basis for $\mathcal{Y}$
$\rightsquigarrow$

$$Bv = (\mathbf{Bv})^T \tilde{\Psi} = \sum_{\lambda \in \mathbb{I}} \langle Bv, \psi_\lambda \rangle \, \tilde{\psi}_\lambda = \sum_{\lambda \in \mathbb{I}} \langle B(\sum_{\mu \in \Lambda} v_\mu \psi_\mu, \psi_\lambda \rangle) \, \tilde{\psi}_\lambda = \sum_{\lambda \in \mathbb{I}} \sum_{\mu \in \Lambda} v_\mu \langle B\psi_\mu, \psi_\lambda \rangle \, \tilde{\psi}_\lambda$$

$\rightsquigarrow$ compute $\langle B\psi_\mu, \psi_\lambda \rangle$ for given $\mu \in \Lambda$ (finite) and all $\lambda \in \mathbb{I}$

Compressibility of $B$: $|\langle B\psi_\mu, \psi_\lambda \rangle| \leq C_{\|\mathbf{v}\|} \sup_{\mu:\ S_\lambda \cap S_\mu \neq \emptyset} 2^{-\gamma(|\lambda| - |\mu|)} \, |v_\mu|$ $\quad \gamma > \frac{d}{2} + 1$
follows from wavelet property (CP)

Essential data structure (for nonlinear operators): tree-type index sets
input $\mathbf{v} \rightsquigarrow$ prediction of tree index set based on supp $\mathbf{v}$ and properties of $\mathbf{B}$
$\rightsquigarrow$ computation of $(\mathbf{Bv})_\lambda$ after transformation to piecewise polynomials
$\rightsquigarrow$ application of $\mathbf{B}$ in optimal linear complexity

# Numerical Example for One Parabolic Linear PDE

[Chegini, Stevenson '11], [Stapel '11]

Compute $y = y(t,x)$ such that

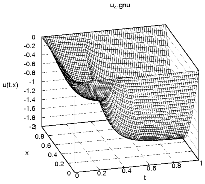
$$\begin{aligned} y_t(t,x) - y_{xx}(t,x) &= g(t) \otimes (-\pi^2)\sin(\pi x) && \text{in } I \times \Omega := (0,1)^2 \\ y(t,0) &= y(t,1) = 0 && \text{for } t \geq 0 \\ y(0,x) &= 0 && \text{for } x \in (0,1) \end{aligned}$$

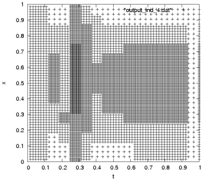
and $g(t) := \begin{cases} 1 & t \in [0, \frac{1}{3}) \\ 2 & t \in [\frac{1}{3}, 1] \end{cases}$

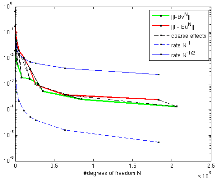
Problem formulation and implementation:

- Modified problem with zero initial conditions $\rightsquigarrow$
  solution space $\mathcal{Y} = (L_2(I) \otimes H^1(\Omega)) \cap (H^1_{(0}(I) \otimes H^{-1}(\Omega))$ and test space $\mathcal{Q} = L_2(I) \otimes Y$
- Inhomogeneous initial data: homogenization of initial conditions $\rightsquigarrow$ modification of r.h.s.
- Implementation based on AWM Toolbox by [Vorloeper '10]
  biorthogonal isotropic wavelets of order $m = 2$, $\tilde{m} = 4$
- Iterative solution by GMRES

## Plot of Solution, Refined Grid and Residual Error Reduction



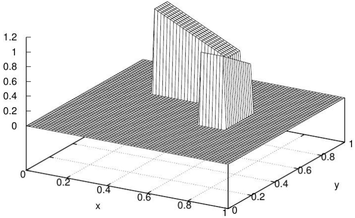

8526 degrees of freedom Expected rate in $H^1$ (isotropic wavelets): $1/2$ red: after coarsening

# Application of Nonlinear Operator in Wavelet Bases: Numerical Example

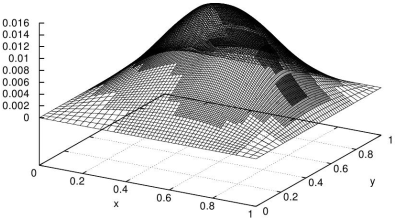
[Mollet, Pabel '12], [Pabel '15]

PDE with nonlinear term

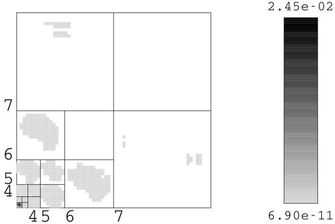
$$\begin{aligned} -\Delta y + y^3 &= f && \text{in } \Omega := (0,1)^2 \\ y &= 0 && \text{on } \partial\Omega \end{aligned}$$



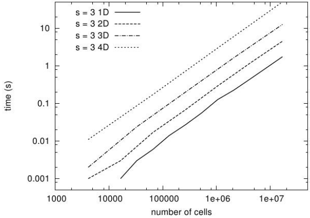

right hand side $f$




solution $y$ (with Richardson scheme and residual error bound $10^{-3}$)




distribution of 7177 active wavelet coefficients




Runtime (seconds) for evaluating $y^3$ for $d \leq 4$

## Application: Optimal Control Problem Constrained by a Parabolic PDE

Given $y_*(t,\cdot)$ $\quad f \quad$ $\omega > 0$ $\quad$ end time $T > 0$ $\quad$ initial condition $y_0$

$$
\begin{array}{llcll}
\text{minimize} & \mathcal{J}(y,u) & = & \frac{1}{2}\int_0^T \|y(t,\cdot) - y_*(t,\cdot)\|_Z^2\, dt + \frac{\omega}{2}\int_0^T \|u(t,\cdot)\|_U^2\, dt & \\
\text{subject to} & y'(t) + A(t)y(t) & = & f(t) + u(t) \qquad \text{a.e. } t \in (0,T) =: I & \text{(PDE)} \\
 & y(0) & = & y_0 &
\end{array}
$$

$y' := \frac{\partial}{\partial t} y \qquad y = y(t,x)$ state $\qquad u = u(t,x)$ control

$Y = H_0^1(\Omega)$ state space $\quad Z = Y = H_0^1(\Omega)$ observation space $\quad U = Y' = H^{-1}(\Omega)$ control space

$A(t) : Y \to Y' \qquad \langle A(t)v(t,\cdot), w(t,\cdot)\rangle := \int_\Omega [\nabla v(t,x) \cdot \nabla w(t,x) + v(t,x)w(t,x)]\, dx$ on $\Omega \subset \mathbb{R}^d$

$A(t)$ 2nd order linear selfadjoint coercive & continuous operator on $Y$

## Application: Optimal Control Problem Constrained by a Parabolic PDE

Given $y_*(t,\cdot)$ $\quad f \quad$ $\omega > 0$ $\quad$ end time $T > 0$ $\quad$ initial condition $y_0$

$$\begin{array}{rrcll} \text{minimize} & \mathcal{J}(y,u) & = & \frac{1}{2}\int_0^T \|y(t,\cdot) - y_*(t,\cdot)\|_Z^2 \, dt + \frac{\omega}{2}\int_0^T \|u(t,\cdot)\|_U^2 \, dt & \\ \text{subject to} & y'(t) + A(t)y(t) & = & f(t) + u(t) \qquad \text{a.e. } t \in (0,T) =: I & \text{(PDE)} \\ & y(0) & = & y_0 & \end{array}$$

$y' := \frac{\partial}{\partial t} y \qquad y = y(t,x)$ state $\qquad u = u(t,x)$ control

$Y = H_0^1(\Omega)$ state space $\quad Z = Y = H_0^1(\Omega)$ observation space $\quad U = Y' = H^{-1}(\Omega)$ control space

$A(t) : Y \to Y' \qquad \langle A(t)v(t,\cdot), w(t,\cdot)\rangle := \int_\Omega [\nabla v(t,x) \cdot \nabla w(t,x) + v(t,x)w(t,x)]\, dx$ on $\Omega \subset \mathbb{R}^d$

$A(t)$ 2nd order linear selfadjoint coercive & continuous operator on $Y$

PDE-constrained control problem $\rightsquigarrow$ requires repeated solution of (PDE)

$$\begin{aligned} y'(t) + A(t)y(t) &= f(t) + u(t) \\ y(0) &= y_0 \end{aligned}$$

$\rightsquigarrow$ requires fast solver as core ingredient

Conventional time discretizations (e.g., Crank-Nicolson method) $\rightsquigarrow$
requires fast solver for elliptic PDE

# Necessary and Sufficient Conditions for Optimality

Optimal control problem constrained by parabolic PDE

$\rightsquigarrow$ System of parabolic PDEs coupled globally in time (and space)

$$\begin{aligned}
y'(t) + A(t)\, y(t) &= f(t) + u(t) && \text{a.e. } t \in I \\
y(0) &= y_0 \\
\omega\, \tilde{R}^{-1} u(t) + p(t) &= 0 && \text{a.e. } t \in I \\
-p'(t) + A(t)^T p(t) &= \tilde{R}\,(y_*(t) - y(t)) && \text{a.e. } t \in I \\
p(T) &= 0
\end{aligned}$$

Riesz operator $\tilde{R}$ defined by $\langle v, \tilde{R}w \rangle_{Y \times Y'} := (v, w)_Y$ for all $v, w \in Y$

## Necessary and Sufficient Conditions for Optimality

Optimal control problem constrained by parabolic PDE

$\rightsquigarrow$ System of parabolic PDEs coupled globally in time (and space)

$$
\begin{aligned}
y'(t) + A(t)\, y(t) &= f(t) + u(t) && \text{a.e. } t \in I \\
y(0) &= y_0 \\
\omega\, \tilde{R}^{-1} u(t) + p(t) &= 0 && \text{a.e. } t \in I \\
-p'(t) + A(t)^T p(t) &= \tilde{R}\,(y_*(t) - y(t)) && \text{a.e. } t \in I \\
p(T) &= 0
\end{aligned}
$$

Riesz operator $\tilde{R}$ defined by $\langle v, \tilde{R}w\rangle_{Y \times Y'} := (v, w)_Y$ for all $v, w \in Y$

Obstructions for numerical solution:

- convential time discretizations: time-marching methods
  $\rightsquigarrow$ need storage of $y(t_i), u(t_i), p(t_i)$ for all discrete times $0 = t_0, \ldots, T = t_N$
- in each time step: solve elliptic PDE $\rightsquigarrow$ large linear system of equations
  $\rightsquigarrow$ iterative solver $\rightsquigarrow$ need preconditioning in (conjugate) gradient method
- singularities in data/domain: adaptive (FE) mesh(es) for $y(t_i), u(t_i), p(t_i)$ for all $t_i$
  one mesh for all variables, refinement/coarsening ? ... [Oeltz '06], [Meidner, Vexler '07], ...
  convergence ? complexity ??

# Necessary and Sufficient Conditions for Optimality

Optimal control problem constrained by parabolic PDE

$\rightsquigarrow$ System of parabolic PDEs coupled globally in time (and space)

$$\begin{aligned}
y'(t) + A(t)\,y(t) &= f(t) + u(t) && \text{a.e. } t \in I \\
y(0) &= y_0 && \\
\omega\,\tilde{R}^{-1} u(t) + p(t) &= 0 && \text{a.e. } t \in I \\
-p'(t) + A(t)^T p(t) &= \tilde{R}\,(y_*(t) - y(t)) && \text{a.e. } t \in I \\
p(T) &= 0 &&
\end{aligned}$$

Riesz operator $\tilde{R}$ defined by $\langle v, \tilde{R}w\rangle_{Y\times Y'} := (v,w)_Y$ for all $v, w \in Y$

Obstructions for numerical solution:

- convential time discretizations: time-marching methods
  $\rightsquigarrow$ need storage of $y(t_i), u(t_i), p(t_i)$ for all discrete times $0 = t_0, \ldots, T = t_N$
- in each time step: solve elliptic PDE $\rightsquigarrow$ large linear system of equations
  $\rightsquigarrow$ iterative solver $\rightsquigarrow$ need preconditioning in (conjugate) gradient method
- singularities in data/domain: adaptive (FE) mesh(es) for $y(t_i), u(t_i), p(t_i)$ for all $t_i$
  one mesh for all variables, refinement/coarsening ? ... [Oeltz '06], [Meidner, Vexler '07], ...
  convergence ? complexity ??

---

Solution Ansatz here: full weak space-time form of parabolic PDE constraint and setup of control problem in (infinite) wavelet coordinates

# PDE-Constrained Optimal Control Problem in Wavelet Coordinates

minimize $\quad \mathbf{J}(\mathbf{y},\mathbf{u}) = \frac{1}{2}\|\mathbf{R}^{1/2}(\mathbf{y}-\mathbf{y}_*)\|^2 + \frac{\omega}{2}\|\mathbf{R}^{-1/2}\mathbf{u}\|^2$

subject to $\quad \mathbf{B}\mathbf{y} = \mathbf{f} + \mathbf{u}$ $\qquad \mathbf{B} : \ell_2 \to \ell_2$ automorphism $\qquad \|\cdot\| := \|\cdot\|_{\ell_2}$

## Necessary and Sufficient Conditions — Karush-Kuhn-Tucker (KKT) system

$$\mathbf{L}(\mathbf{y},\mathbf{u},\mathbf{p}) := \mathbf{J}(\mathbf{y},\mathbf{u}) + \langle \mathbf{p}, \mathbf{B}\mathbf{y} - (\mathbf{f}+\mathbf{u})\rangle$$

$$\delta\mathbf{L} = 0 \rightsquigarrow \boxed{\begin{aligned} \mathbf{B}\mathbf{y} &= \mathbf{f} + \mathbf{u} \\ \omega\mathbf{R}^{-1}\mathbf{u} &= \mathbf{p} \\ \mathbf{B}^*\mathbf{p} &= \mathbf{R}(\mathbf{y}_* - \mathbf{y}) \end{aligned}} \iff \boxed{\mathbf{Q}\,\mathbf{u} = \mathbf{g}}$$

$$\iff \boxed{\begin{pmatrix} \mathbf{R} & \mathbf{0} & \mathbf{B}^* \\ \mathbf{0} & \omega\mathbf{R}^{-1} & -\mathbf{E} \\ \mathbf{B} & -\mathbf{E} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{R}\mathbf{y}_* \\ \mathbf{0} \\ \mathbf{f} \end{pmatrix} \quad \text{(SPP)}} \qquad \mathbf{Q} : \ell_2 \to \ell_2 \text{ automorphism}$$

where
$$\mathbf{Q} := \mathbf{B}^{-*}\mathbf{R}\mathbf{B}^{-1} + \omega\mathbf{R}^{-1}$$
$$\mathbf{g} := \mathbf{B}^{-*}(\mathbf{R}\mathbf{y}_* - \mathbf{R}\mathbf{B}^{-1}\mathbf{f})$$

## Convergence and Complexity Analysis for Control Problem with Elliptic or Parabolic PDE Constraints

Essential ideas: $\textsc{Res}$ for $\textsc{Solve}\,[\ldots,\mathbf{Q},\ldots]$ reduced to $\textsc{Res}$ for $\textsc{Solve}\,[\ldots,\mathbf{B},\ldots]$

applied to normal equations

and KKT system $\longleftrightarrow$ condensed system $\mathbf{Q}\mathbf{u} = \mathbf{g}$

**Theorem** [Dahmen, Kunoth '05], [Gunzburger, Kunoth '11]

For any target accuracy $\varepsilon > 0$ $\textsc{Solve}\,[\varepsilon,\mathbf{Q},\mathbf{g}] \to \mathbf{u}_\varepsilon$ converges in finitely many steps

$$\|\mathbf{u}-\mathbf{u}_\varepsilon\| \le \varepsilon \quad \|\mathbf{y}-\mathbf{y}_\varepsilon\| \lesssim \varepsilon \quad \|\mathbf{p}-\mathbf{p}_\varepsilon\| \lesssim \varepsilon \quad \mathbf{u}_\varepsilon, \mathbf{y}_\varepsilon, \mathbf{p}_\varepsilon \text{ finitely supported}$$

$\mathbf{u}, \mathbf{y}, \mathbf{p} \in \mathcal{A}^s \Longrightarrow$

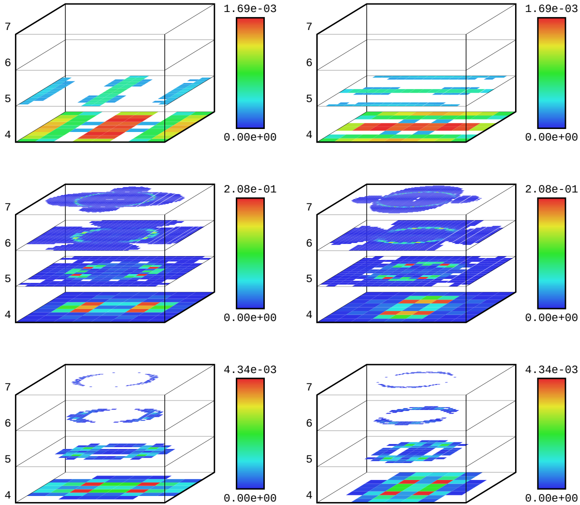
$$(\#\,\mathrm{supp}\,\mathbf{u}_\varepsilon) + (\#\,\mathrm{supp}\,\mathbf{y}_\varepsilon) + (\#\,\mathrm{supp}\,\mathbf{p}_\varepsilon) \lesssim \varepsilon^{-1/s}\left(\|\mathbf{u}\|_{\mathcal{A}^s}^{1/s} + \|\mathbf{y}\|_{\mathcal{A}^s}^{1/s} + \|\mathbf{p}\|_{\mathcal{A}^s}^{1/s}\right)$$

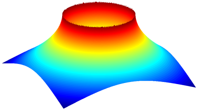
$$\|\mathbf{u}_\varepsilon\|_{\mathcal{A}^s} + \|\mathbf{y}_\varepsilon\|_{\mathcal{A}^s} + \|\mathbf{p}_\varepsilon\|_{\mathcal{A}^s} \lesssim \|\mathbf{u}\|_{\mathcal{A}^s} + \|\mathbf{y}\|_{\mathcal{A}^s} + \|\mathbf{p}\|_{\mathcal{A}^s}$$

$\#\text{flops} \sim \varepsilon^{-1/s}$

# Numerical Example for Elliptic Control Problem (2D)

type $\mathbf{e} = (1, 0)$   type $\mathbf{e} = (0, 1)$



target state $y_*$

[Burstedde '05]

## Summary: PDE-Constrained Control Problems

- Control problem constrained by parabolic PDE
  Full weak space-time formulation of evolution PDE
  $\rightsquigarrow$ saddle point system of PDEs coupled globally in time and space
- For smooth solutions: multilevel/wavelet preconditioners + nested iteration
  $\rightsquigarrow$ numerical solution scheme with optimal complexity
- For non-smooth solutions:
  proofs of convergence and optimal complexity based on adaptive wavelets

## Beyond Wavelets

- Optimal preconditioning: multilevel and multigrid methods (for normal equations);
  fast iterative solvers on (non)uniform grids
- (A posteriori) error estimates for PDE constrained control problems [Liu et al . . . et al . . . ]
- Convergence theory of adaptive (finite element) method for control problem
  with linear elliptic PDE constraints ?
  One or different meshes for all variables ? Refinement / coarsening of meshes ?
- Complexity estimates ? Optimal complexity ? Application of PDE operator ?
- Convergence theory of adaptive (FE/DG) methods for control problems
  constrained by linear evolutionary PDE ?