

Computational Information Games

Houman Owhadi

TexAMP 2015



Main Question

Can we turn the process of discovery of a scalable numerical method into a UQ problem and, to some degree, solve it as such in an automated fashion?

Can we use a computer, not only to implement a numerical method but also to find the method itself?

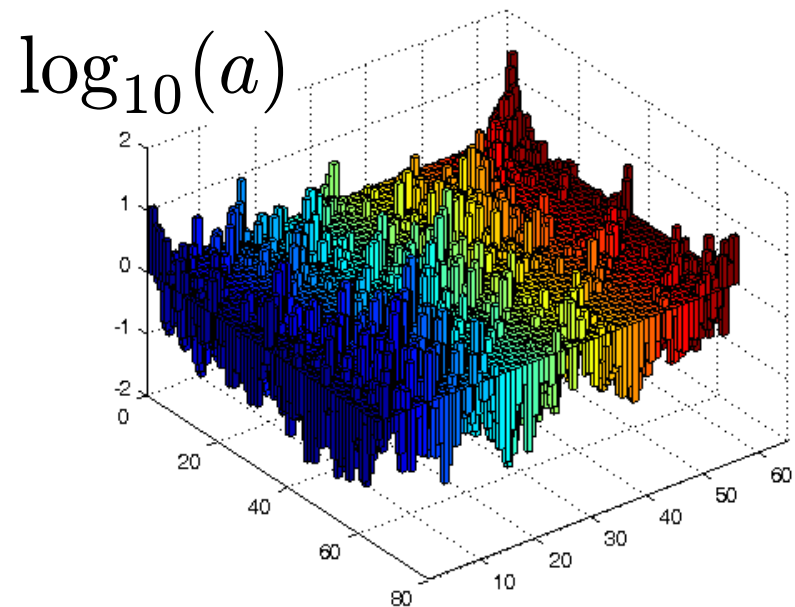
**Problem: Find a method for solving (1)
as fast as possible to a given accuracy**

$$(1) \quad \begin{cases} -\operatorname{div}(a \nabla u) = g, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

$\Omega \subset \mathbb{R}^d$ $\partial\Omega$ is piec. Lip.

a unif. ell.

$a_{i,j} \in L^\infty(\Omega)$



Multigrid Methods

Multigrid: [Fedorenko, 1961, Brandt, 1973, Hackbusch, 1978]

Multiresolution/Wavelet based methods

[Brewster and Beylkin, 1995, Beylkin and Coult, 1998, Averbuch et al., 1998]

\mathbb{R}^m

- **Linear complexity with smooth coefficients**

Problem Severely affected by lack of smoothness

Robust/Algebraic multigrid

[Mandel et al., 1999, Wan-Chan-Smith, 1999, Xu and Zikatanov, 2004, Xu and Zhu, 2008], [Ruge-Stüben, 1987]
[Panayot - 2010]

Stabilized Hierarchical bases, Multilevel preconditioners

[Vassilevski - Wang, 1997, 1998]

[Panayot - Vassilevski, 1997]

[Chow - Vassilevski, 2003]

[Aksoylu- Holst, 2010]

- Some degree of robustness but problem remains open with rough coefficients

Why? Interpolation operators are unknown

Don't know how to bridge scales with rough coefficients!

Low Rank Matrix Decomposition methods

Fast Multipole Method: [Greengard and Rokhlin, 1987]

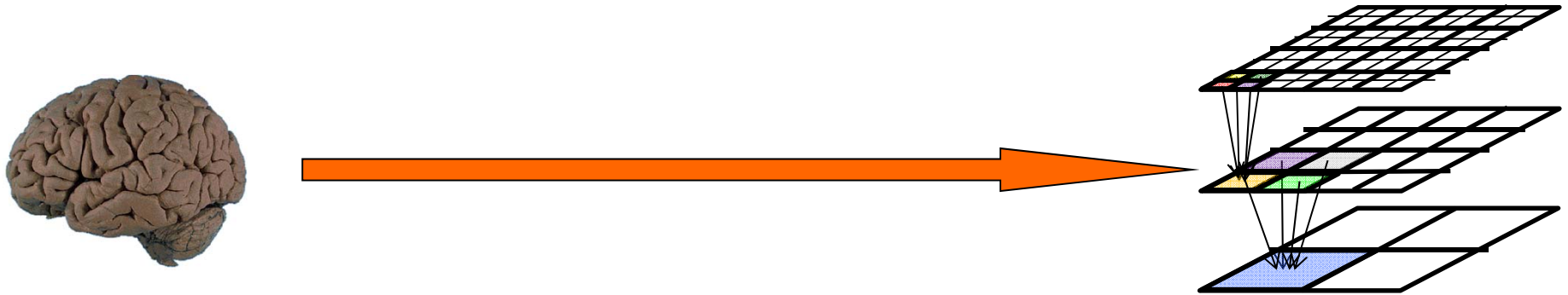
Hierarchical Matrix Method: [Hackbusch et al., 2002]

[Bebendorf, 2008]:

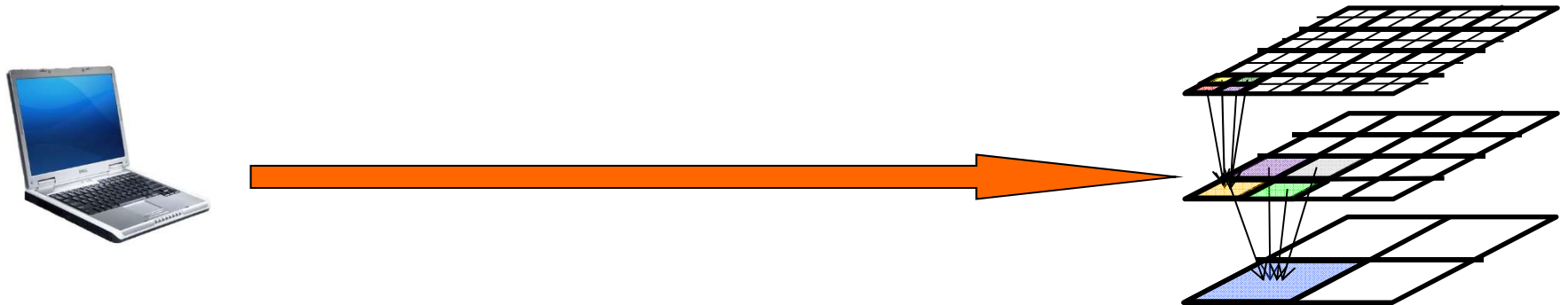
$N \ln^{d+3} N$ complexity

Common theme between these methods

Their process of discovery is based on intuition, brilliant insight, and guesswork



Can we turn this process of discovery into an algorithm?



Answer: YES

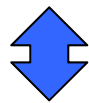
Compute fast



**Play adversarial
Information game**



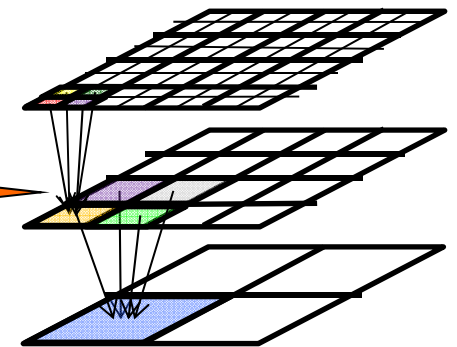
**Compute with
partial information**



Identify game



**Find optimal
strategy**



[Owhadi 2015, Multi-grid with rough coefficients and Multiresolution PDE decomposition from Hierarchical Information Games, arXiv:1503.03467]

Resulting method:

$N \ln^2 N$ complexity

This is a theorem

Resulting method:

$$\begin{cases} -\operatorname{div}(a\nabla u) = g \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases}$$

$$H_0^1(\Omega) = \mathfrak{W}^{(1)} \oplus_a \mathfrak{W}^{(2)} \oplus_a \cdots \oplus_a \mathfrak{W}^{(k)} \oplus_a \cdots$$

$$\langle \psi, \chi \rangle_a := \int_{\Omega} (\nabla \psi)^T a \nabla \chi = 0 \text{ for } (\psi, \chi) \in \mathfrak{W}^{(i)} \times \mathfrak{W}^{(j)}, i \neq j$$

Theorem For $v \in \mathfrak{W}^{(k)}$

$$\frac{C_1}{2^k} \leq \frac{\|v\|_a}{\|\operatorname{div}(a\nabla v)\|_{L^2(\Omega)}} \leq \frac{C_2}{2^k}$$

$$\|v\|_a^2 := \langle v, v \rangle_a = \int_{\Omega} (\nabla v)^T a \nabla v$$

Looks like an eigenspace decomposition

$$u = w^{(1)} + w^{(2)} + \dots + w^{(k)} + \dots$$

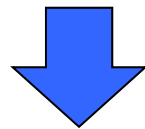
$w^{(k)}$ = F.E. sol. of PDE in $\mathfrak{W}^{(k)}$

Can be computed independently

$B^{(k)}$: Stiffness matrix of PDE in $\mathfrak{W}^{(k)}$

Theorem

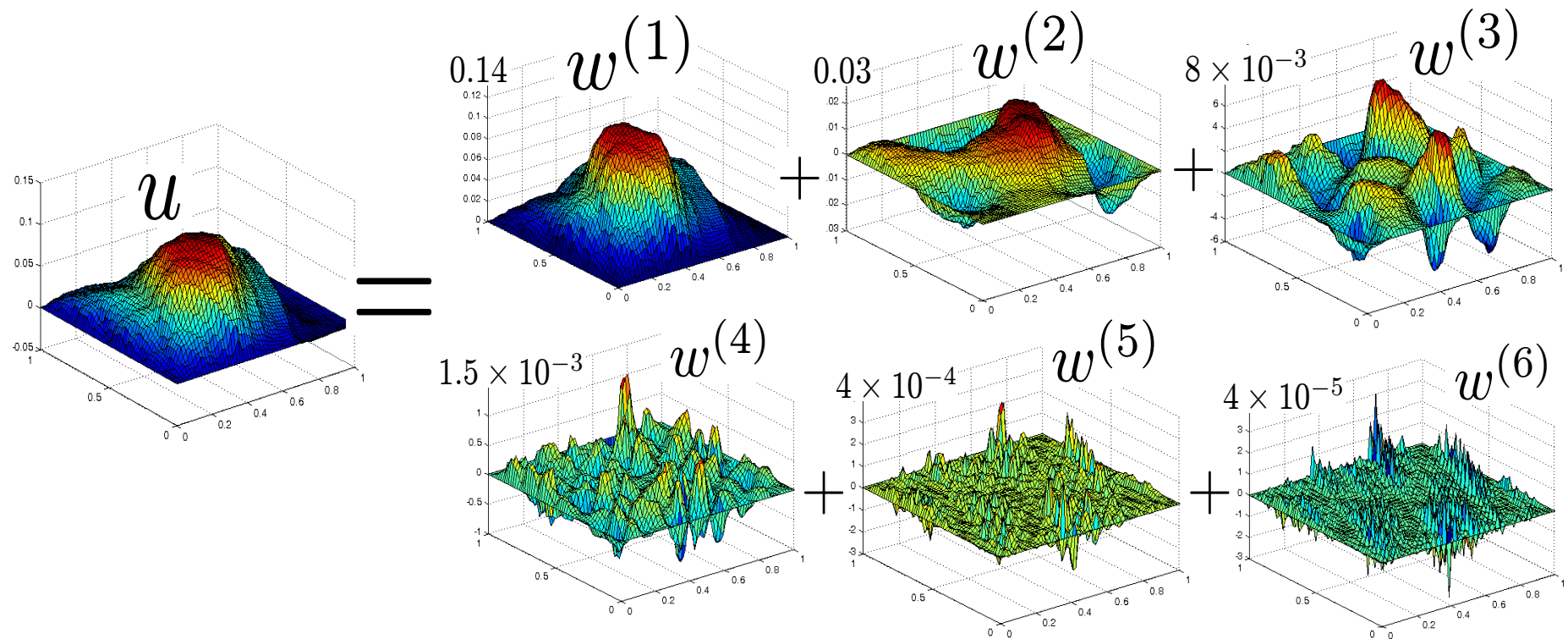
$$\frac{\lambda_{\max}(B^{(k)})}{\lambda_{\min}(B^{(k)})} \leq C$$



Just relax in $\mathfrak{W}^{(k)}$ to find $w^{(k)}$

Quacks like an eigenspace decomposition

Multiresolution decomposition of solution space



Solve time-discretized wave equation (implicit time steps)
with rough coefficients in $\mathcal{O}(N \ln^2 N)$ -complexity

Swims like an eigenspace decomposition

\mathfrak{W} : F.E. space of $H_0^1(\Omega)$ of dim. N

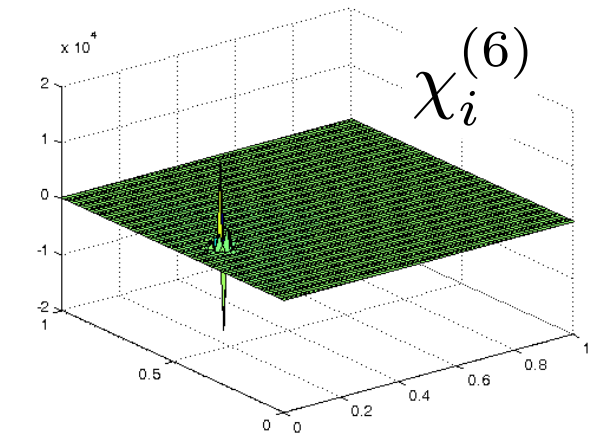
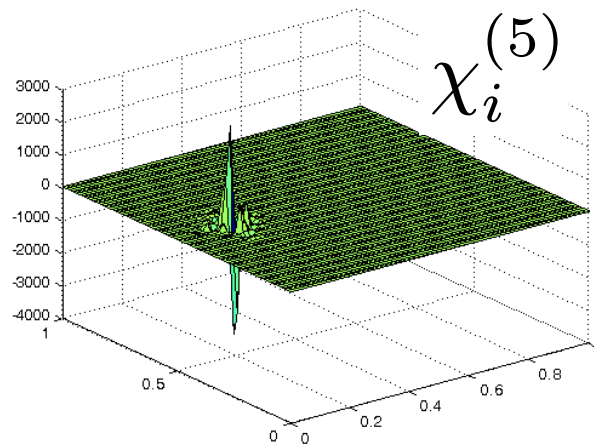
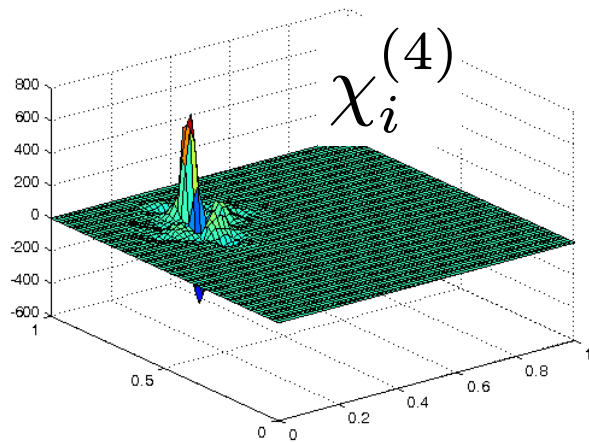
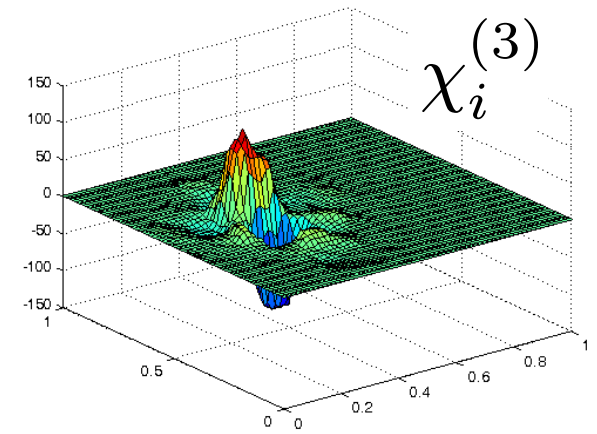
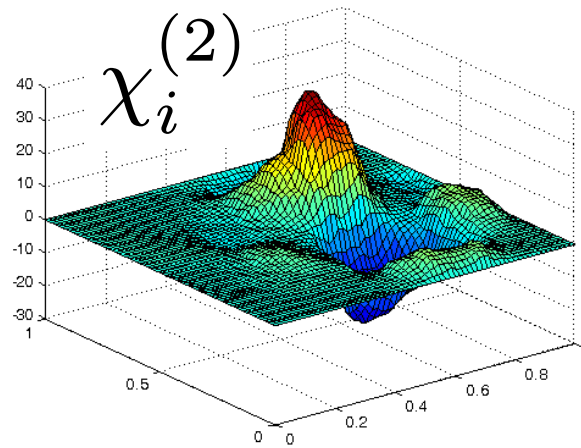
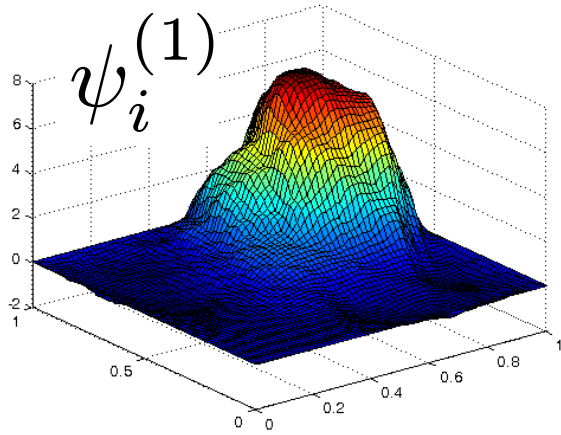
Theorem The decomposition

$$\mathfrak{W} = \mathfrak{W}^{(1)} \oplus_a \mathfrak{W}^{(2)} \oplus_a \cdots \oplus_a \mathfrak{W}^{(k)}$$

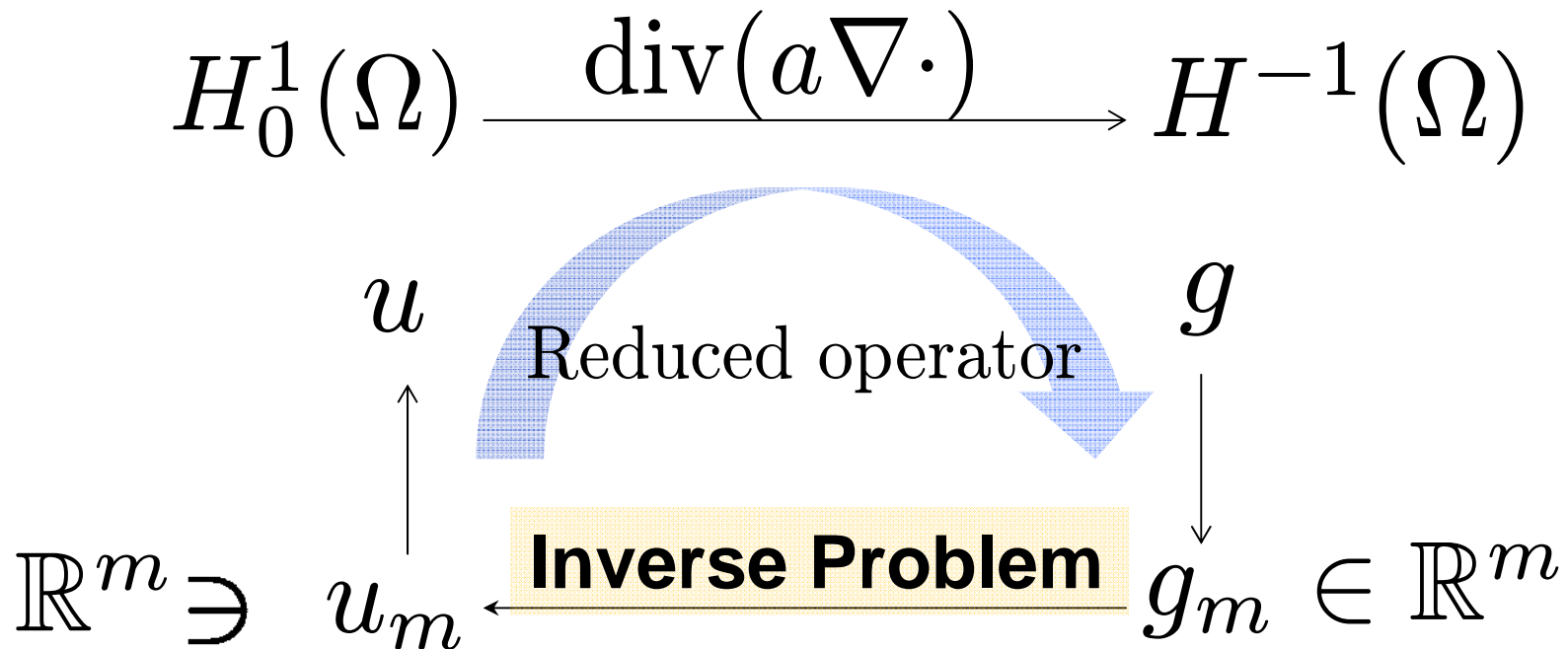
Can be performed and stored in

$$\mathcal{O}(N \ln^2 N) \text{ operations}$$

Doesn't have the complexity of an eigenspace decomposition



**Basis functions look like and behave like wavelets:
Localized and can be used to compress the operator
and locally analyze the solution space**



Numerical implementation requires computation with partial information.

$$\phi_1, \dots, \phi_m \in L^2(\Omega)$$

$$u_m = \left(\int_{\Omega} \phi_1 u, \dots, \int_{\Omega} \phi_m u \right)$$

$$u_m \in \mathbb{R}^m \xrightarrow{\text{Missing information}} u \in H_0^1(\Omega)$$

Discovery process

Identify underlying information game

$$\begin{cases} -\operatorname{div}(a\nabla u) = g \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases}$$

Measurement functions: $\phi_1, \dots, \phi_m \in L^2(\Omega)$

Player A

Chooses

$$g \in L^2(\Omega)$$

$$\|g\|_{L^2(\Omega)} \leq 1$$

Player B

Sees $\int_{\Omega} u\phi_1, \dots, \int_{\Omega} u\phi_m$

Chooses $u^* \in L^2(\Omega)$

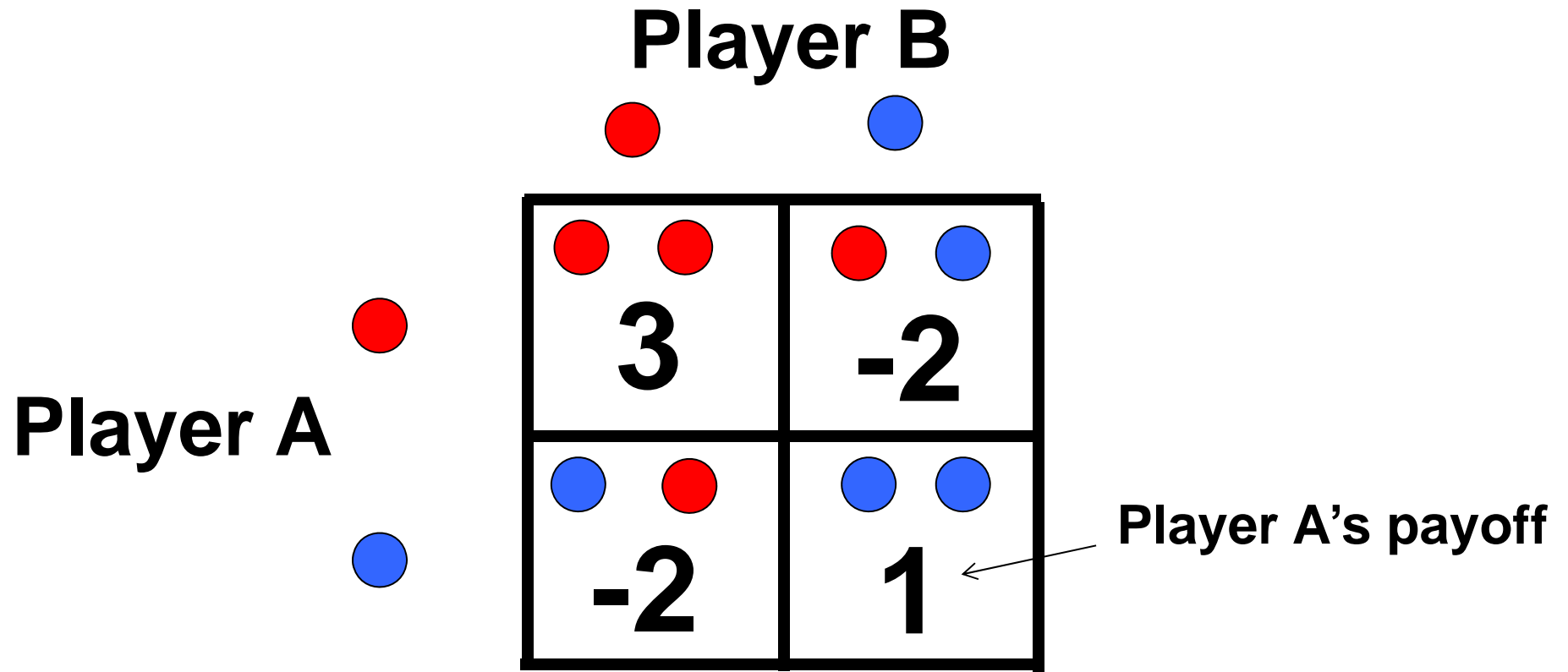
Max

Min

$$\|u - u^*\|_a$$

$$\|f\|_a^2 := \int_{\Omega} (\nabla f)^T a \nabla f$$

Deterministic zero sum game



Player A & B both have a blue and a red marble
At the same time, they show each other a marble

How should A & B play the (repeated) game?

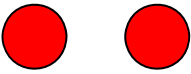
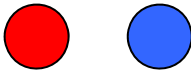

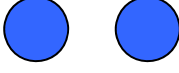
Game theory

Optimal strategies
are mixed strategies

Optimal way to
play is at random

Player B

q ● ● $1 - q$

p ●	 3	 -2
$1 - p$ ●	 -2	 1



John Von Neumann



John Nash

A's expected payoff

$$\begin{aligned} &= 3pq + (1 - p)(1 - q) - 2p(1 - q) - 2q(1 - p) \\ &= 1 - 3q + p(8q - 3) = -\frac{1}{8} \quad \text{for } q = \frac{3}{8} \end{aligned}$$

Player A

Chooses
 $g \in L^2(\Omega)$

$$\|g\|_{L^2(\Omega)} \leq 1$$

Player B

Sees $\int_{\Omega} u\phi_1, \dots, \int_{\Omega} u\phi_m$

Chooses $u^* \in L^2(\Omega)$

$$\|u - u^*\|_a$$

Continuous game but as in decision theory under compactness it can be approximated by a finite game



Abraham Wald

The best strategy for A is to play at random

Player B's best strategy live in the Bayesian class of estimators

Player B's class of mixed strategies

Pretend that player A is choosing g at random

$$g \in L^2(\Omega) \quad \longleftrightarrow \quad \xi: \text{Random field}$$

$$\begin{cases} -\operatorname{div}(a\nabla u) = g \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases} \quad \longleftrightarrow \quad \begin{cases} -\operatorname{div}(a\nabla v) = \xi \text{ in } \Omega, \\ v = 0 \text{ on } \partial\Omega, \end{cases}$$

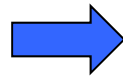
Player B's bet

$$u^*(x) := \mathbb{E} \left[v(x) \mid \int_{\Omega} v(y) \phi_i(y) dy = \int_{\Omega} u(y) \phi_i(y) dy, \forall i \right]$$

Player's B optimal strategy?

Player B's best bet? \longleftrightarrow min max problem
over distribution of ξ

Computational efficiency



$$\xi \sim \mathcal{N}(0, \Gamma)$$



Elementary gambles form deterministic basis functions for player B's bet



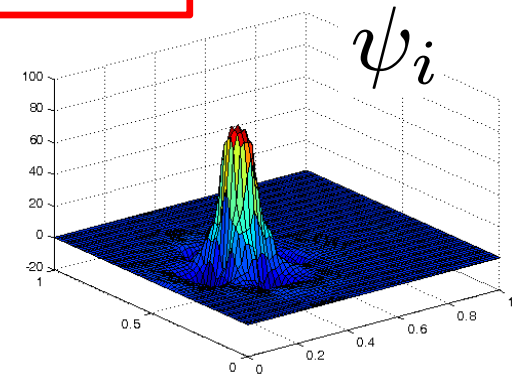
Theorem

$$u^*(x) = \sum_{i=1}^m \psi_i(x) \int_{\Omega} u(y) \phi_i(y) dy$$

Gamblets

ψ_i : Elementary gambles/bets

Player B's bet if $\int_{\Omega} u \phi_j = \delta_{i,j}$, $j = 1, \dots, m$



$$\psi_i(x) := \mathbb{E}_{\xi \sim \mathcal{N}(0, \Gamma)} \left[v(x) \mid \int_{\Omega} v(y) \phi_j(y) dy = \delta_{i,j}, j \in \{1, \dots, m\} \right]$$

What are these gamblets?

Depend on

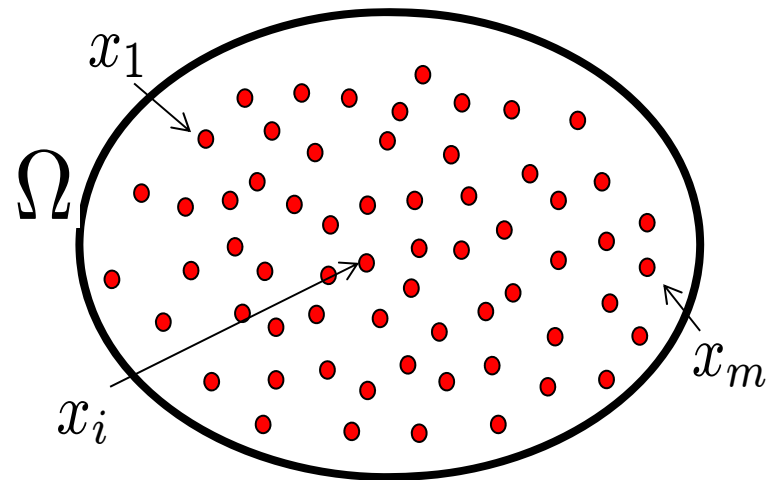
- Γ : Covariance function of ξ (Player B's decision)
- $(\phi_i)_{i=1}^m$: Measurements functions (rules of the game)

Example

[Owhadi, 2014]
arXiv:1406.6668

$$\Gamma(x, y) = \delta(x - y)$$

$$\phi_i(x) = \delta(x - x_i)$$



$a = I_d \iff \psi_i$: Polyharmonic splines

[Harder-Desmarais, 1972] [Duchon 1976, 1977, 1978]

$a_{i,j} \in L^\infty(\Omega) \iff \psi_i$: Rough Polyharmonic splines
[Owhadi-Zhang-Berlyand 2013]

What is Player B's best strategy?

What is Player B's best choice for

$$\Gamma(x, y) = \mathbb{E}[\xi(x)\xi(y)] \quad ?$$

$$\Gamma = \mathcal{L}$$



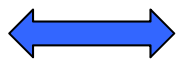
$$\int_{\Omega} \xi(x) f(x) dx \sim \mathcal{N}(0, \|f\|_a^2)$$

$$\|f\|_a^2 := \int_{\Omega} (\nabla f)^T a \nabla f$$



$$\mathcal{L} = -\operatorname{div}(a \nabla \cdot)$$

Why?



See algebraic generalization

The recovery is optimal (Galerkin projection)

Theorem If $\Gamma = \mathcal{L}$ then

$u^*(x)$ is the F.E. solution of (1) in $\text{span}\{\mathcal{L}^{-1}\phi_i | i = 1, \dots, m\}$

$$\|u - u^*\|_a = \inf_{\psi \in \text{span}\{\mathcal{L}^{-1}\phi_i : i \in \{1, \dots, m\}\}} \|u - \psi\|_a$$

$$\mathcal{L} = -\text{div}(a\nabla\cdot)$$

$$(1) \quad \begin{cases} -\text{div}(a\nabla u) = g, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

Optimal variational properties

Theorem

$\sum_{i=1}^m w_i \psi_i$ minimizes $\|\psi\|_a$
over all ψ such that $\int_{\Omega} \phi_j \psi = w_j$ for $j = 1, \dots, m$

Variational characterization

Theorem ψ_i : Unique minimizer of

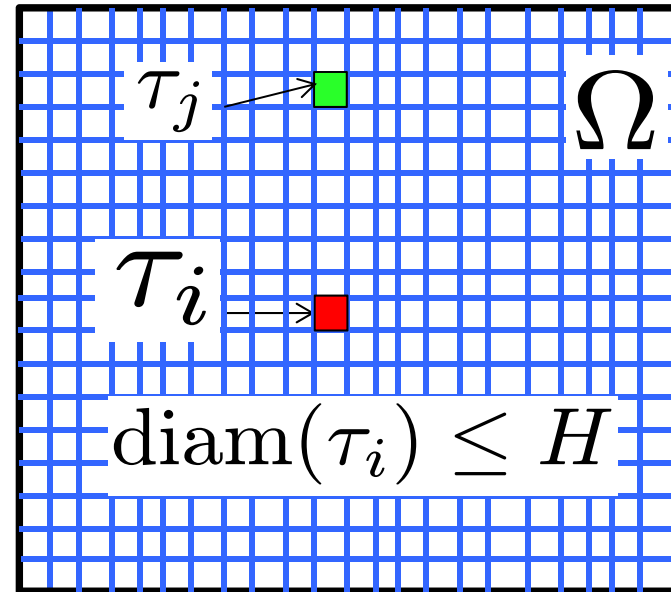
$$\begin{cases} \text{Minimize} & \|\psi\|_a \\ \text{Subject to} & \psi \in H_0^1(\Omega) \text{ and } \int_{\Omega} \phi_j \psi = \delta_{i,j}, \quad j = 1, \dots, m \end{cases}$$

Selection of measurement functions

Example Indicator functions of a
Partition of Ω of resolution H



$$\phi_i = \mathbf{1}_{\tau_i}$$



Theorem

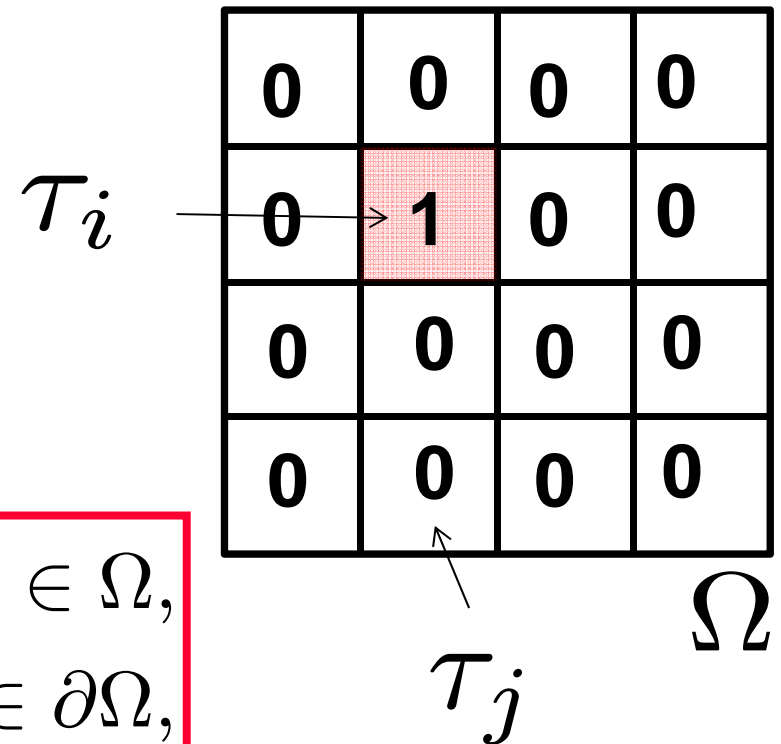
$$\|u - u^*\|_a \leq \frac{H}{\lambda_{\min}(a)} \|g\|_{L^2(\Omega)}$$

Elementary gamble

ψ_i

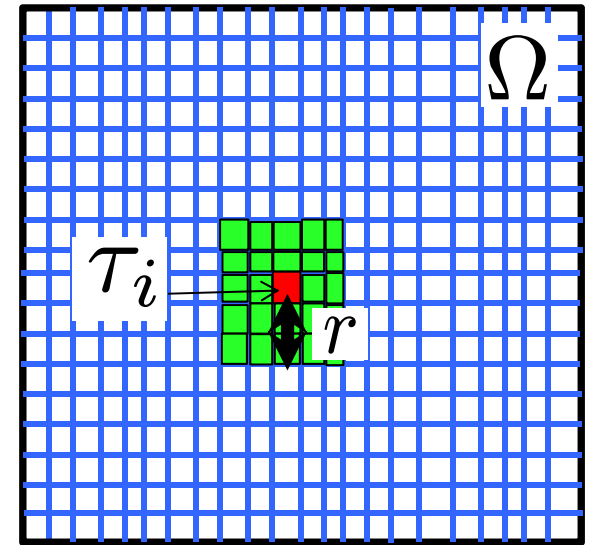
Your best bet on the value of u given the information that

$$\int_{\tau_i} u = 1 \text{ and } \int_{\tau_j} u = 0 \text{ for } j \neq i$$



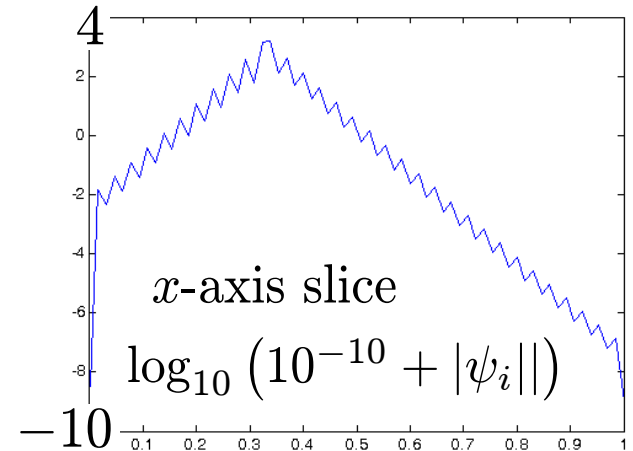
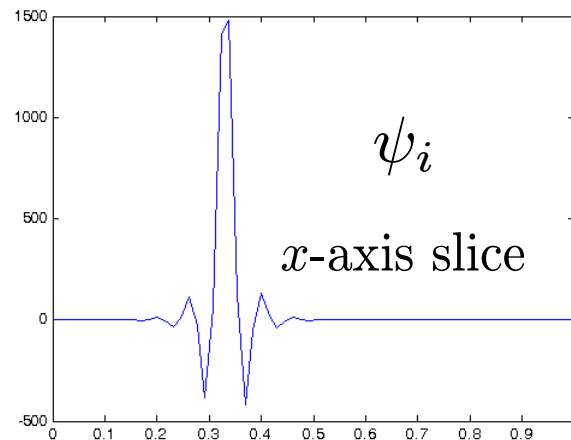
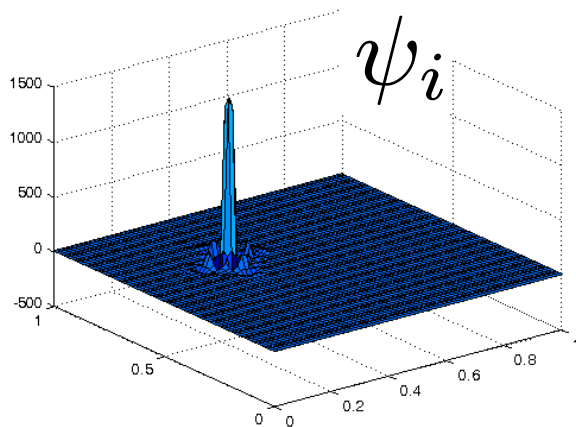
$$(1) \quad \begin{cases} -\operatorname{div}(a\nabla u) = g, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases}$$

Exponential decay of gamblets



Theorem

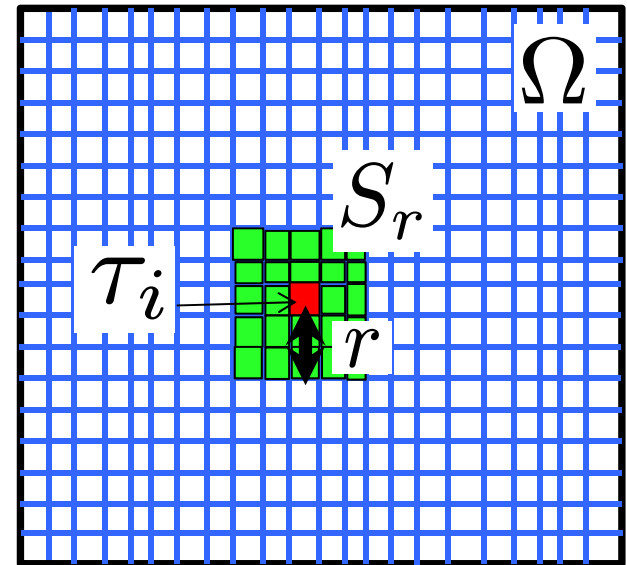
$$\int_{\Omega \cap (B(\tau_i, r))^c} (\nabla \psi_i)^T a \nabla \psi_i \leq e^{-\frac{r}{tH}} \|\psi_i\|_a^2$$



Localization of the computation of gamblets

$\psi_i^{\text{loc},r}$: Minimizer of

$$\begin{cases} \text{Minimize} & \|\psi\|_a \\ \text{Subject to} & \psi \in H_0^1(S_r) \text{ and } \int_{S_r} \phi_j \psi = \delta_{i,j} \\ & \text{for } \tau_j \in S_r \end{cases}$$



No loss of accuracy if
localization $\sim H \ln \frac{1}{H}$

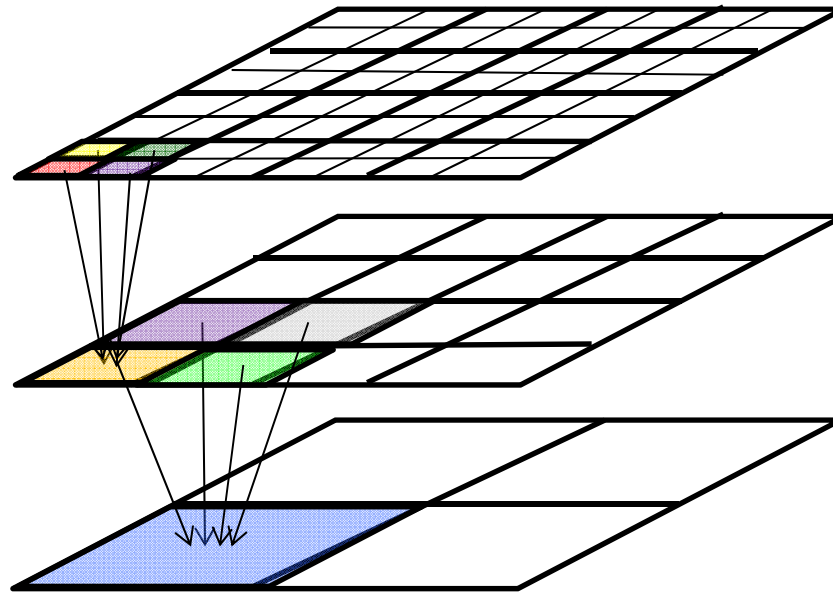
$$u^{*,\text{loc}}(x) = \sum_{i=1}^m \psi_i^{\text{loc},r}(x) \int_{\Omega} u(y) \phi_i(y) dy$$

Theorem

If $r \geq CH \ln \frac{1}{H}$

$$\|u - u^{*,\text{loc}}\|_a \leq \frac{1}{\sqrt{\lambda_{\min}(a)}} H \|g\|_{L^2(\Omega)}$$

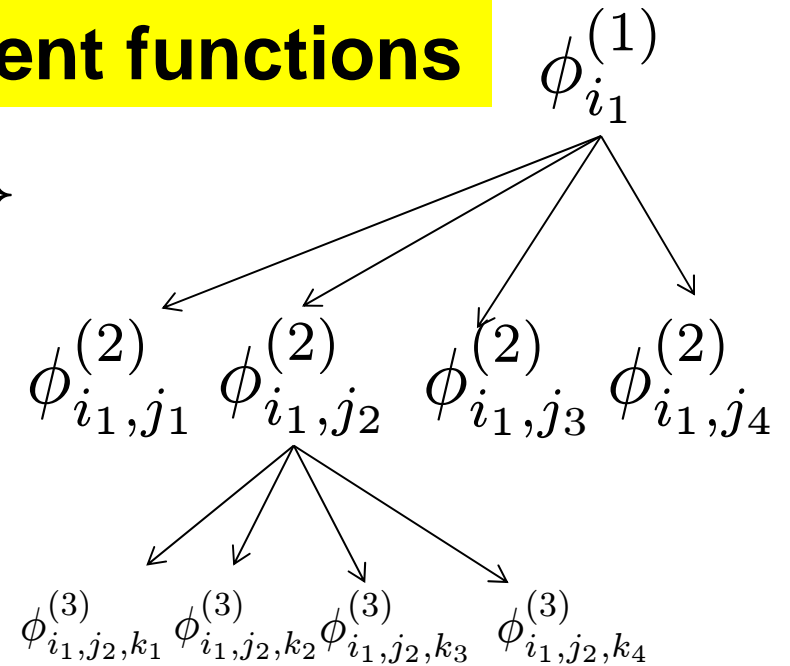
Formulation of the hierarchical game



Hierarchy of nested Measurement functions

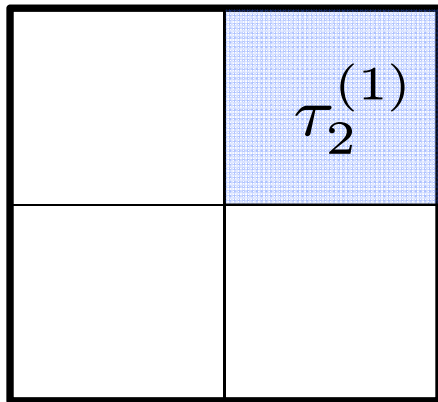
$$\phi_{i_1, \dots, i_k}^{(k)} \text{ with } k \in \{1, \dots, q\}$$

$$\phi_i^{(k)} = \sum_j c_{i,j} \phi_{i,j}^{(k+1)}$$

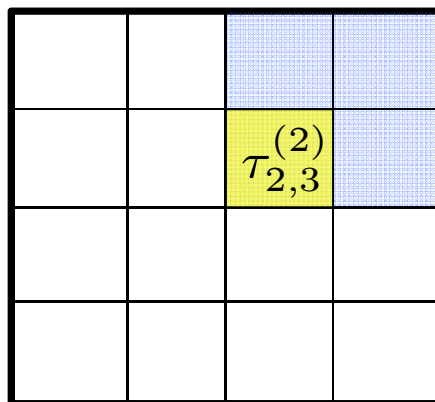


Example

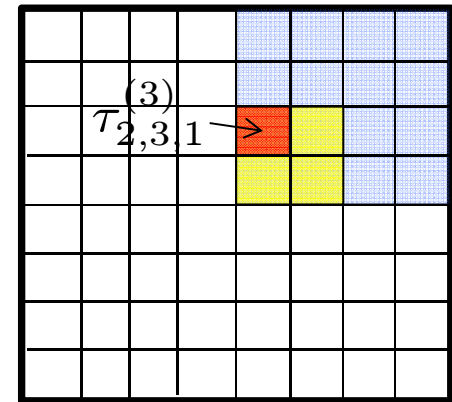
$\phi_i^{(k)}$: Indicator functions of a hierarchical nested partition of Ω of resolution $H_k = 2^{-k}$



$$\phi_2^{(1)} = 1_{\tau_2^{(1)}}$$

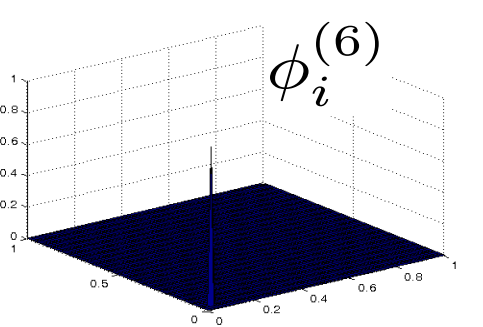
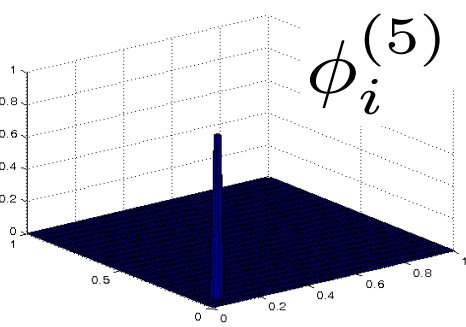
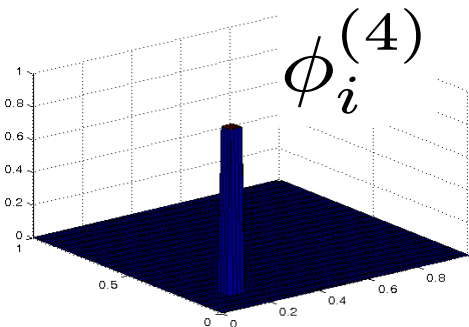
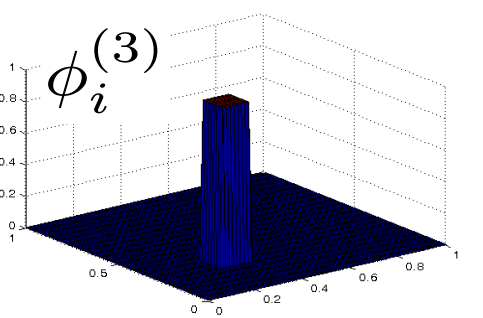
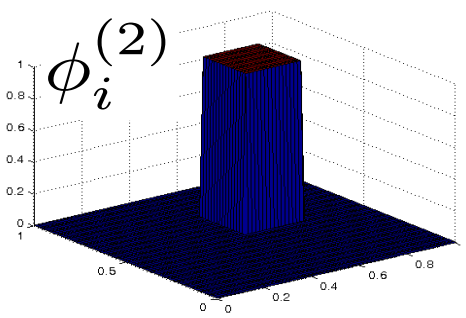
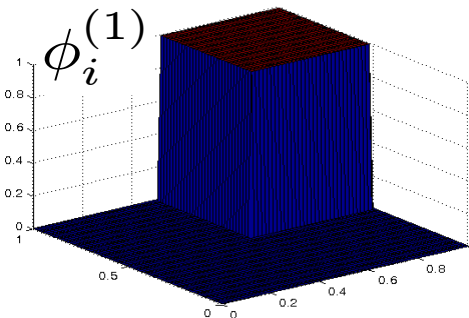
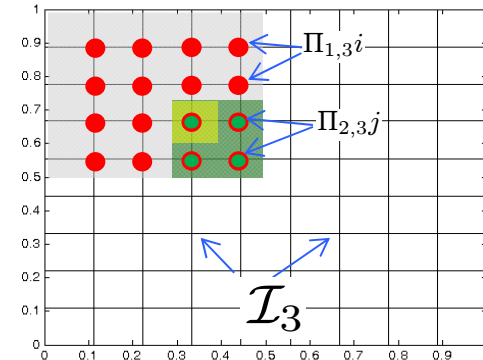
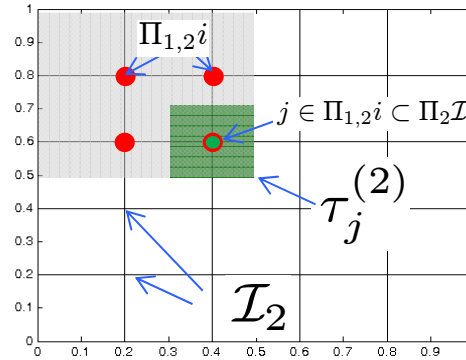
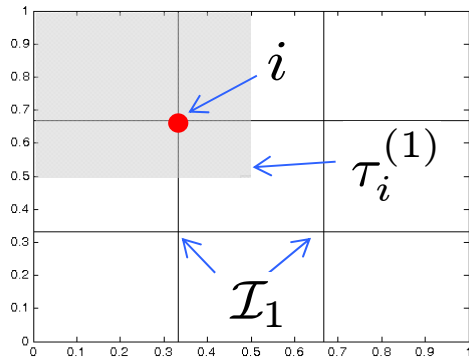


$$\phi_{2,3}^{(2)} = 1_{\tau_{2,3}^{(2)}}$$



$$\phi_{2,3,1}^{(3)} = 1_{\tau_{2,3,1}^{(3)}}$$

In the discrete setting simply aggregate elements (as in algebraic multigrid)



Formulation of the hierarchy of games

Player A

Chooses

$$g \in L^2(\Omega)$$

$$\|g\|_{L^2(\Omega)} \leq 1$$

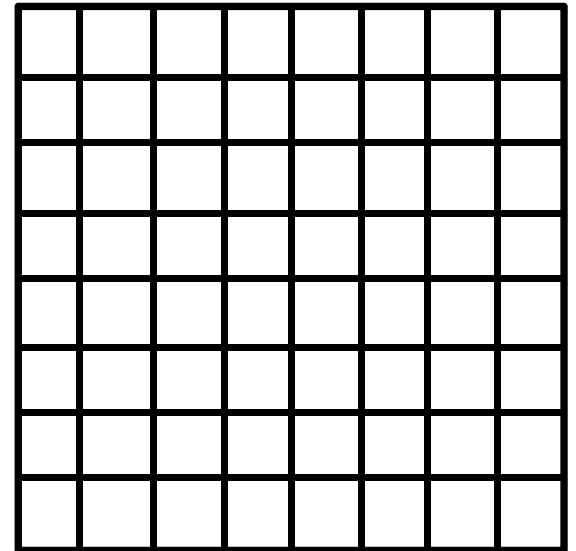
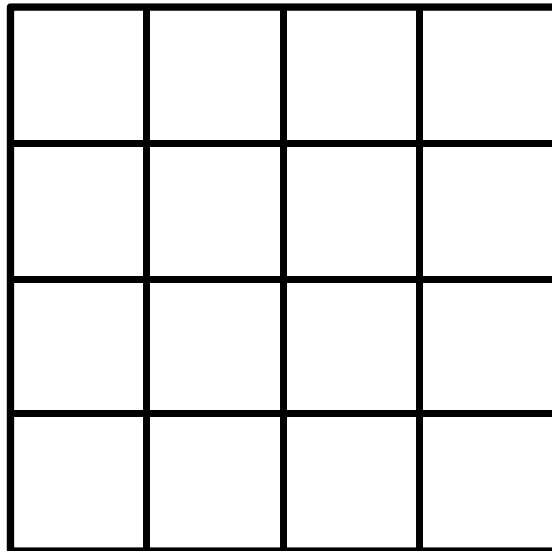
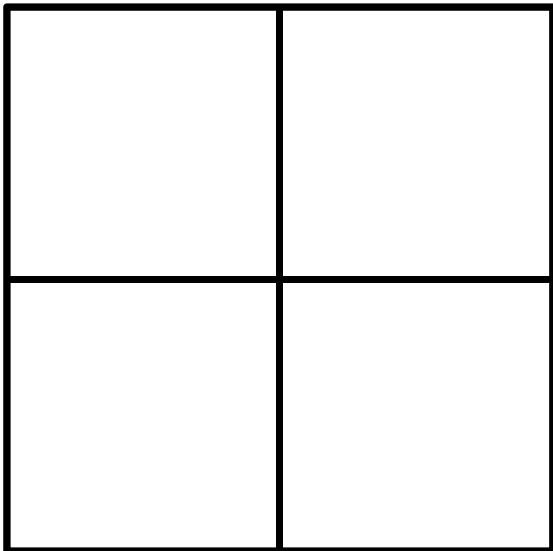
$$\begin{cases} -\operatorname{div}(a\nabla u) = g \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases}$$

Player B

Sees $\{\int_{\Omega} u \phi_i^{(k)}, i \in \mathcal{I}_k\}$

Must predict

u and $\{\int_{\Omega} u \phi_j^{(k+1)}, j \in \mathcal{I}_{k+1}\}$



Player B's best strategy

$$\begin{cases} -\operatorname{div}(a\nabla u) = g & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$



$$\begin{cases} -\operatorname{div}(a\nabla v) = \xi & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega, \end{cases}$$

$$\xi \sim \mathcal{N}(0, \mathcal{L})$$

Player B's bets

$$u^{(k)}(x) := \mathbb{E}\left[v(x) \mid \int_{\Omega} v(y) \phi_i^{(k)}(y) dy = \int_{\Omega} u(y) \phi_i^{(k)}(y) dy, i \in \mathcal{I}_k\right]$$

The sequence of approximations forms a martingale under the mixed strategy emerging from the game

$$\mathcal{F}_k = \sigma\left(\int_{\Omega} v \phi_i^{(k)}, i \in \mathcal{I}_k\right)$$

$$v^{(k)}(x) := \mathbb{E}\left[v(x) \mid \mathcal{F}_k\right]$$

Theorem

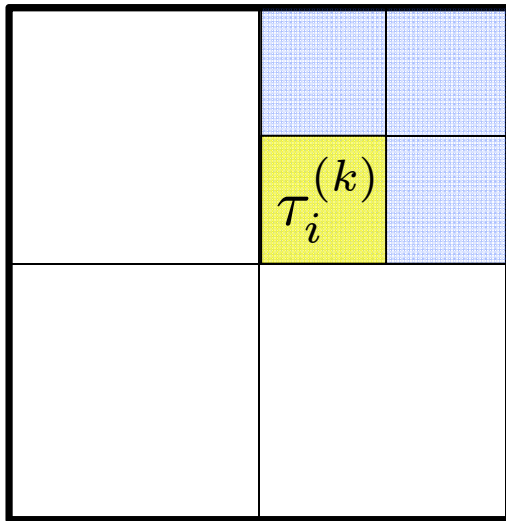
$$\mathcal{F}_k \subset \mathcal{F}_{k+1}$$

$$v^{(k)}(x) := \mathbb{E}\left[v^{(k+1)}(x) \mid \mathcal{F}_k\right]$$

Accuracy of the recovery

Theorem

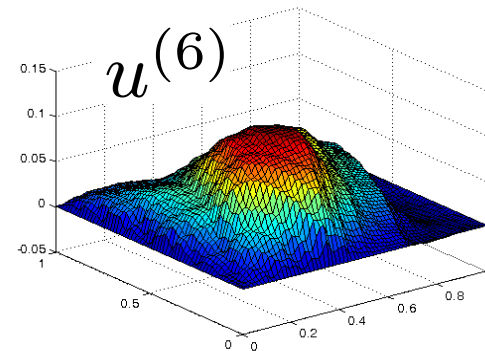
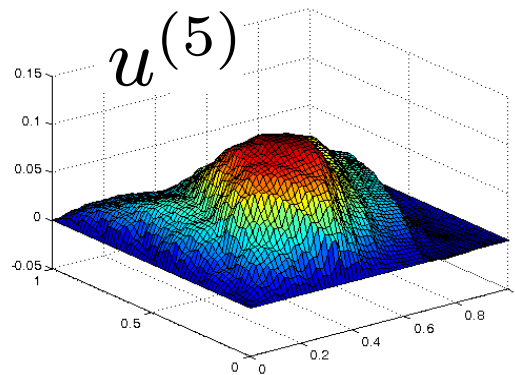
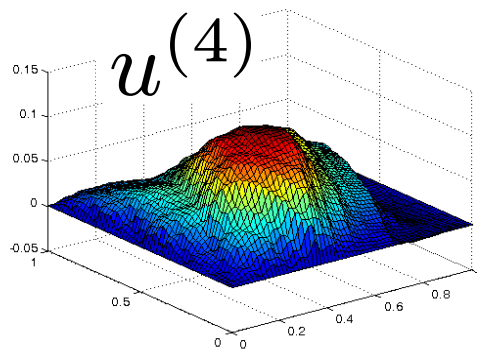
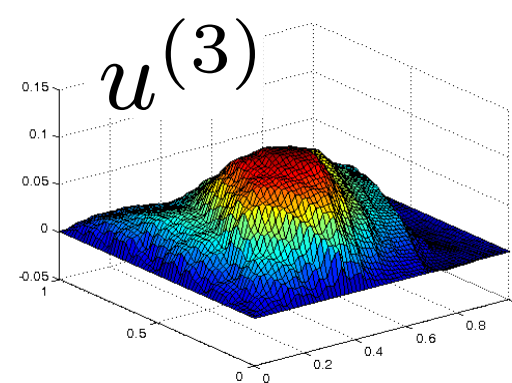
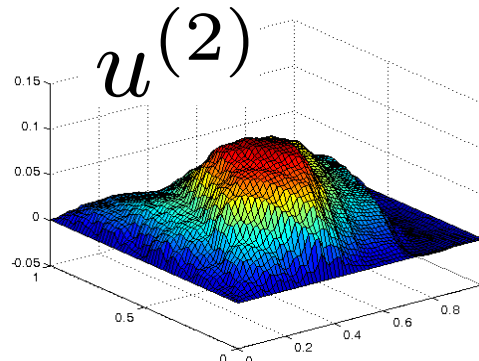
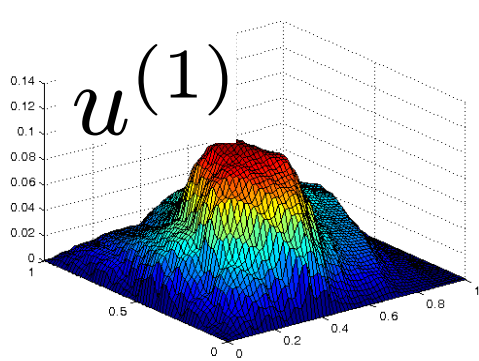
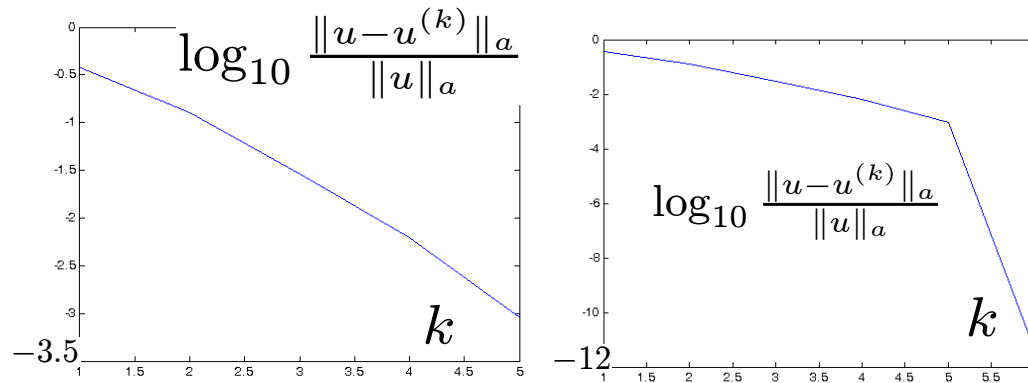
$$\|u - u^{(k)}\|_a \leq \frac{H_k}{\lambda_{\min}(a)} \|g\|_{L^2(\Omega)}$$



$$H_k := \max_i \text{diam}(\tau_i^{(k)})$$

$$\phi_i^{(k)} = 1_{\tau_i^{(k)}} \quad \text{diam}(\tau_i^{(k)}) \leq H_k$$

In a discrete setting the last step of the game recovers the solution to numerical precision



Gamblets

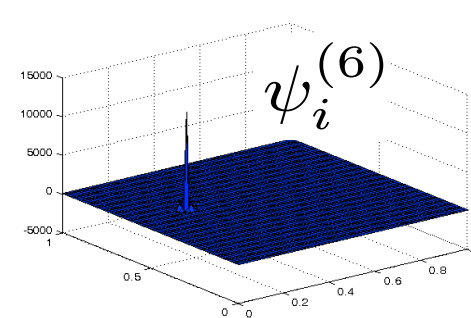
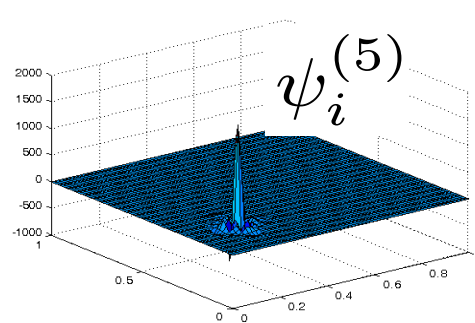
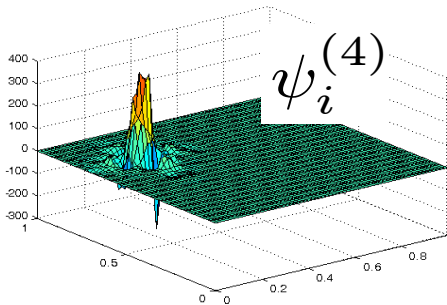
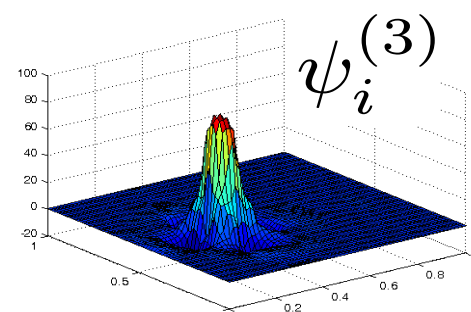
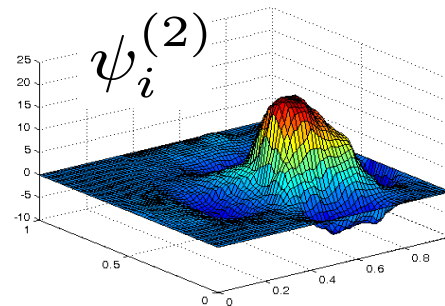
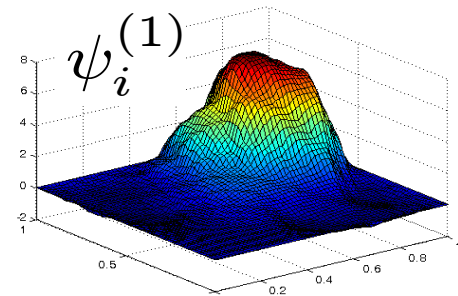
Elementary gambles form a hierarchy of deterministic basis functions for player B's hierarchy of bets

Theorem

$$u^{(k)}(x) = \sum_i \psi_i^{(k)}(x) \int_{\Omega} u(y) \phi_i^{(k)}(y) dy$$

$\psi_i^{(k)}$: Elementary gambles/bets at resolution $H_k = 2^{-k}$

$$\psi_i^{(k)}(x) := \mathbb{E} \left[v(x) \mid \int_{\Omega} v(y) \phi_j^{(k)}(y) dy = \delta_{i,j}, j \in \mathcal{I}_k \right]$$

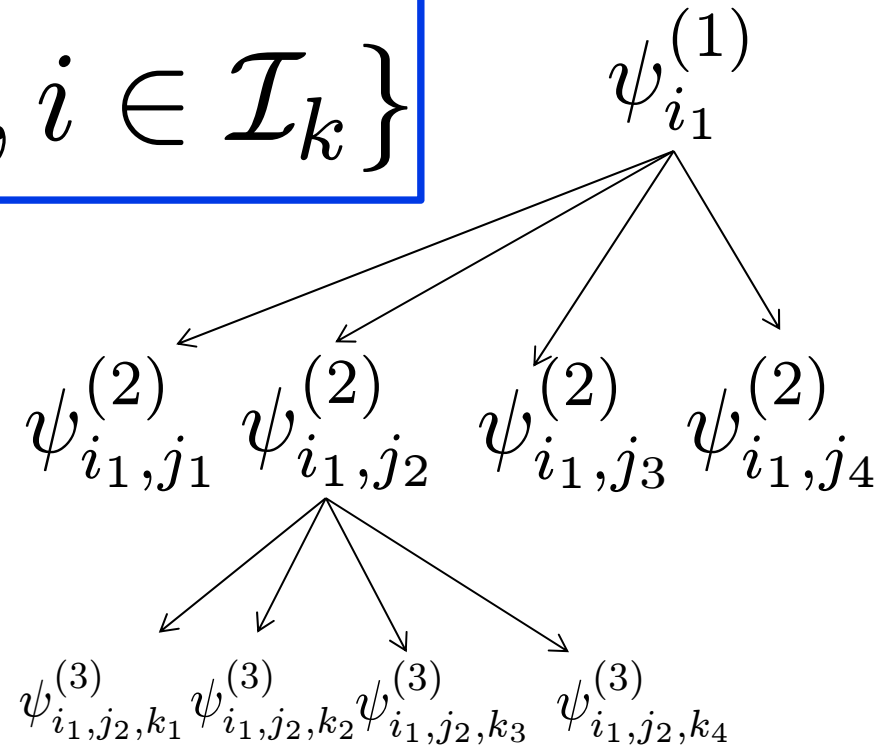


Gamblets are nested

$$\mathfrak{G}^{(k)} := \text{span} \{ \psi_i^{(k)}, i \in \mathcal{I}_k \}$$

Theorem

$$\mathfrak{G}^{(k)} \subset \mathfrak{G}^{(k+1)}$$

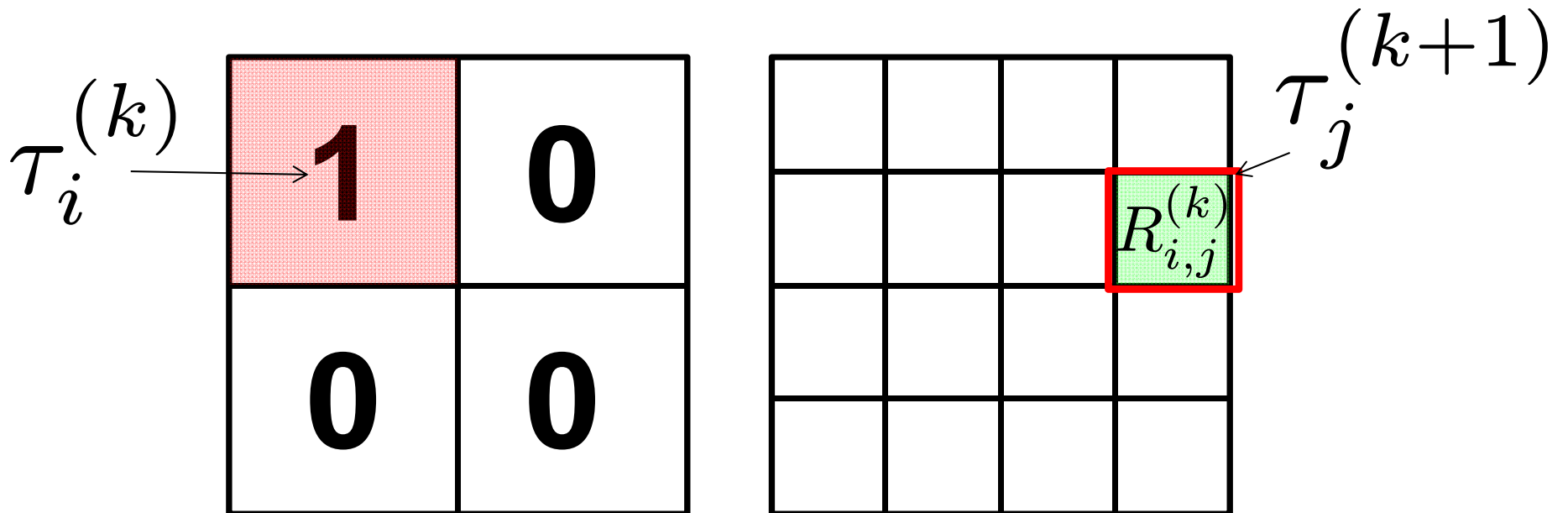


$$\psi_i^{(k)}(x) = \sum_{j \in \mathcal{I}_{k+1}} R_{i,j}^{(k)} \psi_j^{(k+1)}(x)$$

Interpolation/Prolongation operator

$$R_{i,j}^{(k)} = \mathbb{E} \left[\int_{\Omega} v(y) \phi_j^{(k+1)}(y) dy \mid \int_{\Omega} v(y) \phi_l^{(k)}(y) dy = \delta_{i,l}, l \in \mathcal{I}_k \right]$$

$R_{i,j}^{(k)}$ Your best bet on the value of $\int_{\tau_j^{(k+1)}} u$ given the information that $\int_{\tau_i^{(k)}} u = 1$ and $\int_{\tau_l} u = 0$ for $l \neq i$



**At this stage you can finish with
classical multigrid**

But we want multiresolution decomposition

Elementary gamble

$$\chi_i^{(k)}$$

Your best bet on the value of u given the information that

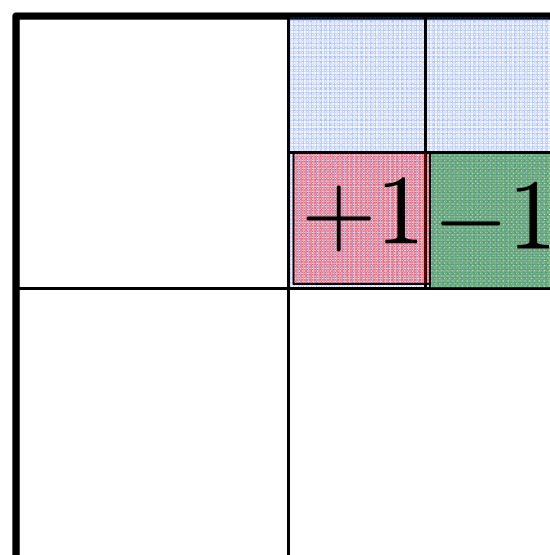
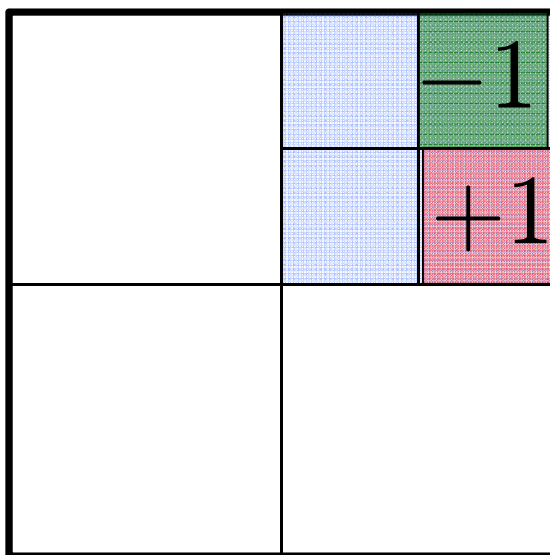
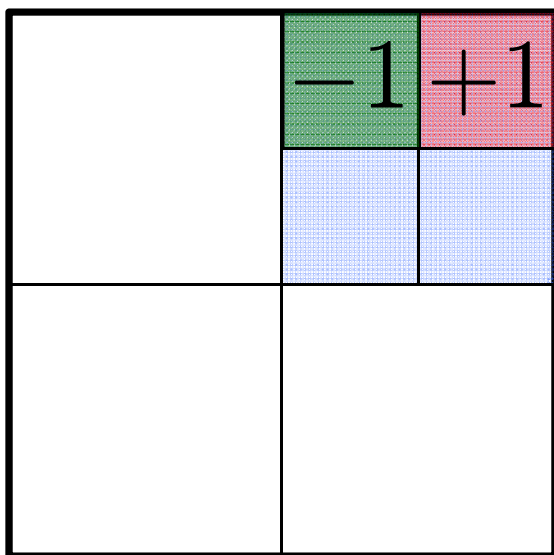
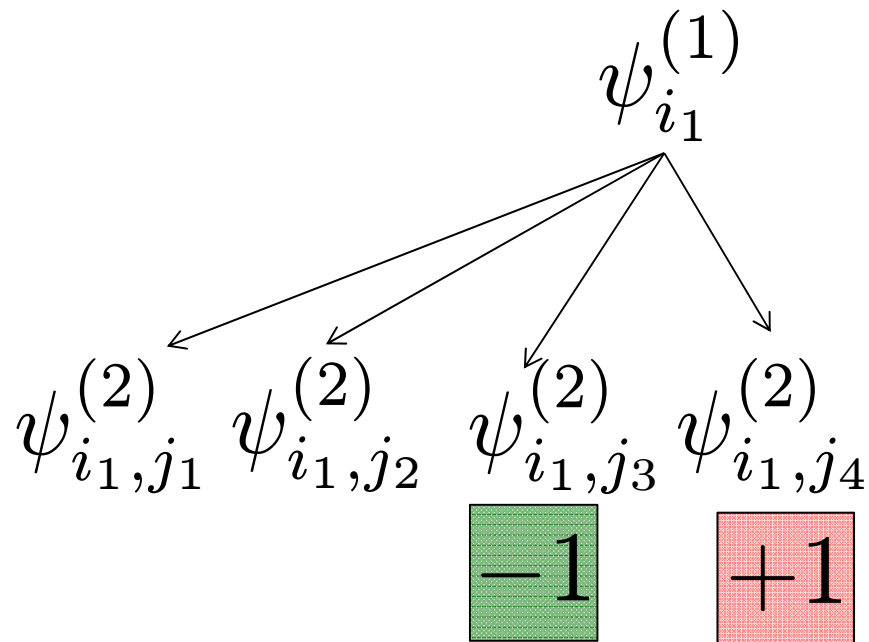
$$\int_{\tau_i^{(k)}} u = 1, \int_{\tau_{i^-}^{(k)}} u = -1 \text{ and } \int_{\tau_j^{(k)}} u = 0 \text{ for } j \neq i$$

		$\tau_{i^-}^{(k)}$		
	0	-1	0	0
$\tau_i^{(k)}$	0	1	0	0
	0	0	0	0
	0	0	0	0
				Ω

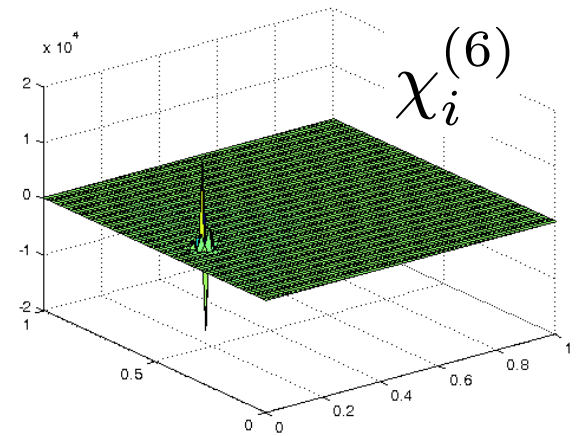
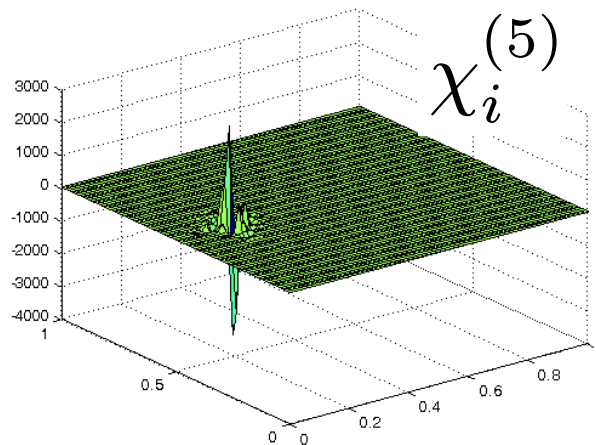
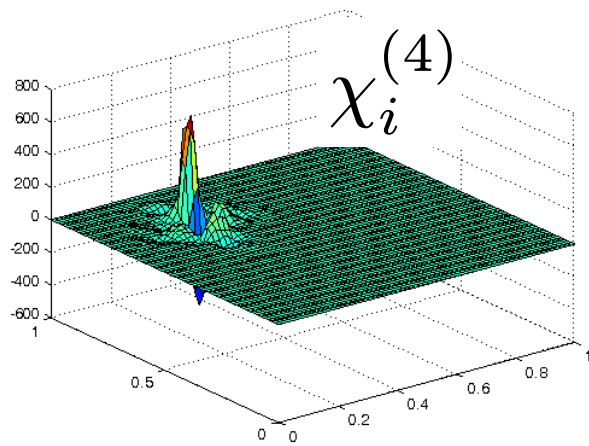
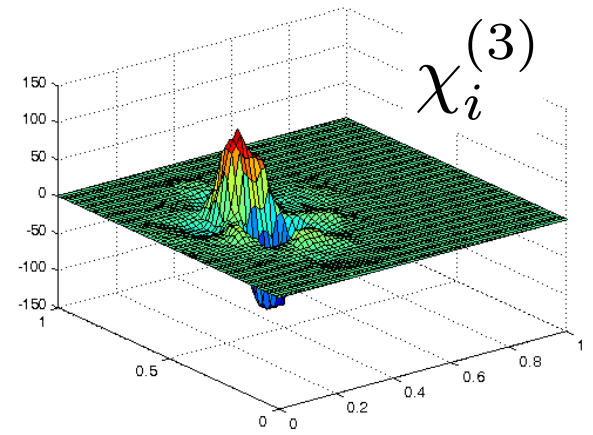
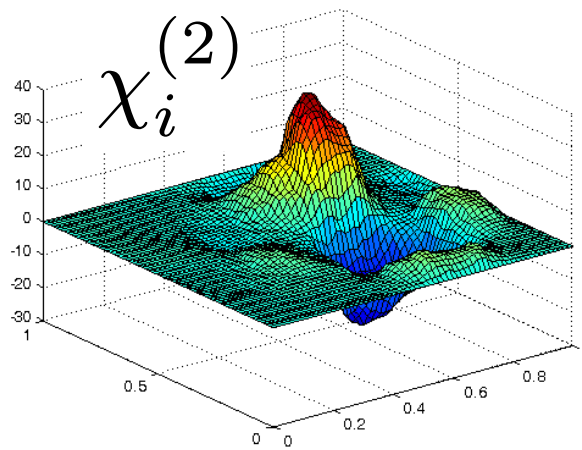
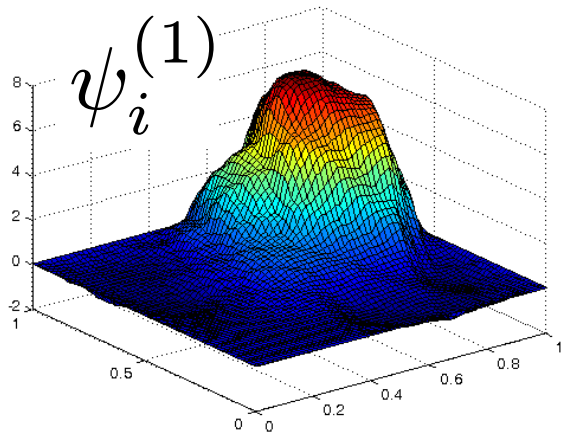
$$\chi_i^{(k)} = \psi_i^{(k)} - \psi_{i^-}^{(k)}$$

$$i = (i_1, \dots, i_{k-1}, i_k)$$

$$i^- = (i_1, \dots, i_{k-1}, i_k - 1)$$



$$\chi_i^{(k)} = \psi_i^{(k)} - \psi_{i-}^{(k)}$$



Multiresolution decomposition of the solution space

$$\mathfrak{V}^{(k)} := \text{span}\{\psi_i^{(k)}, i \in \mathcal{I}_k\}$$

$$\mathfrak{W}^{(k)} := \text{span}\{\chi_i^{(k)}, i\}$$

$\mathfrak{W}^{(k+1)}$: Orthogonal complement of $\mathfrak{V}^{(k)}$ in $\mathfrak{V}^{(k+1)}$
with respect to $\langle \psi, \chi \rangle_a := \int_{\Omega} (\nabla \psi)^T a \nabla \chi$

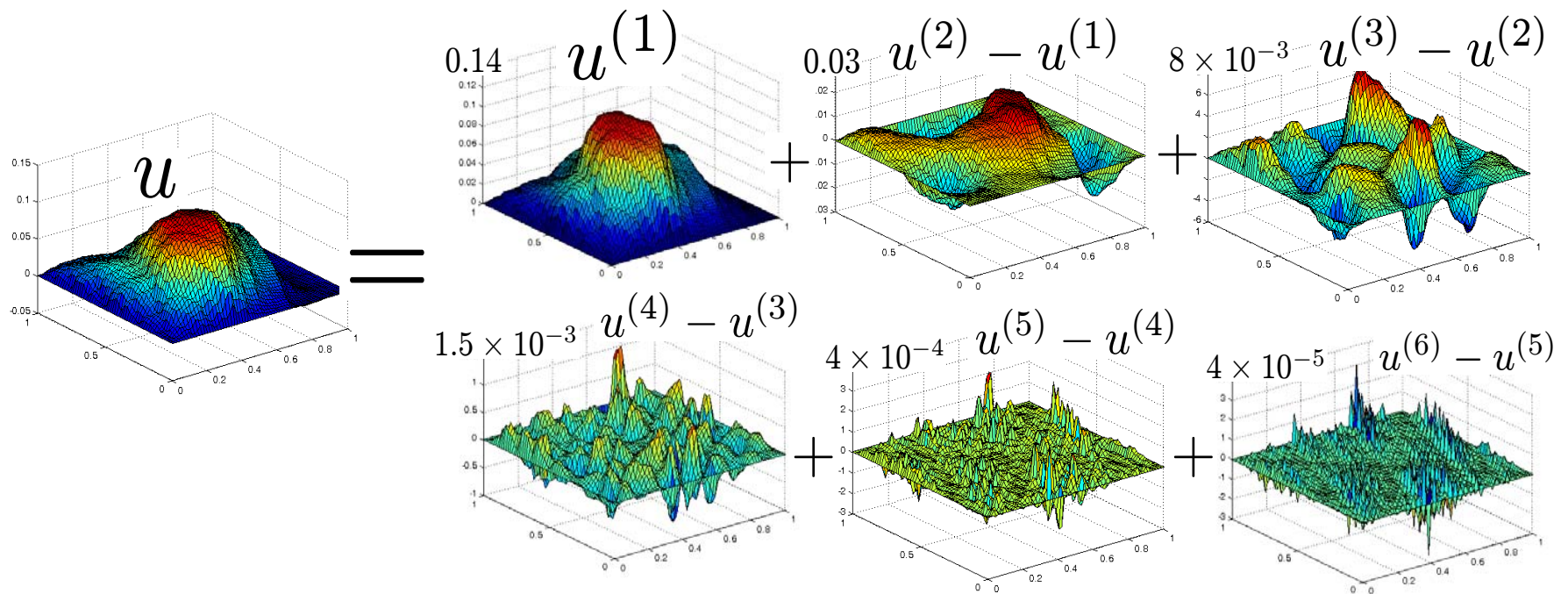
Theorem

$$H_0^1(\Omega) = \mathfrak{V}^{(1)} \oplus_a \mathfrak{W}^{(2)} \oplus_a \cdots \oplus_a \mathfrak{W}^{(k)} \oplus_a \cdots$$

Multiresolution decomposition of the solution

Theorem

$u^{(k+1)} - u^{(k)} = \text{F.E. sol. of PDE in } \mathfrak{W}^{(k+1)}$



Subband solutions $u^{(k+1)} - u^{(k)}$
can be computed independently

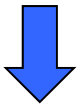
Uniformly bounded condition numbers

$$A_{i,j}^{(k)} := \langle \psi_i^{(k)}, \psi_j^{(k)} \rangle_a$$

$$B_{i,j}^{(k)} := \langle \chi_i^{(k)}, \chi_j^{(k)} \rangle_a$$

Theorem

$$\frac{\lambda_{\max}(B^{(k)})}{\lambda_{\min}(B^{(k)})} \leq C$$

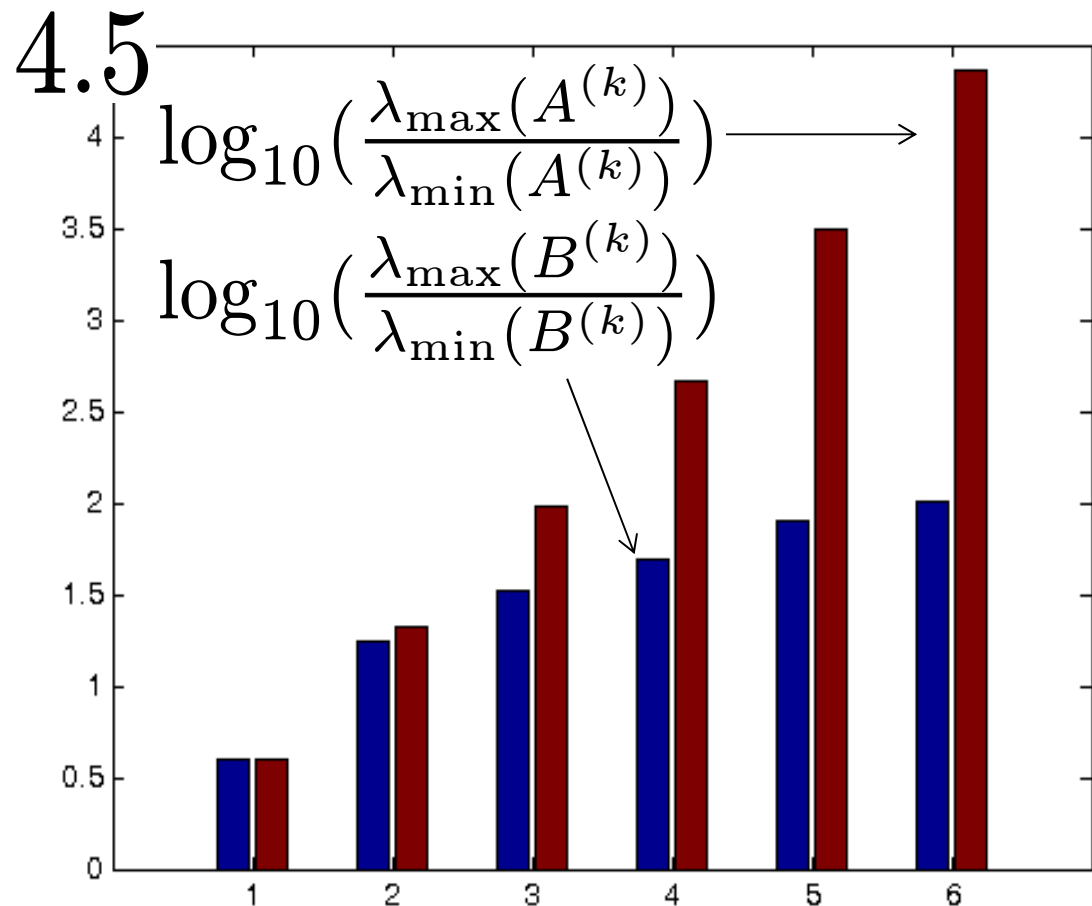


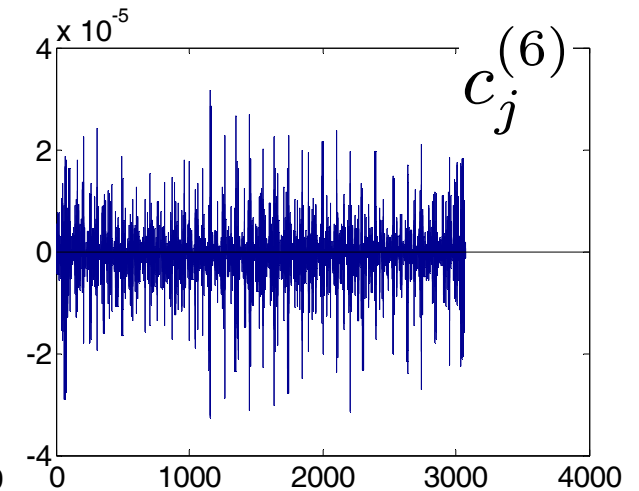
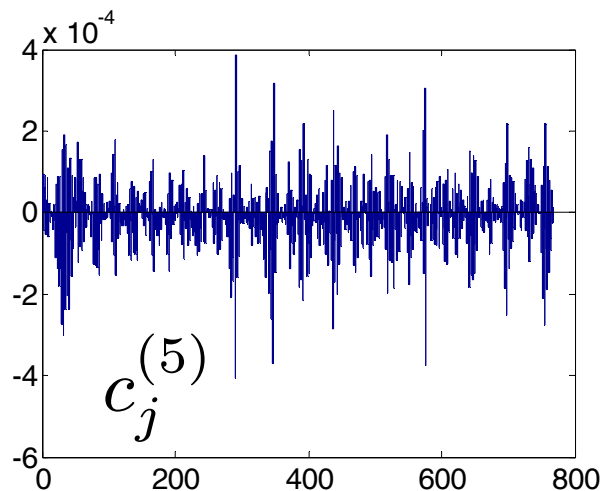
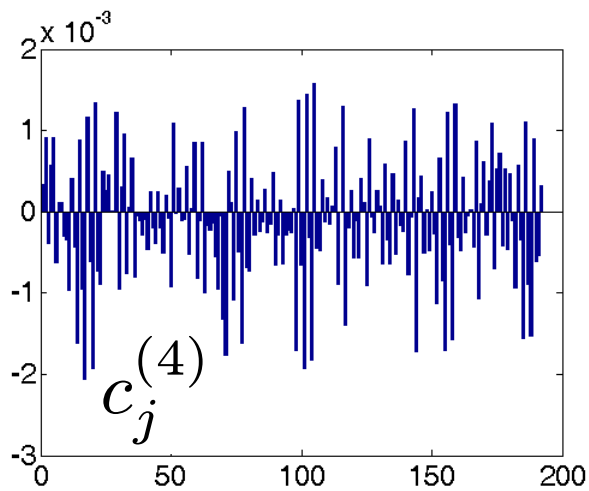
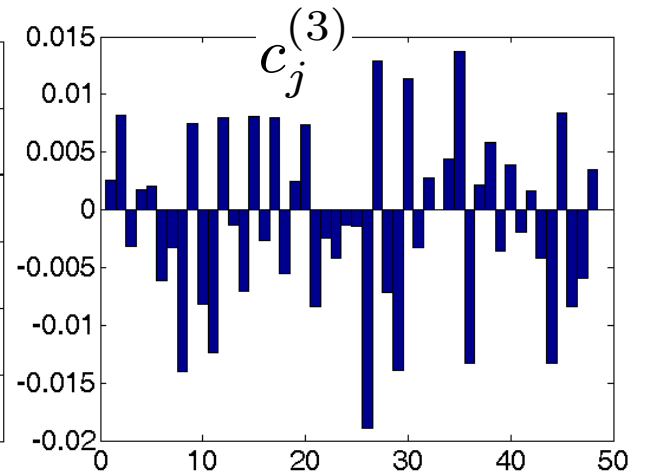
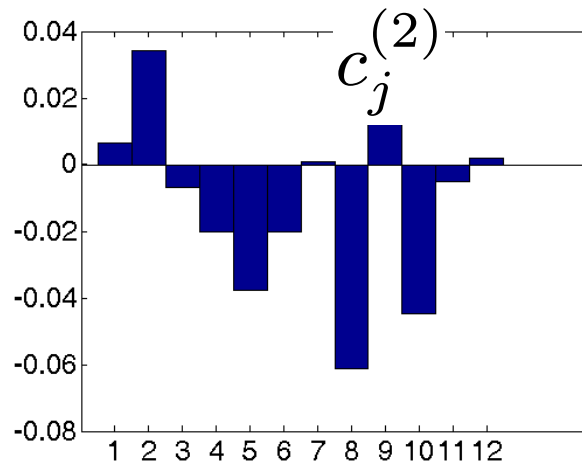
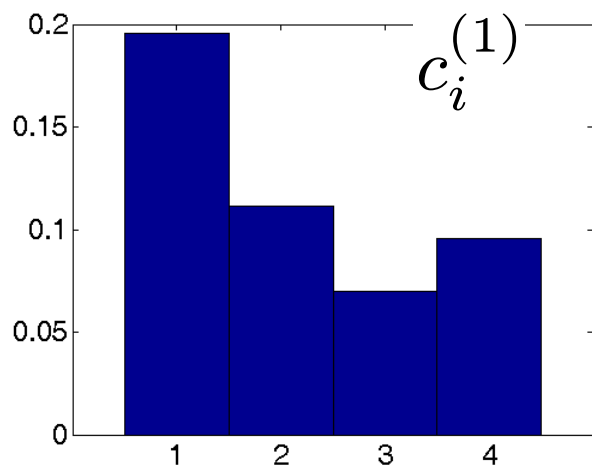
Just relax!

In $v \in \mathfrak{W}^{(k)}$

to get

$$u^{(k)} - u^{(k-1)}$$



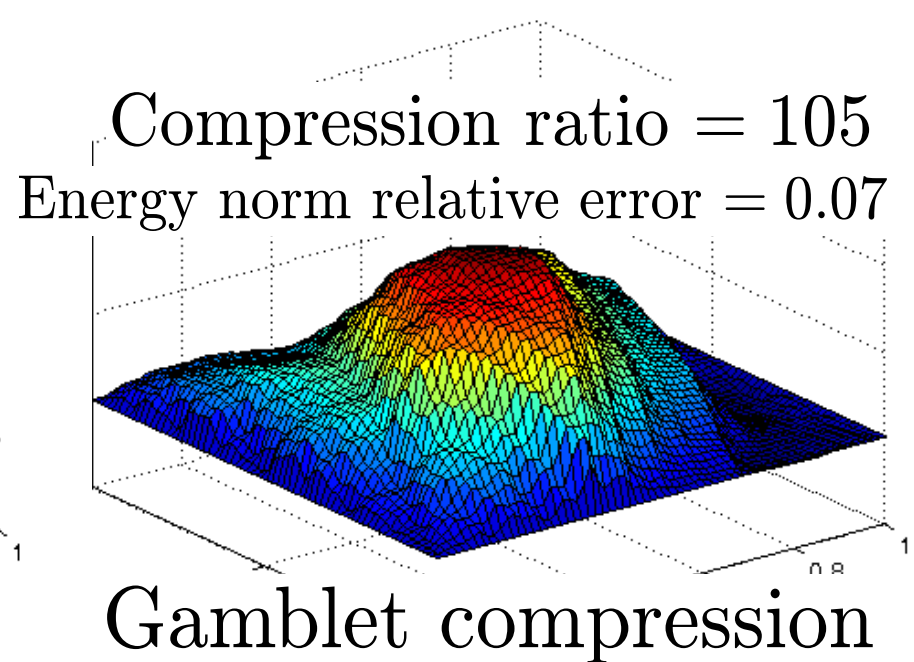
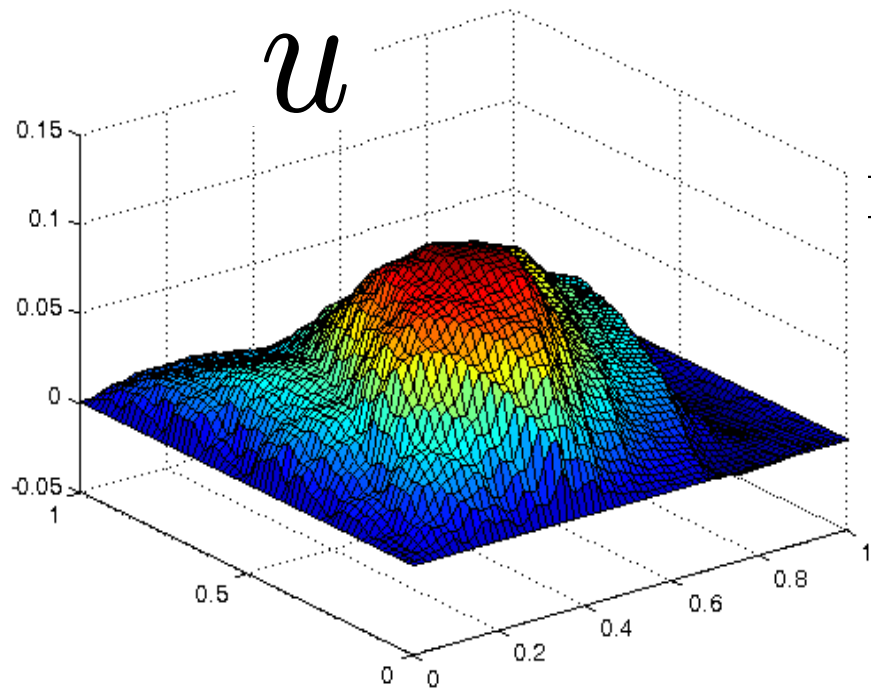


$$u = \sum_i c_i^{(1)} \frac{\psi_i^{(1)}}{\|\psi_i^{(1)}\|_a} + \sum_{k=2}^q \sum_j c_j^{(k)} \frac{\chi_j^{(k)}}{\|\chi_j^{(k)}\|_a}$$

Coefficients of the solution in the gamblet basis

Operator Compression

Gamblets behave like wavelets but they are adapted to the PDE and can compress its solution space



Throw 99% of the coefficients

Fast gamblet transform

$\mathcal{O}(N \ln^2 N)$ complexity

Nesting

$$A^{(k)} = (R^{(k,k+1)})^T A^{(k+1)} R^{(k,k+1)}$$

Level(k) gamblets and stiffness matrices can be computed from level(k+1) gamblets and stiffness matrices

Well conditioned linear systems

Underlying linear systems have uniformly bounded condition numbers

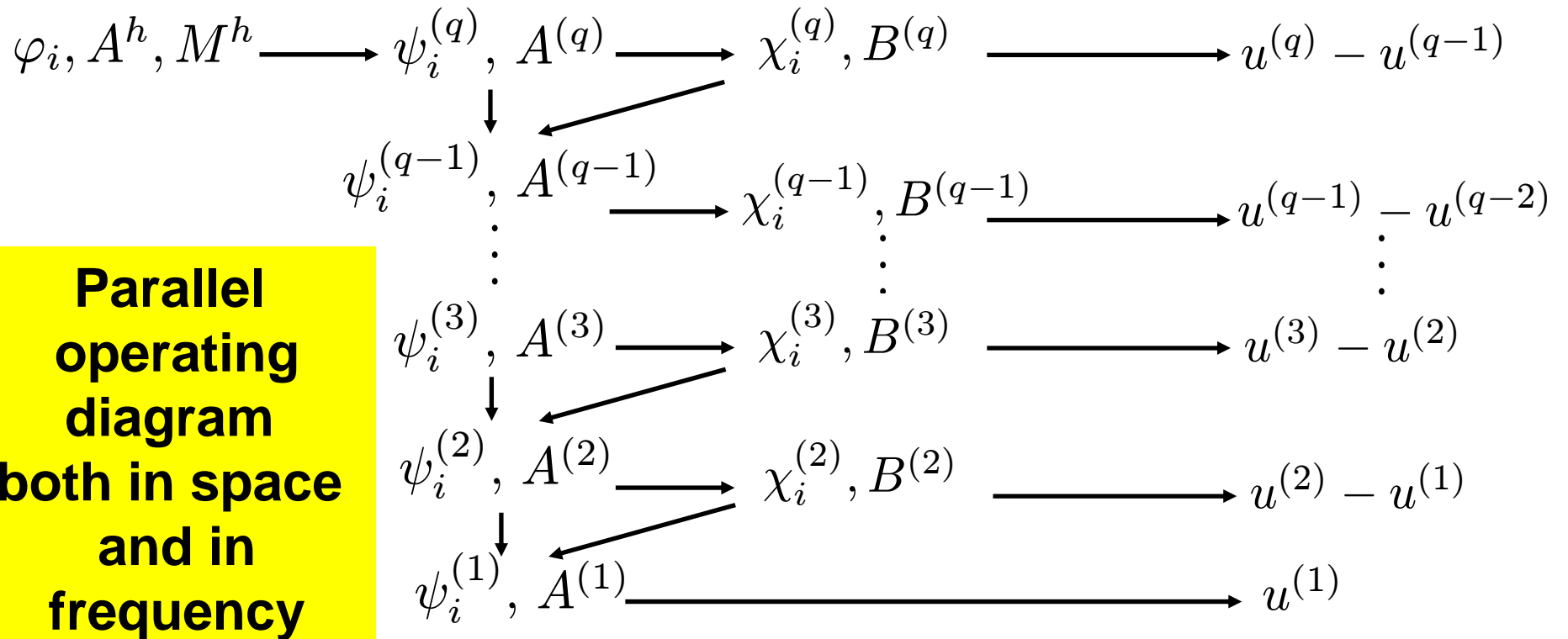
$$\psi_i^{(k)} = \psi_{(i,1)}^{(k+1)} + \sum_j C_{i,j}^{(k+1),\chi} \chi_j^{(k+1)}$$

$$C^{(k+1),\chi} = (B^{(k+1)})^{-1} Z^{(k+1)}$$

$$Z_{j,i}^{(k+1)} := -(e_j^{(k+1)} - e_{j^-}^{(k+1)})^T A^{(k+1)} e_{(i,1)}^{(k+1)}$$

Localization

The nested computation can be localized without compromising accuracy or condition numbers



**Parallel
operating
diagram
both in space
and in
frequency**

