# Principles and Methods of UQ
# A minitutorial, Part II

**Houman Owhadi**

**IPAM Sep 13-14, 2017**

DARPA EQUiPS / AFOSR award no FA9550-16-1-0054
(Computational Information Games)

# Application

(1) $\boxed{\mathcal{L}u = g}$

$\Omega \subset \mathbb{R}^d$

$\mathcal{L}$: arbitrary, symmetric, positive, continuous linear bijection

$$(H_0^s(\Omega), \|\cdot\|_{H_0^s(\Omega)}) \xrightarrow{\quad\mathcal{L}\quad} (H^{-s}(\Omega), \|\cdot\|_{H^{-s}(\Omega)})$$

$\mathcal{L}$ local: $\int_\Omega u\mathcal{L}v = 0$ if $u$ and $v$ have disjoint supports

**You want to solve (1) as fast as possible to a given degree of accuracy**

You want to approximate an eigensubspaces of $\mathcal{L}$

You want to find Wannier functions of $\mathcal{L}$

You want a multiresolution decomposition of $\mathcal{L}$

H. Owhadi and C. Scovel. Universal Scalable Robust Solvers from Computational Information Games and fast eigenspace adapted Multiresolution Analysis, 2017. arXiv:1703.10761

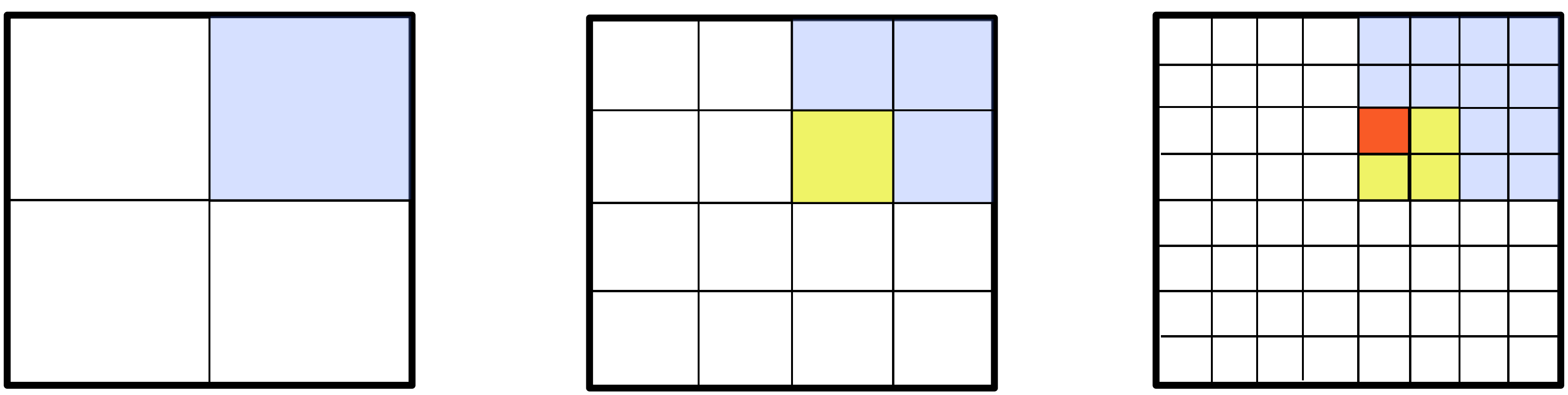
## Hierarchy of measurement functions

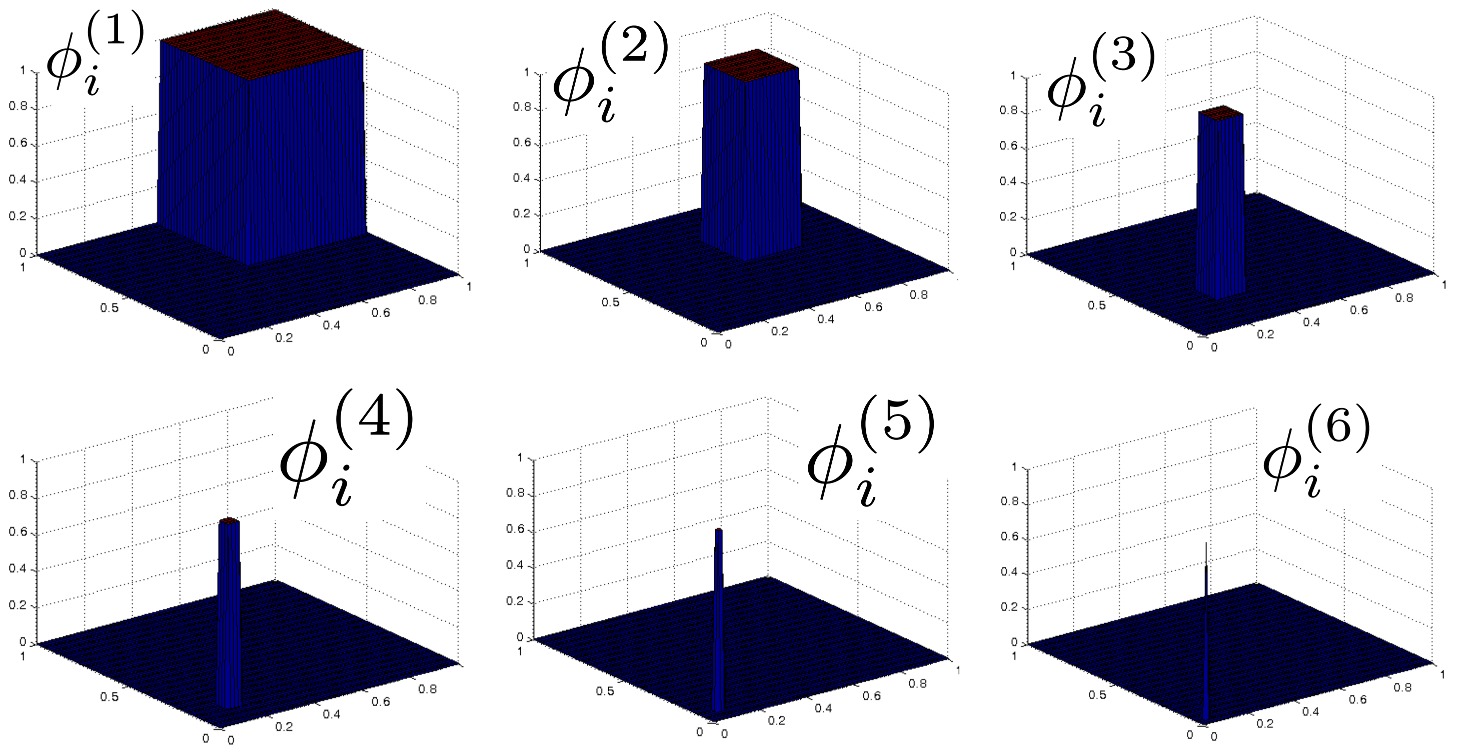
$$\phi_i^{(k)} \in H^{-s}(\Omega) \text{ with } k \in \{1,\ldots,q\}$$

$$\boxed{\phi_i^{(k)} = \sum_j \pi_{i,j}^{(k,k+1)} \phi_j^{(k+1)}}$$

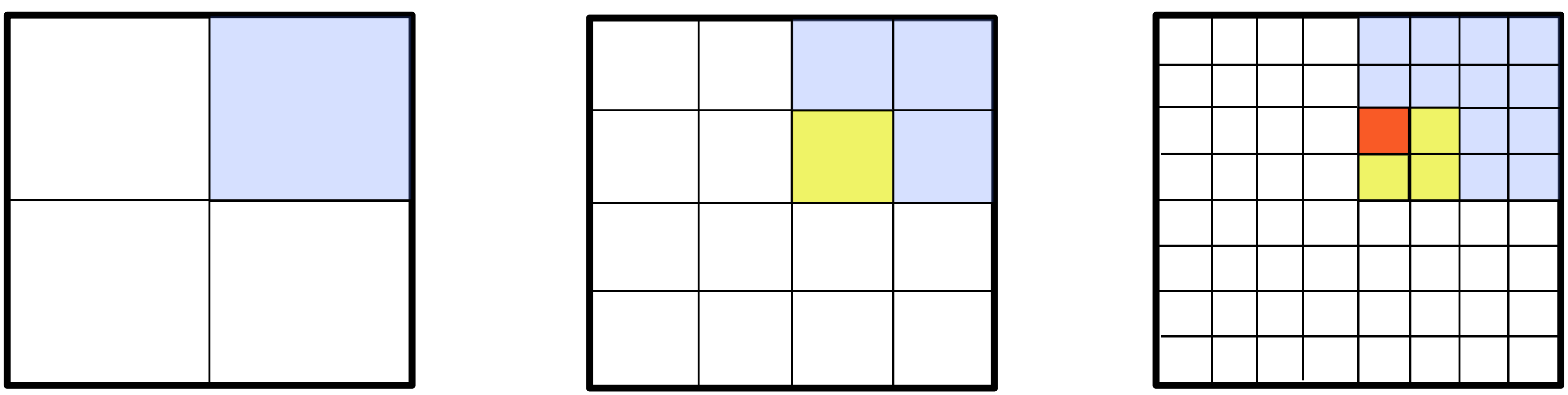
## Example

$\phi_i^{(k)}$ : Weighted indicator functions of a hierarchical nested partition of $\Omega$ of resolution $2^{-k}$



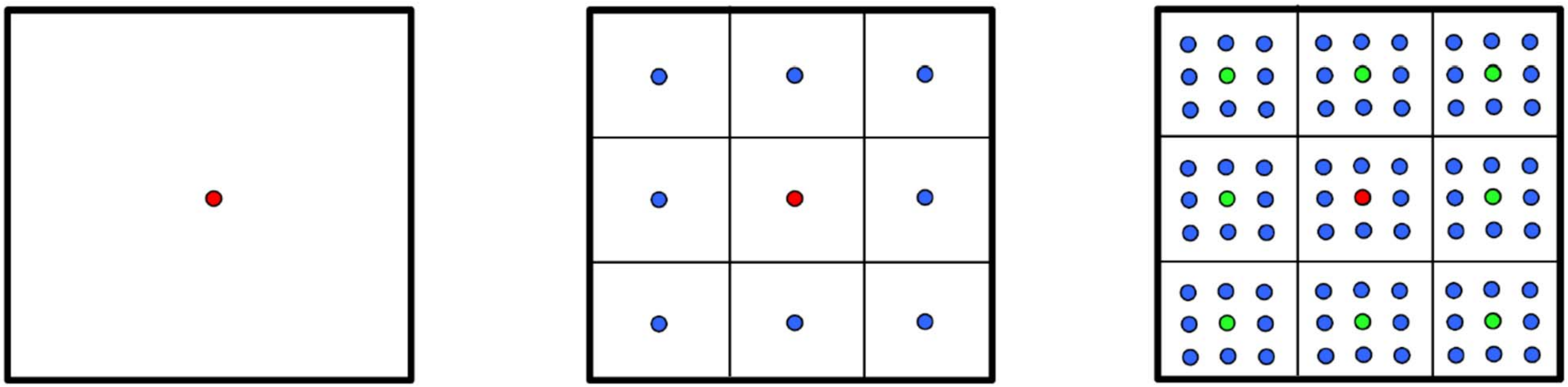



$\phi_i^{(k)}$ : Weighted indicator functions of a hierarchical nested partition of $\Omega$ of resolution $2^{-k}$



$\boxed{H_0^s(\Omega)}$ $\qquad s > d/2$

$\phi_i^{(k)}$ : Subsampled delta Dirac functions



[Schäfer, Sullivan, Owhadi. 2017]: Compression, inversion, and approximate PCA of dense kernel matrices at near-linear computational complexity.

[ H. Owhadi and C. Scovel. Universal Scalable Robust Solvers from Computational Information Games and fast eigenspace adapted Multiresolution Analysis 2017. arXiv:1703.10761]

**Player I**

Chooses

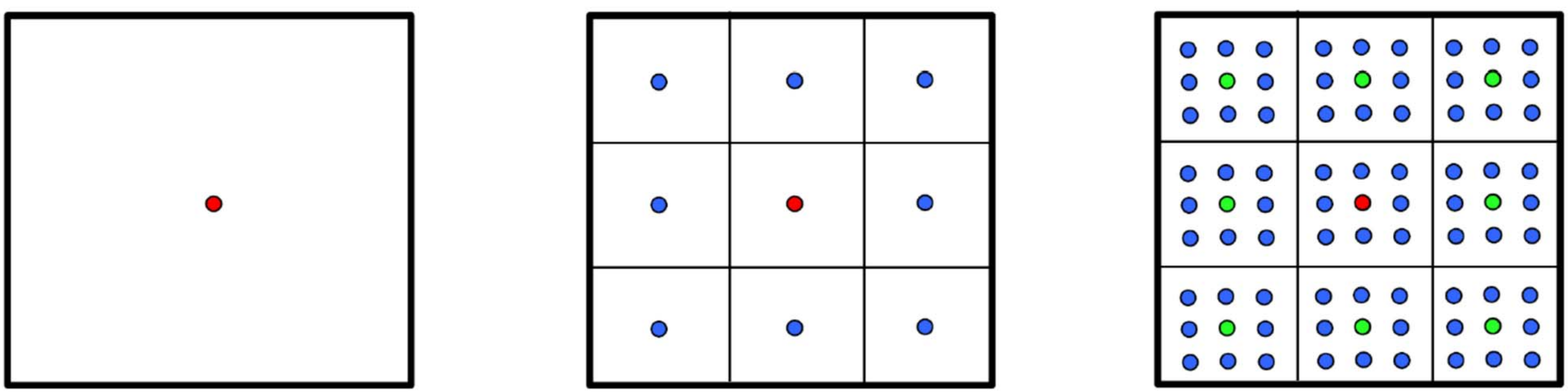
$u \in H_0^s(\Omega)$

**Player II**

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

Must predict

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



Use relative error in operator norm to define loss

$$\boxed{\|u\|^2 := \int_\Omega u \mathcal{L} u}$$

## Example

**Player I**

Chooses

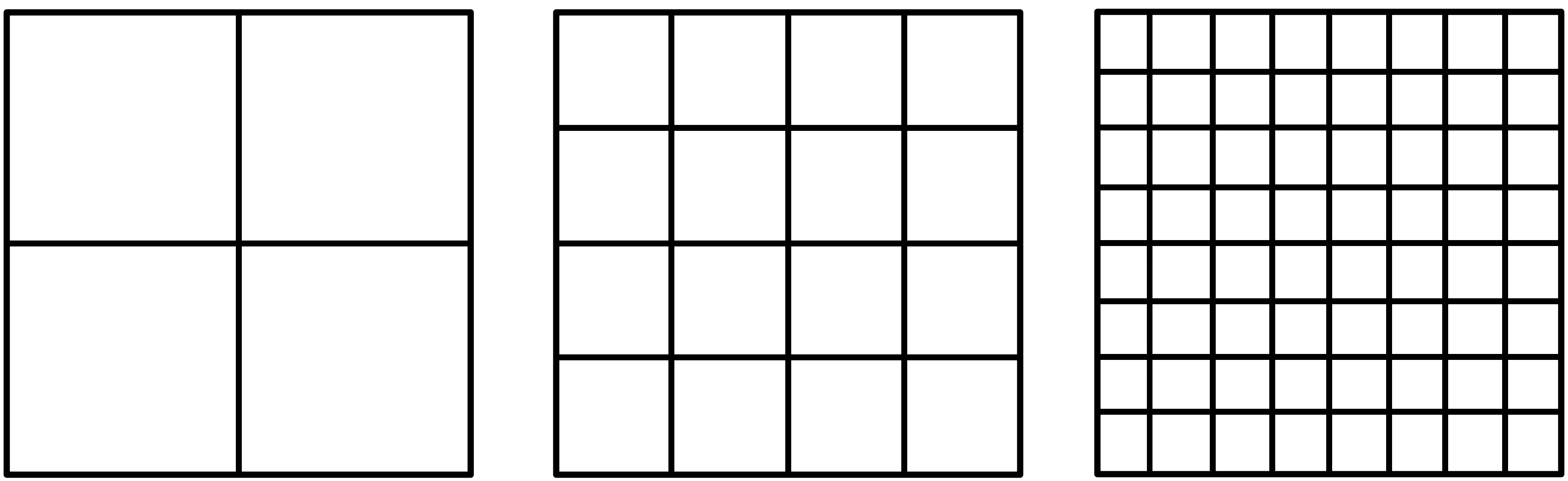
$u \in H_0^s(\Omega)$

**Player II**

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

Must predict

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



## Canonical Gaussian field

$$\xi \sim \mathcal{N}(0, G) \qquad G = \mathcal{L}^{-1}$$

$$\mathbb{E}[\xi(x)] = 0 \qquad \mathrm{Cov}\left(\xi(x), \xi(y)\right) = G(x, y)$$

$$\int_\Omega \phi(x)\xi(x)\,dx \sim \mathcal{N}(0, \int_{\Omega^2} \phi(x)G(x,y)\phi(y)\,dx\,dy)$$

## Player II's bets

$$\boxed{u^{(k)} := \mathbb{E}\left[\xi \middle| [\phi_i^{(k)}, \xi] = [\phi_i^{(k)}, u], \ i \in \mathcal{I}_k\right]}$$

$u^{(k)}$: Martingale

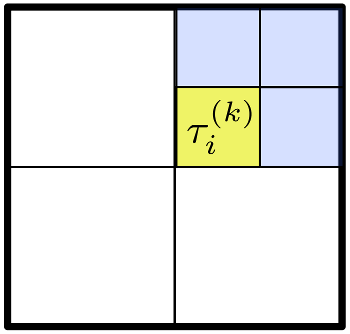
$u^{(k+1)} - u^{(k)}$: Uncorrelated (therefore independent)

**Example**

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

$$H_0^1(\Omega)$$

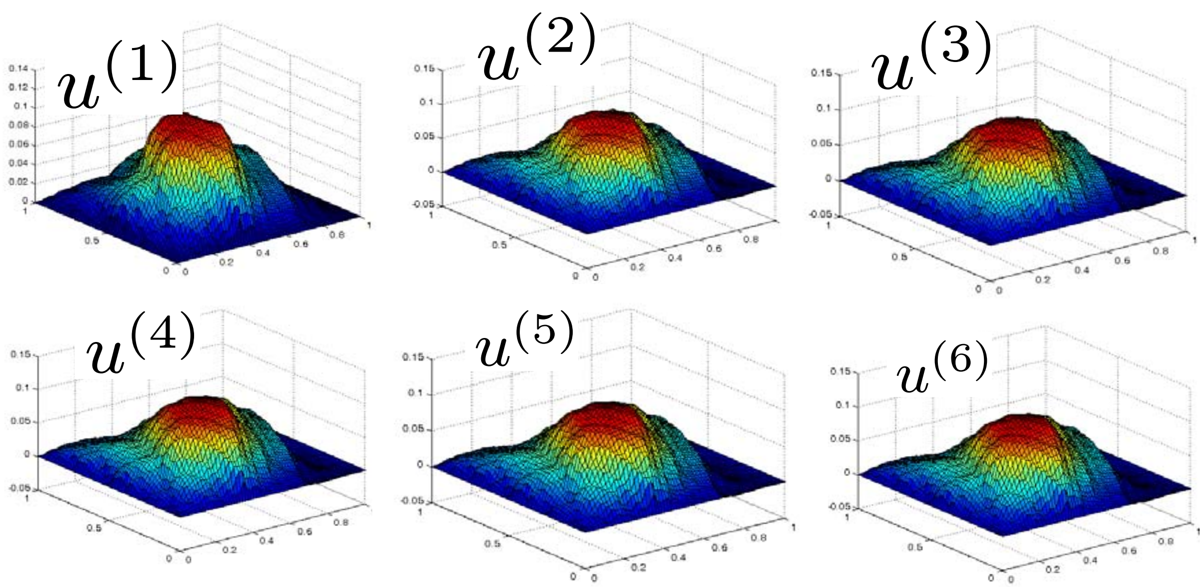
$$\|u\|^2 = \int_\Omega (\nabla u)^T a \nabla u$$




$$\phi_i^{(k)} = 1_{\tau_i^{(k)}} \qquad \operatorname{diam}(\tau_i^{(k)}) \le H^k$$

**Theorem**

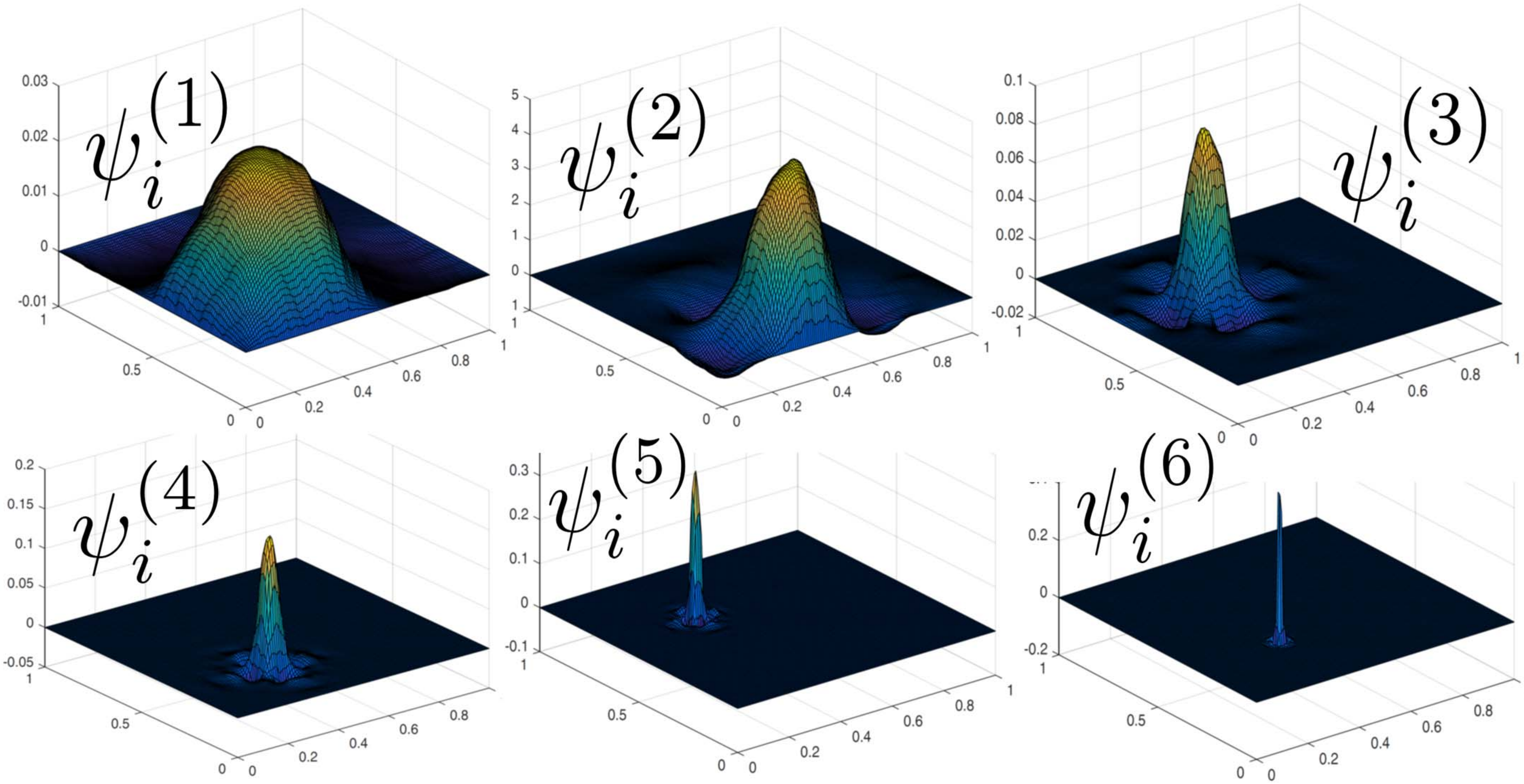
$$\|u - u^{(k)}\| \le \frac{H^k}{\lambda_{\min}(a)} \|g\|_{L^2(\Omega)}$$



# Gamblets

$$\boxed{u^{(k)} = \sum_i [\phi_i^{(k)}, u]\, \psi_i^{(k)}}$$

$$\boxed{\psi_i^{(k)} = \mathbb{E}\Big[\xi \Big| [\phi_l^{(k)}, \xi] = \delta_{i,l},\, l \in \mathcal{I}_k\Big]}$$

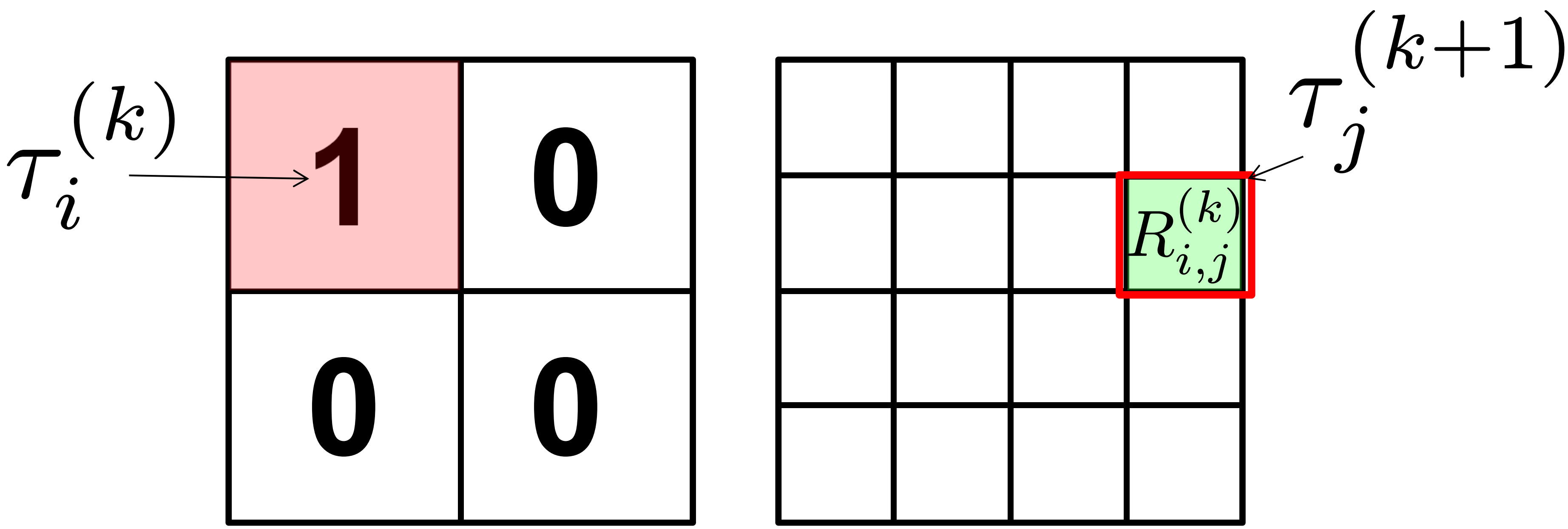


## Gamblets are nested

$$\boxed{\psi_i^{(k)} = \sum_j R_{i,j}^{(k,k+1)} \psi_j^{(k+1)}}$$

## Interpolation/Prolongation operator

$$\boxed{R_{i,j}^{(k,k+1)} = \mathbb{E}\big[[\phi_j^{(k+1)},\xi]\big|[\phi_l^{(k)},\xi] = \delta_{i,l},\, l \in \mathcal{I}_k\big]}$$



$$\mathfrak{V}^{(k)} := \operatorname{span}\{\psi_i^{(k)} \mid i \in \mathcal{I}^{(k)}\}$$

$$\mathfrak{V}^{(k-1)} \subset \mathfrak{V}^{(k)}$$

**Orthogonalized gamblets**

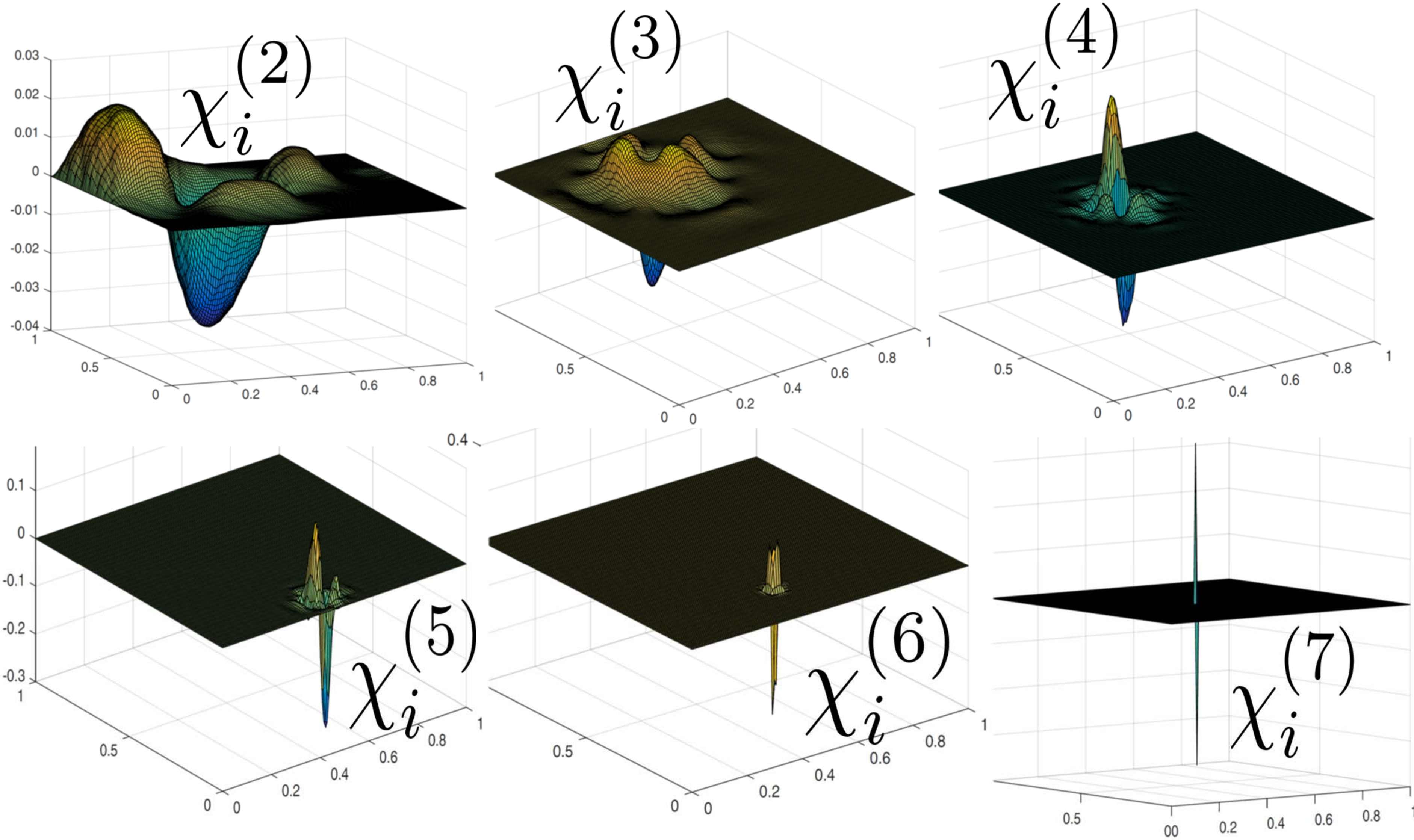
$$\boxed{\chi_i^{(k)} := \sum_{j\in\mathcal{I}^{(k)}} W_{i,j}^{(k)} \psi_j^{(k)}}$$

For $k \geq 2$ $\quad W^{(k)}$: $\quad \operatorname{Img}(W^{(k),T}) = \operatorname{Ker}(\pi^{(k-1,k)})$

$$\operatorname{Cond}(W^{(k)} W^{(k),T}) \leq C$$

$$\boxed{\phi_i^{(k-1)} = \sum_j \pi_{i,j}^{(k-1,k)} \phi_j^{(k)}}$$

$$\chi_i^{(k)} := \sum_{j\in\mathcal{I}^{(k)}} W_{i,j}^{(k)}\,\psi_j^{(k)}$$

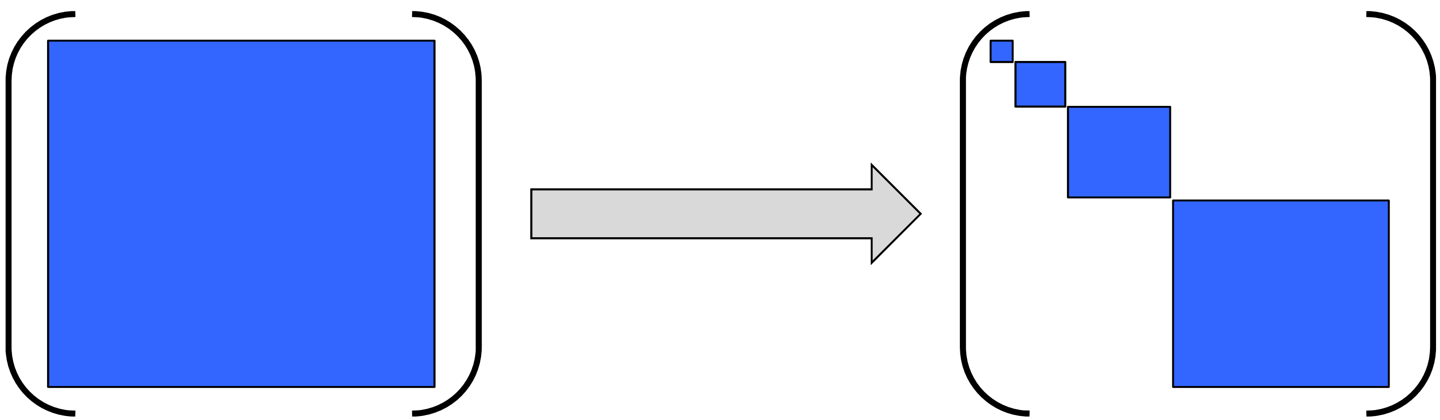


$$\mathfrak{V}^{(k)} := \mathrm{span}\{\psi_i^{(k)} \mid i \in \mathcal{I}^{(k)}\}$$
$$\mathfrak{W}^{(k)} := \mathrm{span}\{\chi_i^{(k)} \mid i \in \mathcal{I}^{(k)}\}$$

**Theorem**

For $k \geq 2$ $\quad \mathfrak{V}^{(k)} = \mathfrak{V}^{(k-1)} \oplus \mathfrak{W}^{(k)}$

$$H_0^s(\Omega) = \mathfrak{V}^{(1)} \oplus \mathfrak{W}^{(2)} \oplus \mathfrak{W}^{(3)} \oplus \cdots$$



$$H_0^s(\Omega) \xrightarrow{\mathcal{L}} H^{-s}(\Omega)$$

$$u \longrightarrow g$$

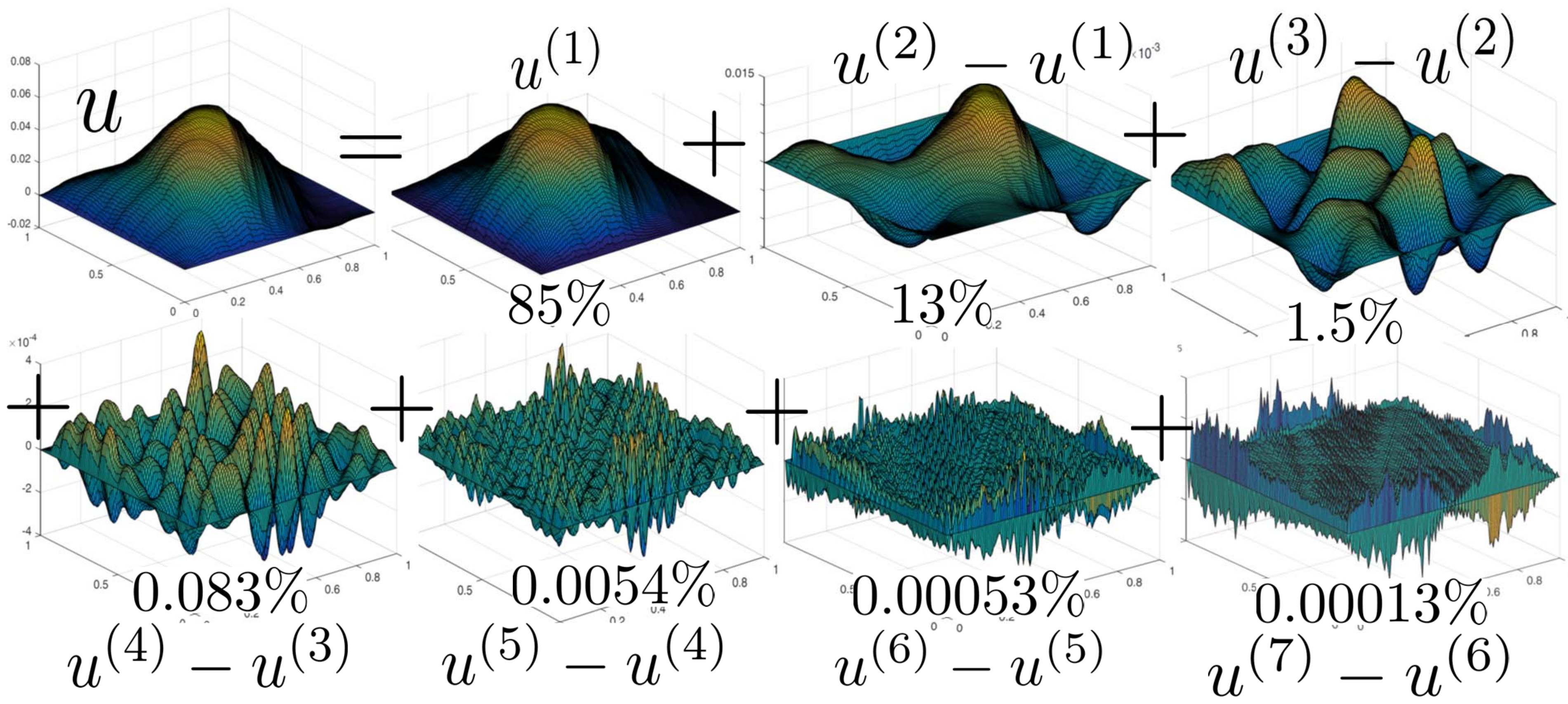
$$\boxed{\mathcal{L}u = g}$$

**Theorem** $u = u^{(1)} + \cdots + \left(u^{(k)} - u^{(k-1)}\right) + \cdots$

$$u^{(k)} - u^{(k-1)} = \textstyle\sum_{i \in \mathcal{I}^{(k)}} w_i^{(k)} \chi_i^{(k)}$$

$$\boxed{B^{(k)} w^{(k)} = g^{(k)}}$$

$$g_i^{(k)} = [g, \chi_i^{(k)}] \qquad B_{i,j}^{(k)} = \langle \chi_i^{(k)}, \chi_j^{(k)} \rangle$$

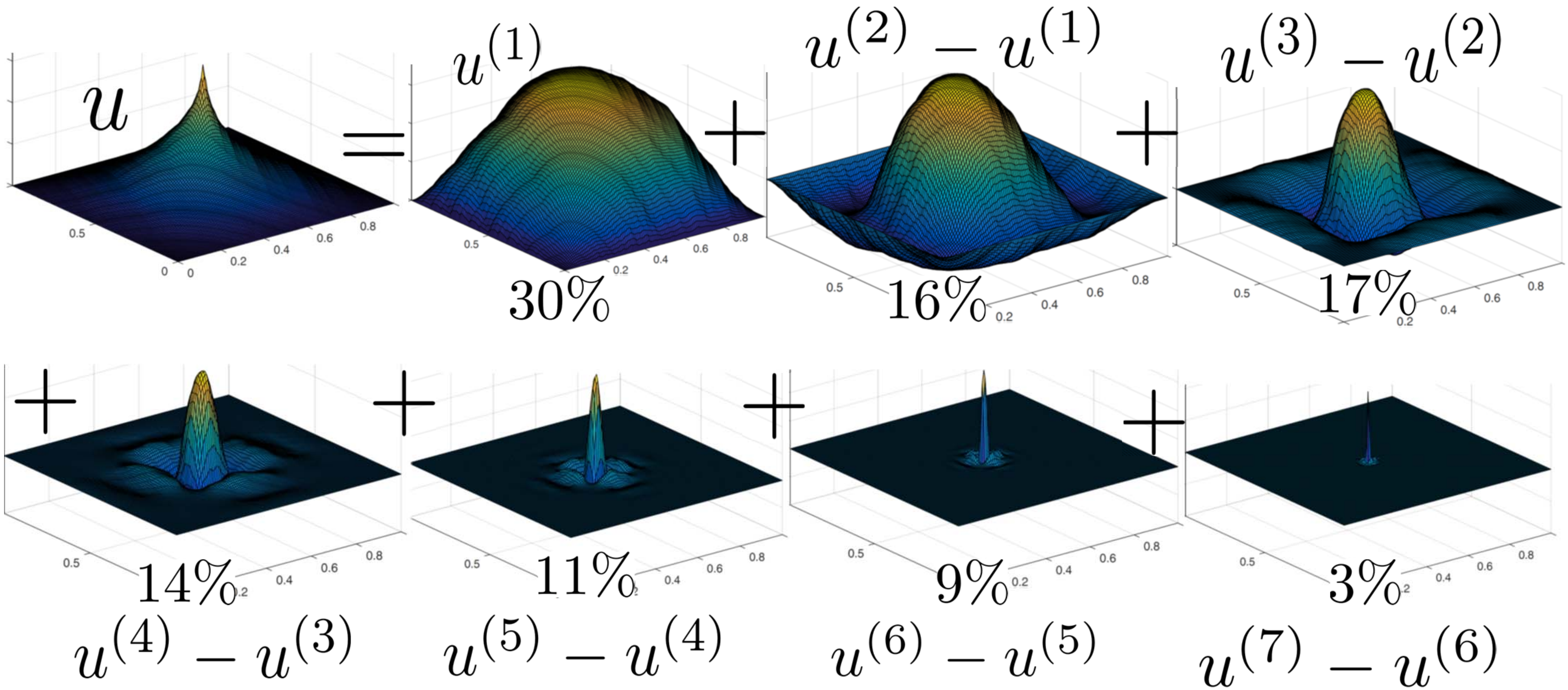



**Energy content**

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

$$g \in C^{\infty}(\Omega)$$

**If r.h.s. is regular we don't need to compute all subbands**




**Energy content**

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

$$g = \delta(x - x_0)$$

# Operator adapted wavelets

## First Generation Wavelets: Signal and imaging processing

[Mallat, 1989] [Daubechies, 1990]

[Coifman, Meyer, and Wickerhauser, 1992]

## First Generation Operator Adapted Wavelets (shift and scale invariant)

[Cohen, Daubechies, Feauveau. Biorthogonal bases of compactly supported wavelets. 1992]

[Beylkin, Coifman, Rokhlin, 1992] [Engquist, Osher, Zhong, 1992]

[Alpert, Beylkin, Coifman, Rokhlin, 1993] [Jawerth, Sweldens, 1993]

[Dahlke, Weinreich, 1993] [Bacry, Mallat, Papanicolaou. 1993]

[Bertoluzza, Maday, Ravel, 1994] [Vasilyev, Paolucci, 1996]

[Dahmen, Kunoth, 2005] [Stevenson, 2009]

## Lazy wavelets (Multiresolution decomposition of solution space)

[Yserentant. Multilevel splitting, 1986]

[Bank, Dupont, Yserentant. Hierarchical basis multigrid method. 1988]

# Operator adapted wavelets

**Second Generation Operator Adapted Wavelets**

[Sweldens. The lifting scheme, 1998] [Dorobantu - Engquist. 1998]

[Vassilevski, Wang. Stabilizing the hierarchical basis, 1997]

[Carnicer, Dahmen, Peña, 1996] [Lounsbery, DeRose, Warren, 1997]

[Vassilevski, Wang. Stabilizing hierarchical basis, 1997-1998]

[Barinka, Barsch, Charton, Cohen, Dahlke, Dahmen, Urban, 2001]

[Cohen, Dahmen, DeVore, 2001] [Chiavassa, Liandrat, 2001]

[Dahmen, Kunoth, 2005] [Schwab, Stevenson, 2008]

[Sudarshan, 2005] [Engquist, Runborg, 2009] [Yin, Liandrat, 2016]

**We want**

1. **Scale-orthogonal wavelets with respect to operator scalar product (leads to block-diagonalization)**
2. **Operator to be well conditioned within each subband**
3. **Wavelets need to be localized (compact support or exp. decay)**

# Wannier functions

[Wannier. Dynamics of band electrons in electric and magnetic fields. 1962]

[Kohn. Analytic properties of Bloch waves and Wannier functions, 1959]

[Marzari, Vanderbilt. Maximally localized generalized Wannier functions for composite energy bands. 1997]

[E, Tiejun, Jianfeng. Localized bases of eigensubspaces and operator compression, 2010]

[Vidvuds, Lai, Caflisch, Osher, Compressed modes for variational problems in mathematics and physics, 2013]

[Owhadi, Multiresolution operator decomposition, SIREV 2017]

[Owhadi, Zhang, gamblets for hyperbolic and parabolic PDEs, 2016]

[Hou, Qin, Zhang, A sparse decomposition of low rank symmetric positive semi-definite matrices, 2016]

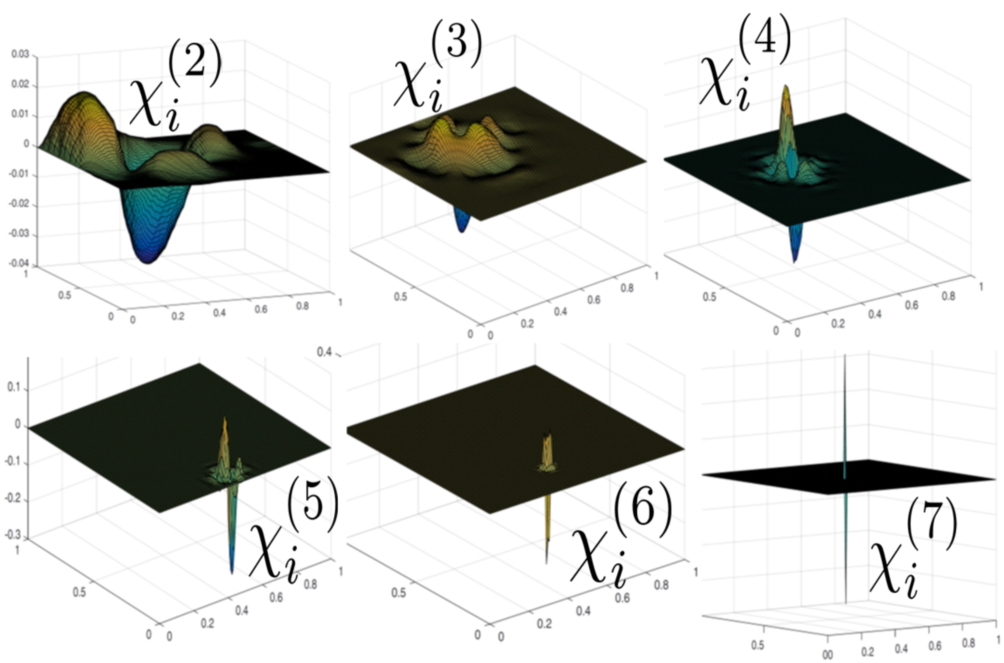
[Hou, Zhang, Sparse operator compression of elliptic operators. 2017]

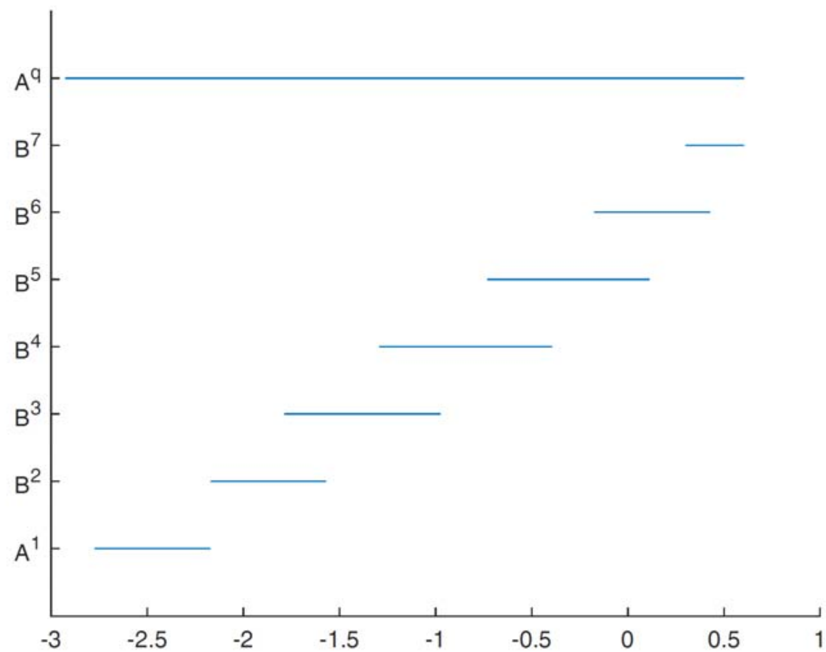
## Gamblets are operator adapted wavelets and Wannier functions

$$\mathfrak{V}^{(k)} := \operatorname{span}\{\psi_i^{(k)} \mid i \in \mathcal{I}^{(k)}\}$$

$$\mathfrak{W}^{(k)} := \operatorname{span}\{\chi_i^{(k)} \mid i \in \mathcal{I}^{(k)}\}$$

$$\boxed{H_0^s(\Omega) = \mathfrak{V}^{(1)} \oplus \mathfrak{W}^{(2)} \oplus \mathfrak{W}^{(3)} \oplus \cdots}$$




- **Scale orthogonal in operator scalar product.**
- **Operator is well conditionned within each sub-band (localized in spectrum)**
- **Decay exponentially fast (localized in space)**

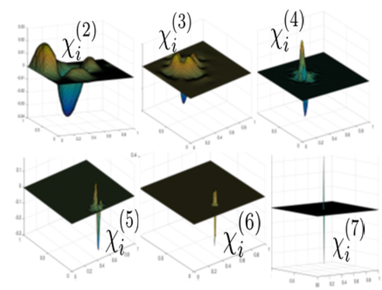
# Sparse factorization of the Green’s function

$$G = \mathcal{L}^{-1}$$

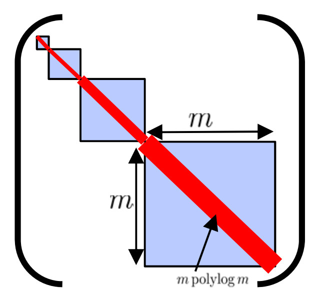
## Theorem

$$\boxed{G(x,y) = \sum_k \sum_{i,j} \chi_i^{(k)}(x) B_{i,j}^{(k),-1} \chi_j^{(k)}(y)}$$

$\chi_i^{(k)}$: Localized (exponentially decaying) to subdomain of resolution $2^{-k}$




$\mathcal{B}^{(k),-1}$: Uniformly well-conditionned and sparse




## Complexity

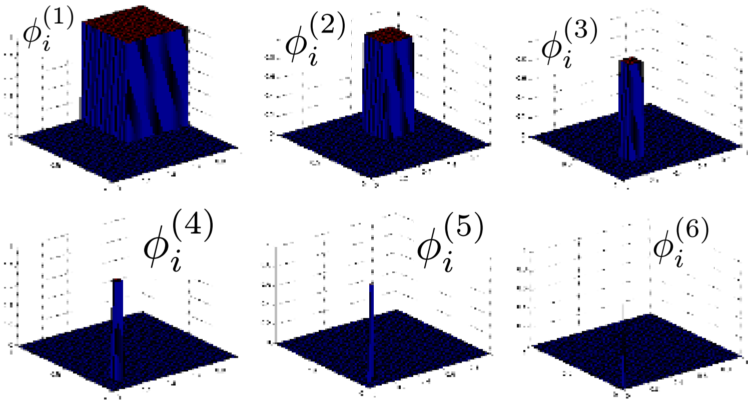
$\mathcal{O}(N \log^{3d} N)$

$$H_0^s(\Omega) \xrightarrow{\qquad\qquad \mathcal{L} \qquad\qquad} H^{-s}(\Omega)$$

$$\cup$$

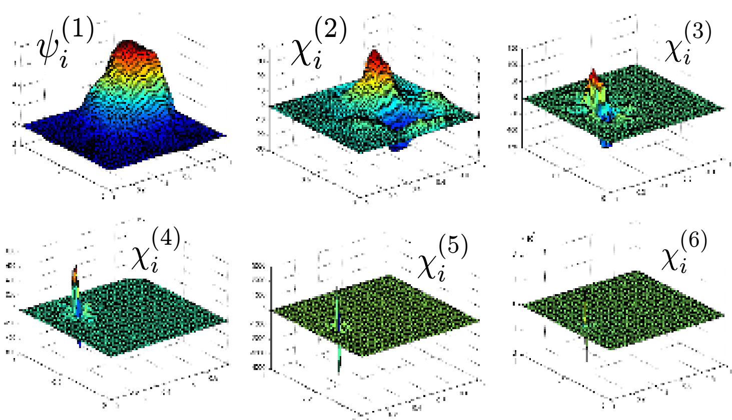
$$L^2(\Omega)$$

## The method




Haar-wavelet decomposition of $L^2(\Omega) \to H^{-s}(\Omega)$

Gamblet transform



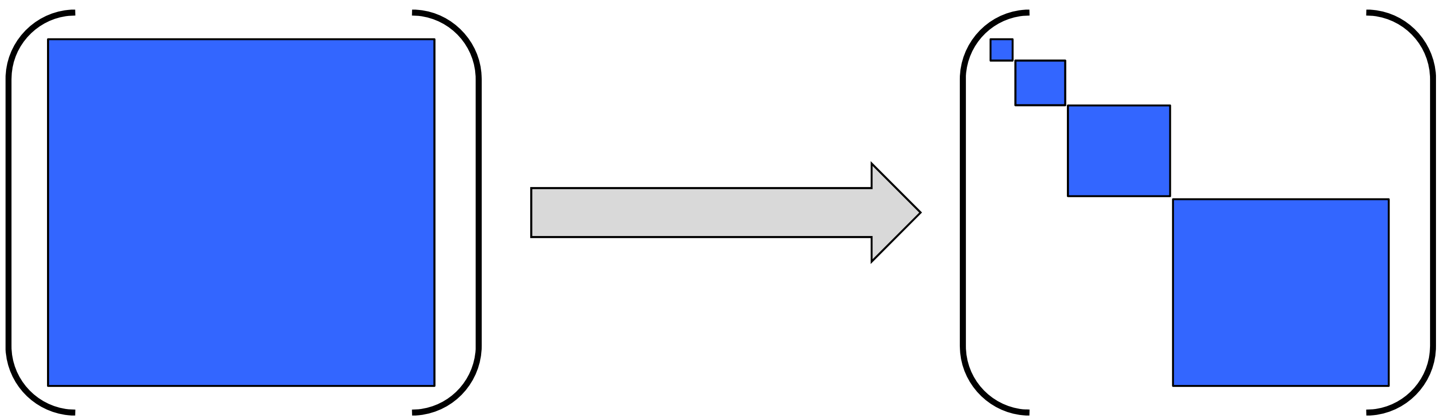

Multi-resolution decomposition of $H_0^s(\Omega) \to H^{-s}(\Omega)$

$$H_0^s(\Omega) \xrightarrow{\quad\mathcal{L}\quad} H^{-s}(\Omega)$$

$$\boxed{H_0^s(\Omega) = \mathfrak{W}^{(1)} \oplus \mathfrak{W}^{(2)} \oplus \mathfrak{W}^{(3)} \oplus \cdots}$$

$$\boxed{\|u\|^2 := \int_\Omega u\mathcal{L}u}$$

**Gamblet transform**

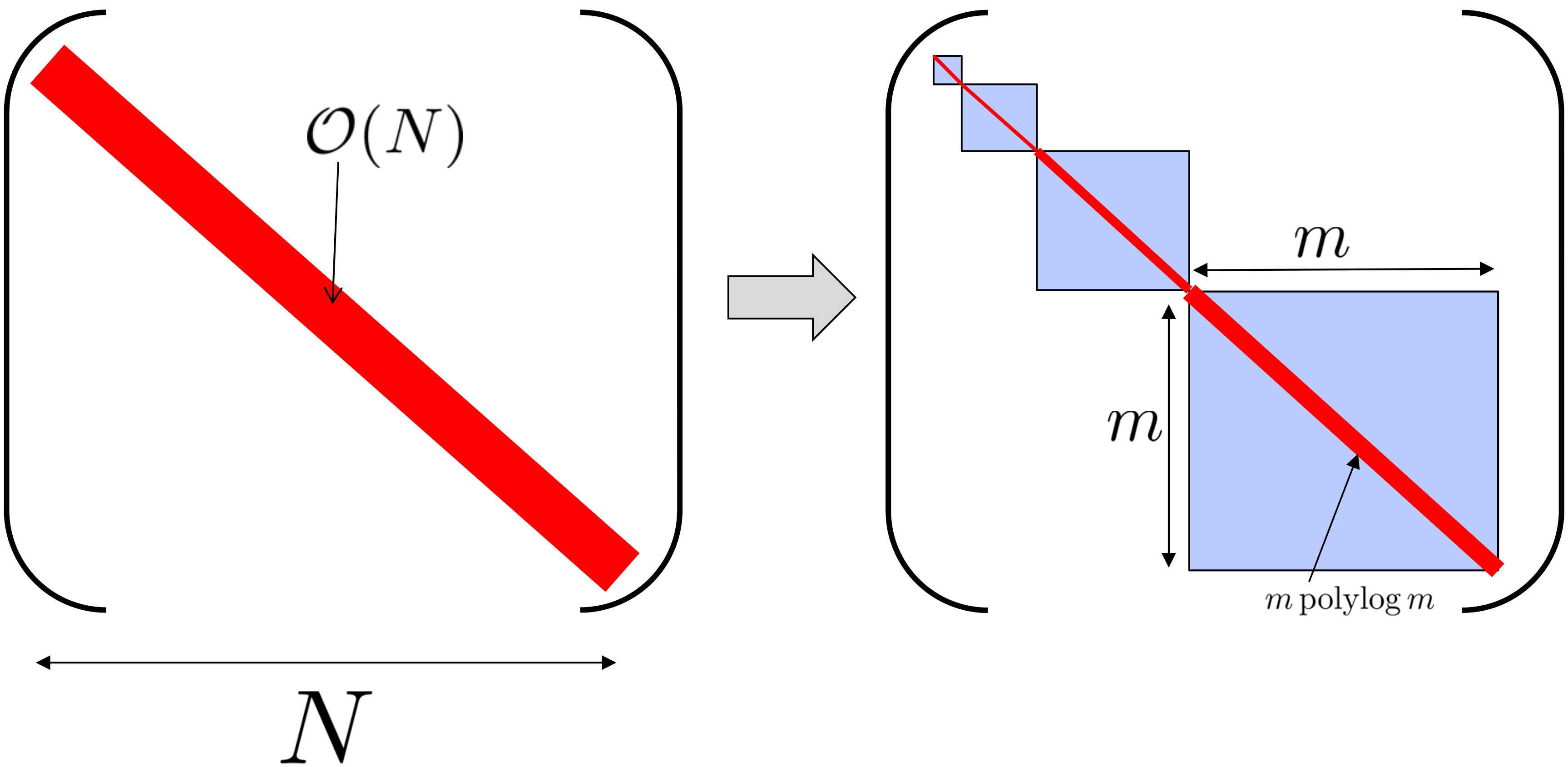


**Theorem**

Blocks have uniformly bounded condition numbers

**Theorem** If $\mathcal{L}$ is local $\int_\Omega u\mathcal{L}v = 0$ if $u$ and $v$ have disjoint supports

then blocks are uniformly sparse




**Complexity** $\mathcal{O}(N \log^{3d} N)$ (Block diagonalization)

$\mathcal{O}(N \log^{d+1} N)$ (Solving $\mathcal{L}u = g$)

## Gamblet Transform/Solve

For $i \in \mathcal{I}^{(q)}$, $\psi_i^{(q)} = \varphi_i$

For $i \in \mathcal{I}^{(q)}$, $g_i^{(q)} = [g, \psi_i^{(q)}]$

For $i, j \in \mathcal{I}^{(q)}$, $A_{i,j}^{(q)} = \langle \psi_i^{(q)}, \psi_j^{(q)} \rangle$

**for** $k = q$ to $2$ **do**

  $B^{(k)} = W^{(k)} A^{(k)} W^{(k),T}$

  $w^{(k)} = B^{(k),-1} W^{(k)} g^{(k)}$

  For $i \in \mathcal{J}^{(k)}$, $\chi_i^{(k)} = \sum_{j \in \mathcal{I}^{(k)}} W_{i,j}^{(k)} \psi_j^{(k)}$

  $u^{(k)} - u^{(k-1)} = \sum_{i \in \mathcal{J}^{(k)}} w_i^{(k)} \chi_i^{(k)}$

  $D^{(k,k-1)} = -B^{(k),-1} W^{(k)} A^{(k)} \bar{\pi}^{(k,k-1)}$

  $R^{(k-1,k)} = \bar{\pi}^{(k-1,k)} + D^{(k-1,k)} W^{(k)}$

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

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

  $g^{(k-1)} = R^{(k-1,k)} g^{(k)}$

**end for**

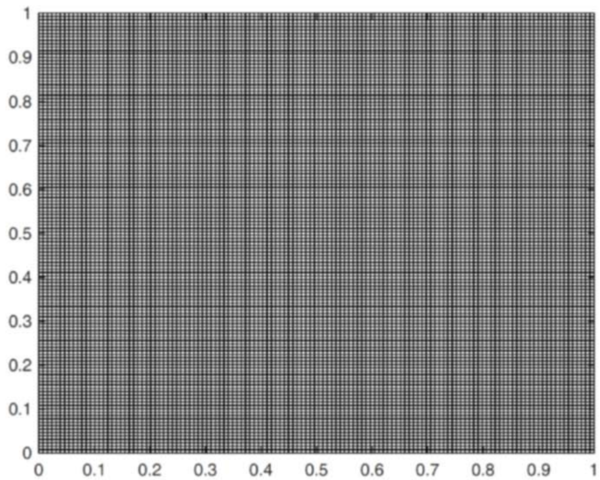
$U^{(1)} = A^{(1),-1} g^{(1)}$

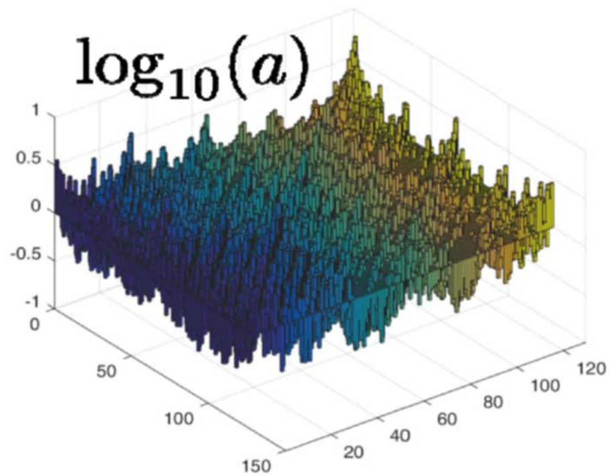
$u^{(1)} = \sum_{i \in \mathcal{I}^{(1)}} U_i^{(1)} \psi_i^{(1)}$

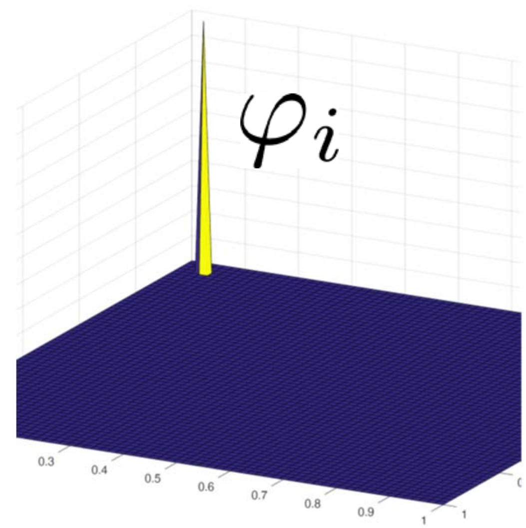
$u = u^{(1)} + (u^{(2)} - u^{(1)}) + \cdots + (u^{(q)} - u^{(q-1)})$

# Example

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



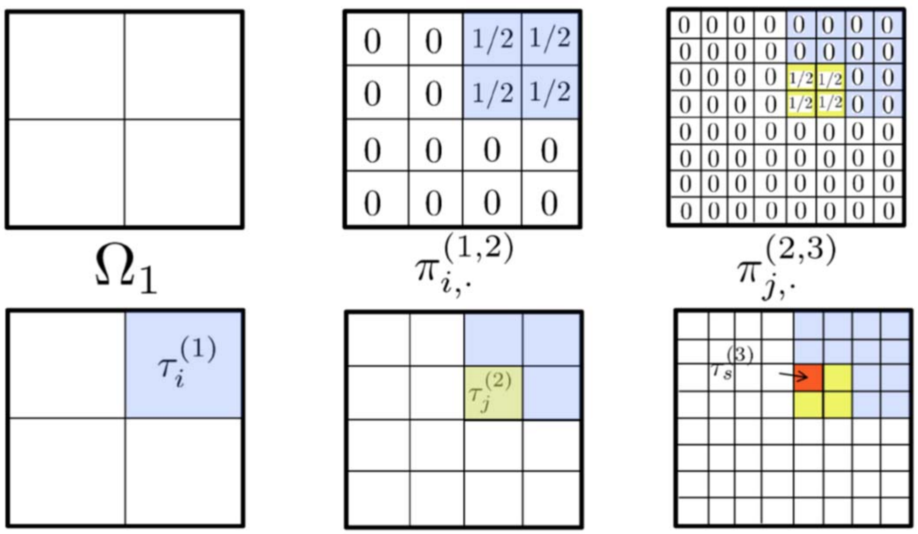




$$\mathcal{B} = \{\varphi_i | i \in \mathcal{I}\} \subset H_0^1(\Omega)$$

$$\|u\|^2 := \int_\Omega (\nabla u)^T a \nabla u$$

# Inputs of the algorithm

$$A_{i,j} = \int_\Omega (\nabla \varphi_i)^T a \nabla \varphi_j$$




$$\pi^{(k-1,k)}$$

$$\pi^{(k-1,k)}(\pi^{(k-1,k)})^T = I^{(k-1)}$$

$$t^{(1)} = l^{(1)} = r^{(1)} = i$$

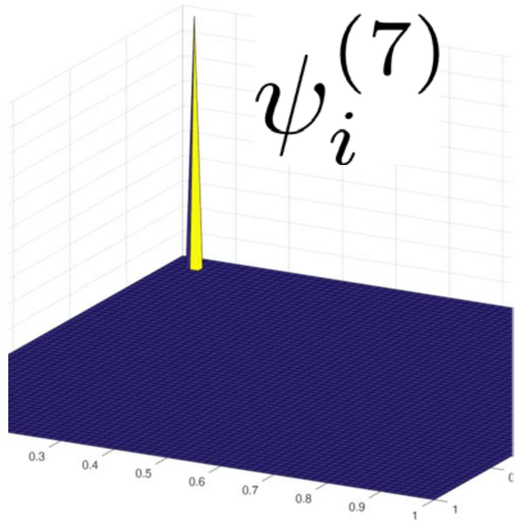
$$\begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & -\frac{2}{\sqrt{6}} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} \\ 0 & 0 & \frac{1}{\sqrt{12}} & -\frac{3}{\sqrt{12}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$W^{(2)}_{t,\cdot} \qquad W^{(2)}_{l,\cdot} \qquad W^{(2)}_{r,\cdot}$$

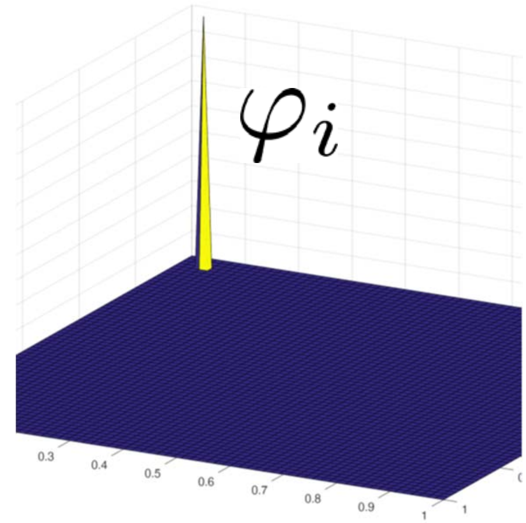
$$W^{(k)}:$$

$$\mathrm{Img}(W^{(k),T}) = \mathrm{Ker}(\pi^{(k-1,k)})$$

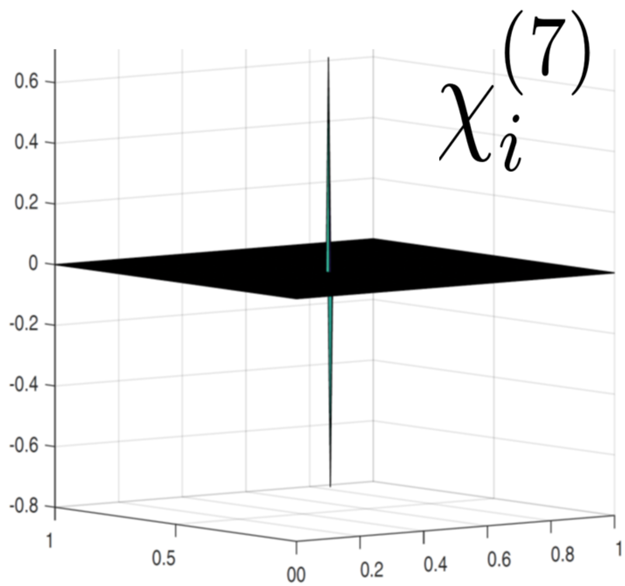
$$W^{(k),T} W^{(k)} = J^{(k)}$$




$$=$$

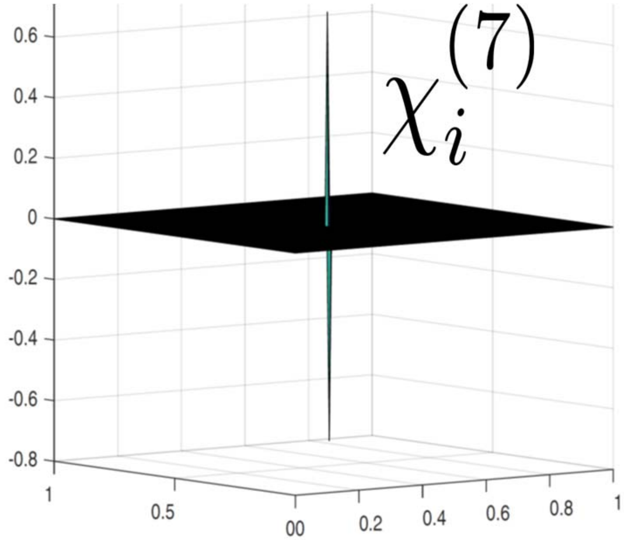



$$\boxed{\chi_i^{(7)} := \sum_{j \in \mathcal{I}^{(7)}} W_{i,j}^{(7)} \psi_j^{(7)}}$$

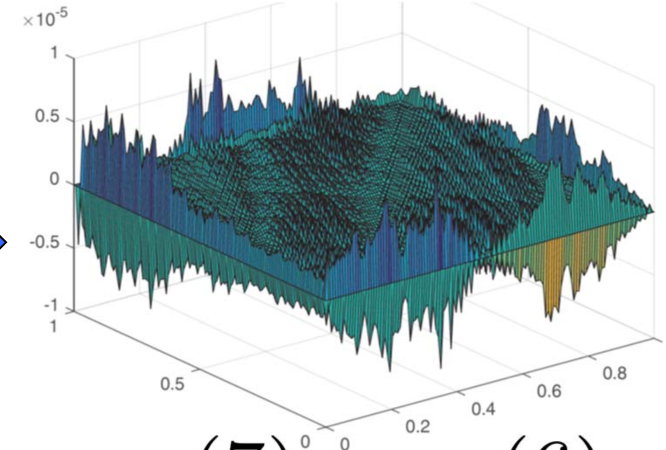



$$t^{(1)} = l^{(1)} = r^{(1)} = i$$

$$\underbrace{\begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{W_{t,\cdot}^{(2)}} \quad \underbrace{\begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & -\frac{2}{\sqrt{6}} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{W_{l,\cdot}^{(2)}} \quad \underbrace{\begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} \\ 0 & 0 & \frac{1}{\sqrt{12}} & -\frac{3}{\sqrt{12}} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{W_{r,\cdot}^{(2)}}$$

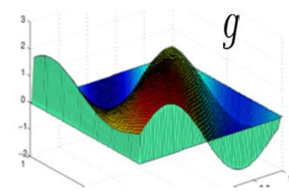


$$\chi_i^{(7)}$$



$$u^{(7)} - u^{(6)}$$

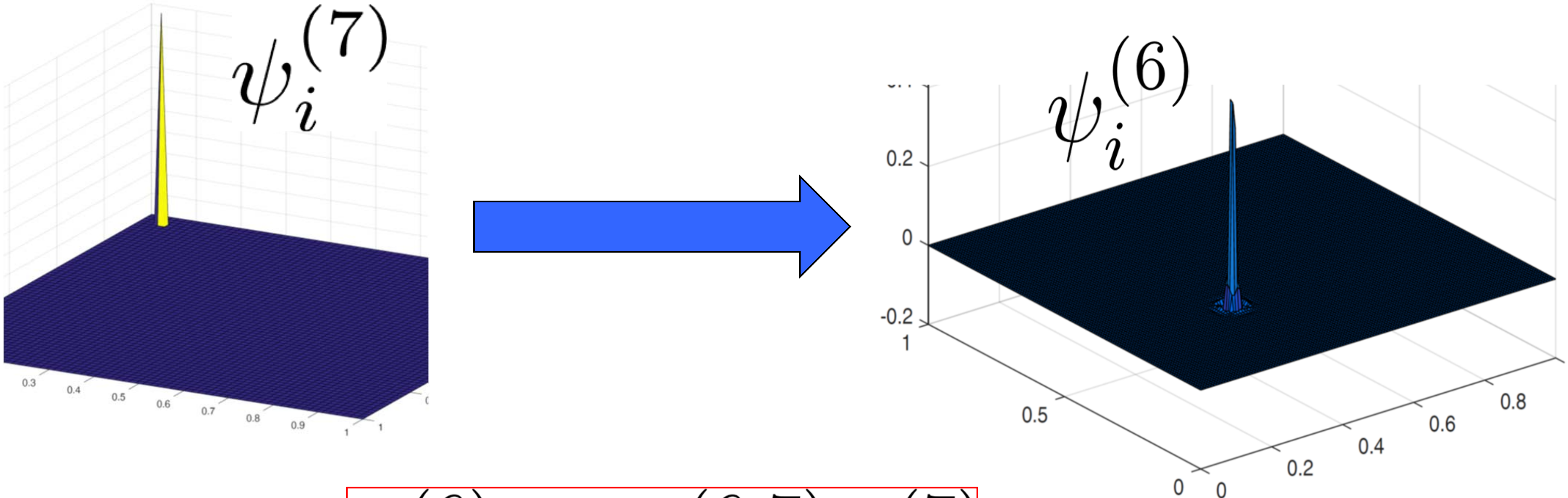
$$\boxed{u^{(7)} - u^{(6)} = \sum_i w_i^{(7)} \chi_i^{(7)}}$$




$$\boxed{B^{(7)} w^{(7)} = W^{(7),T} g^{(7)}}$$

$$g_i^{(7)} = \int_\Omega g \psi_i^{(7)}$$

$$B_{i,j}^{(7)} = \int_\Omega (\nabla \chi_i^{(7)})^T a \nabla \chi_j^{(7)} \qquad B^{(7)} = W^{(7)} A W^{(7),T}$$

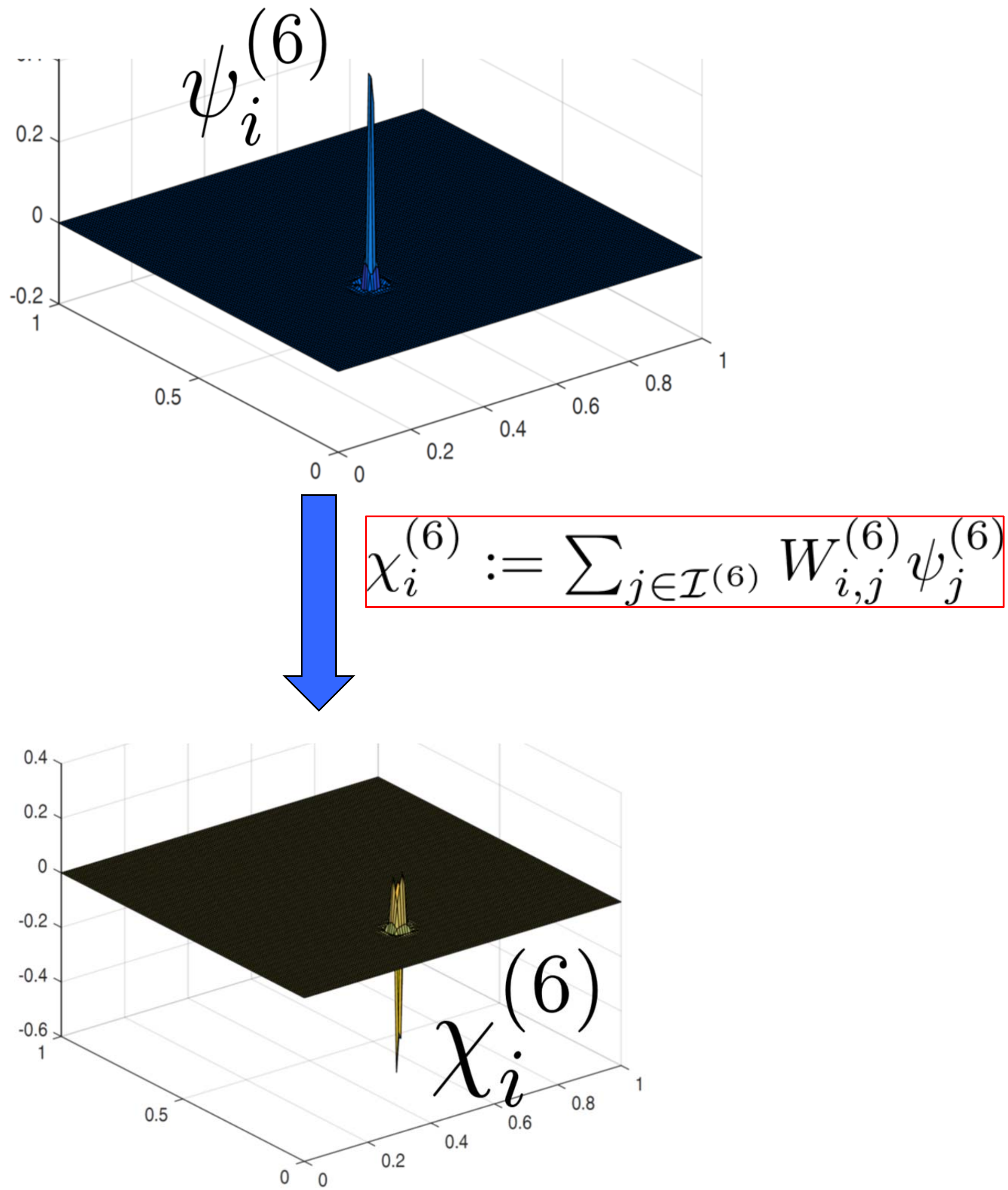



$$\boxed{\psi_i^{(6)} = R_{i,j}^{(6,7)} \psi_j^{(7)}}$$

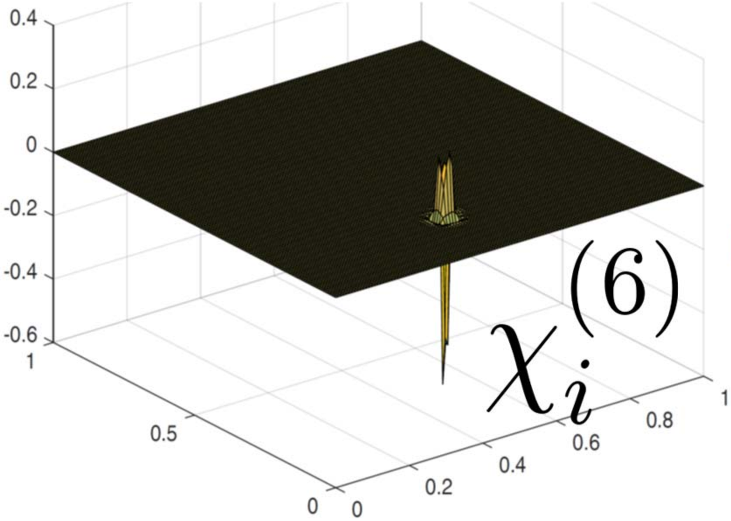
$$A^{(7)} = A$$

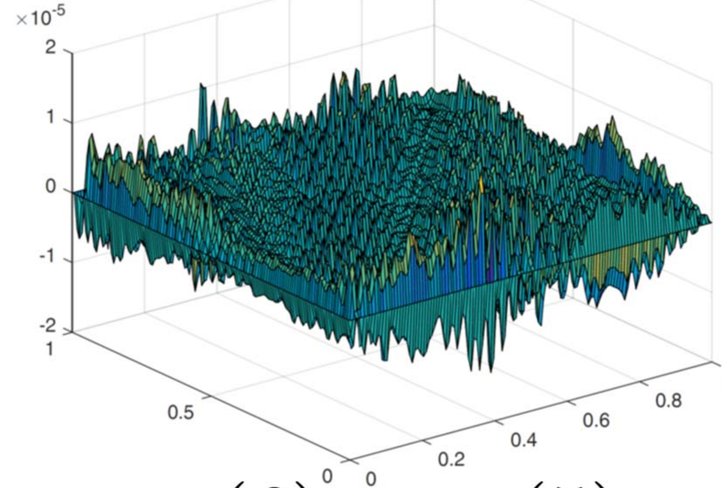
$$R^{(6,7)} = \pi^{(6,7)}(I^{(7)} - A^{(7)} W^{(7),T} B^{(7),-1} W^{(7)})$$

$$\boxed{A^{(6)} = R^{(6,7)} A^{(7)} (R^{(6,7)})^T}$$



$\psi_i^{(6)}$
0.2
0
-0.2
1
0.5
0
0
0.2
0.4
0.6
0.8
1
$\chi_i^{(6)} := \sum_{j\in\mathcal{I}^{(6)}} W_{i,j}^{(6)}\psi_j^{(6)}$
0.4
0.2
0
-0.2
-0.4
-0.6
1
0.5
0
0
0.2
0.4
0.6
0.8
1
$\chi_i^{(6)}$



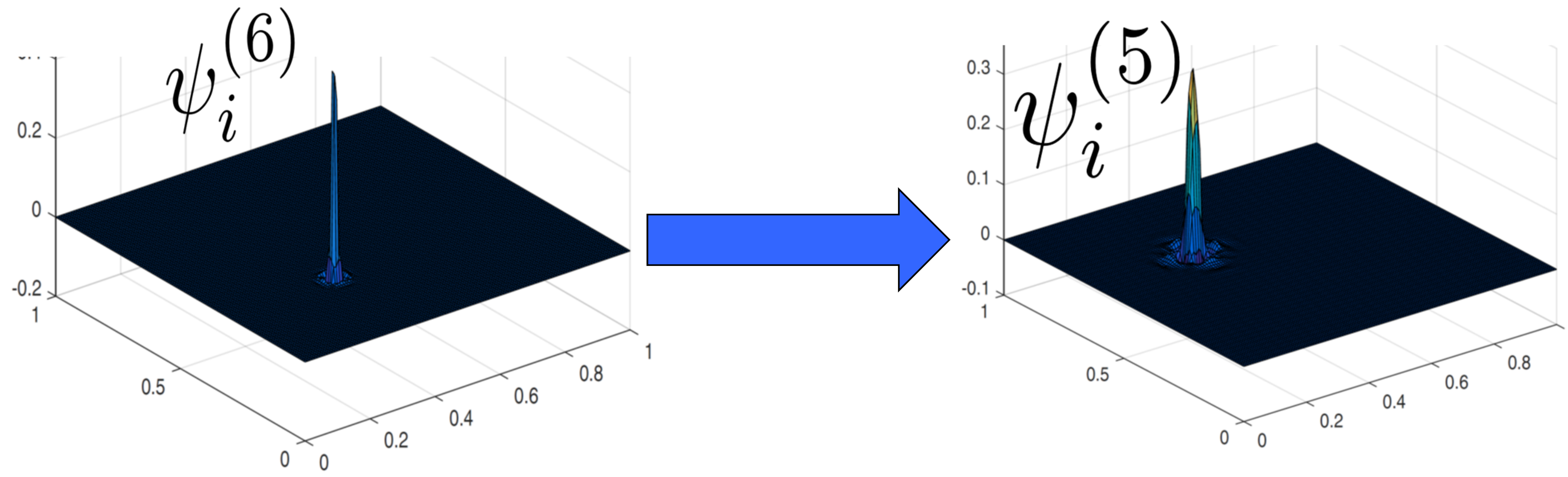




$$u^{(6)} - u^{(5)}$$

$$\boxed{u^{(6)} - u^{(5)} = \sum_i w_i^{(6)} \chi_i^{(6)}}$$

$$\boxed{B^{(6)} w^{(6)} = W^{(6),T} g^{(6)}} \qquad g^{(6)} = R^{(6,7)} g^{(7)}$$

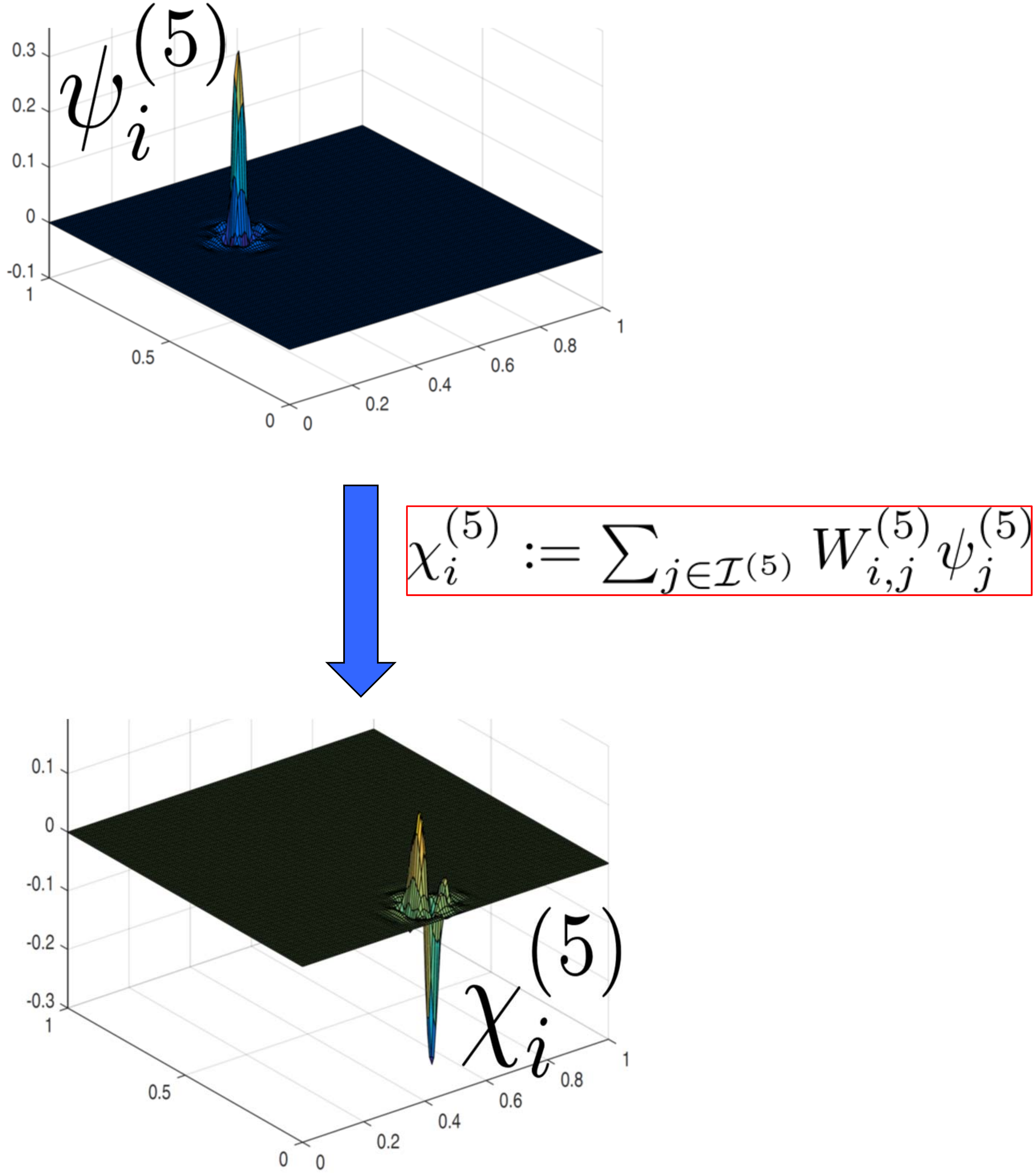
$$B^{(6)} = W^{(6)} A^{(6)} W^{(6),T}$$



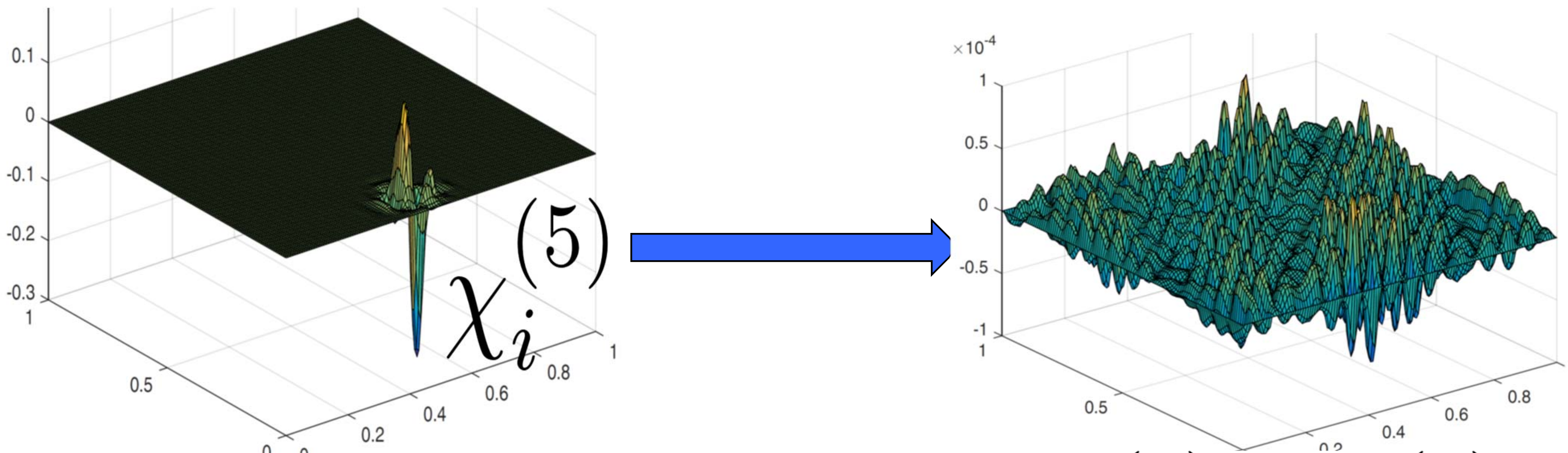

$$\boxed{\psi_i^{(5)} = R_{i,j}^{(5,6)} \psi_j^{(6)}}$$

$$R^{(5,6)} = \pi^{(5,6)} (I^{(6)} - A^{(6)} W^{(6),T} B^{(6),-1} W^{(6)})$$

$$\boxed{A^{(5)} = R^{(5,6)} A^{(6)} (R^{(5,6)})^T}$$



$\psi_i^{(5)}$
0.3
0.2
0.1
0
-0.1
1
0.5
0
0
0.2
0.4
0.6
0.8
1
$\chi_i^{(5)} := \sum_{j\in\mathcal{I}^{(5)}} W_{i,j}^{(5)}\,\psi_j^{(5)}$
0.1
0
-0.1
-0.2
-0.3
1
0.5
0
0
0.2
0.4
0.6
0.8
1
$\chi_i^{(5)}$

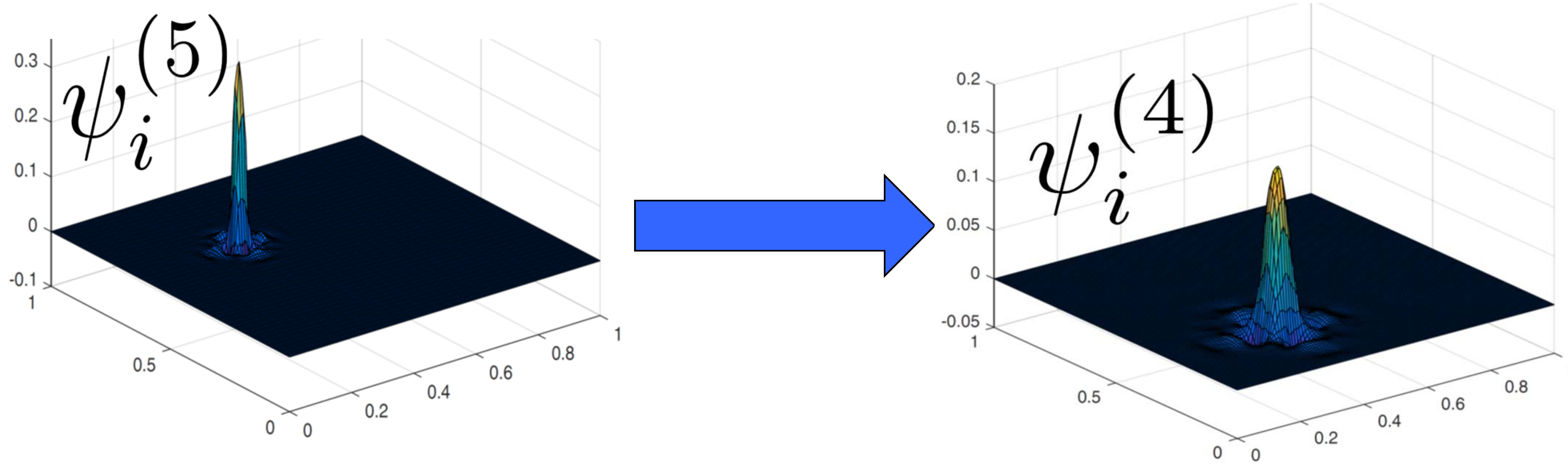



$$u^{(5)} - u^{(4)}$$

$$\boxed{u^{(5)} - u^{(4)} = \sum_i w_i^{(5)} \chi_i^{(5)}}$$

$$\boxed{B^{(5)} w^{(5)} = W^{(5),T} g^{(5)}} \qquad g^{(5)} = R^{(5,6)} g^{(6)}$$

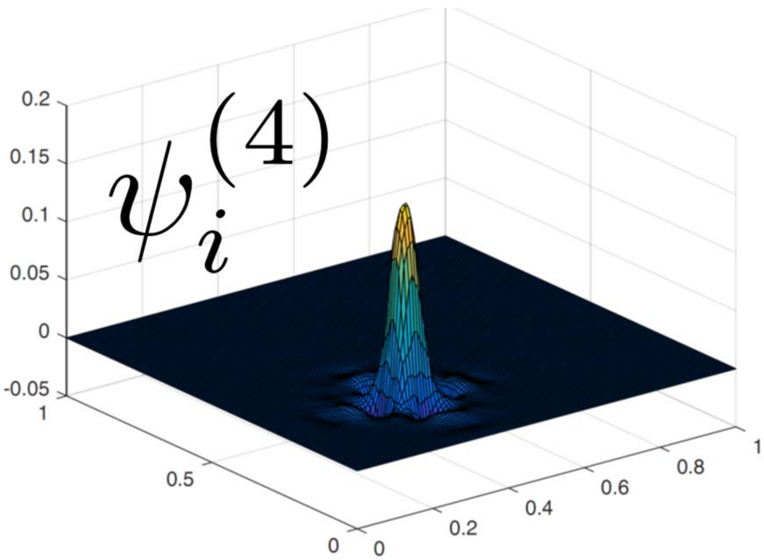
$$B^{(5)} = W^{(5)} A^{(5)} W^{(5),T}$$



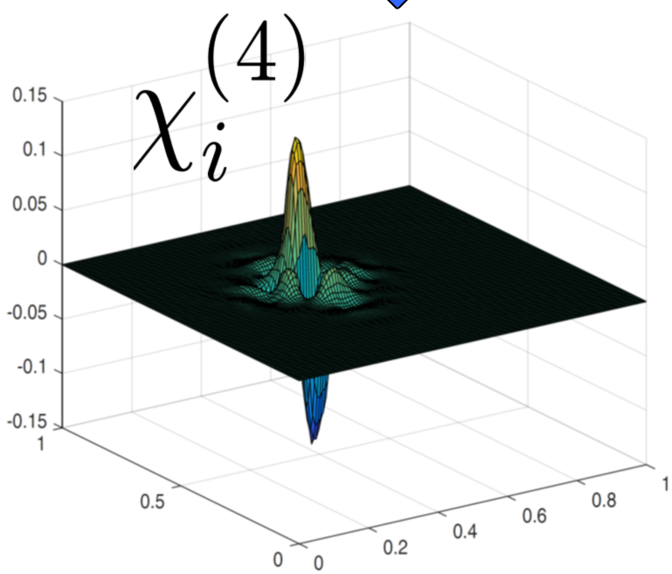

$$\boxed{\psi_i^{(4)} = R_{i,j}^{(4,5)} \psi_j^{(5)}}$$

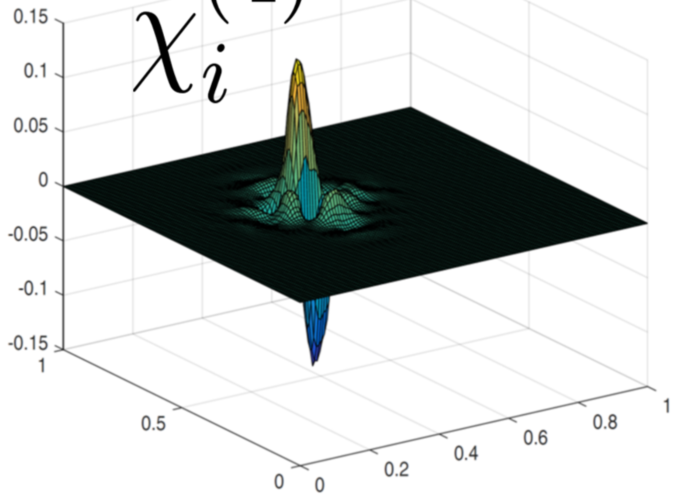
$$R^{(4,5)} = \pi^{(4,5)} \left( I^{(5)} - A^{(5)} W^{(5),T} B^{(5),-1} W^{(5)} \right)$$

$$\boxed{A^{(4)} = R^{(4,5)} A^{(5)} \left( R^{(4,5)} \right)^T}$$

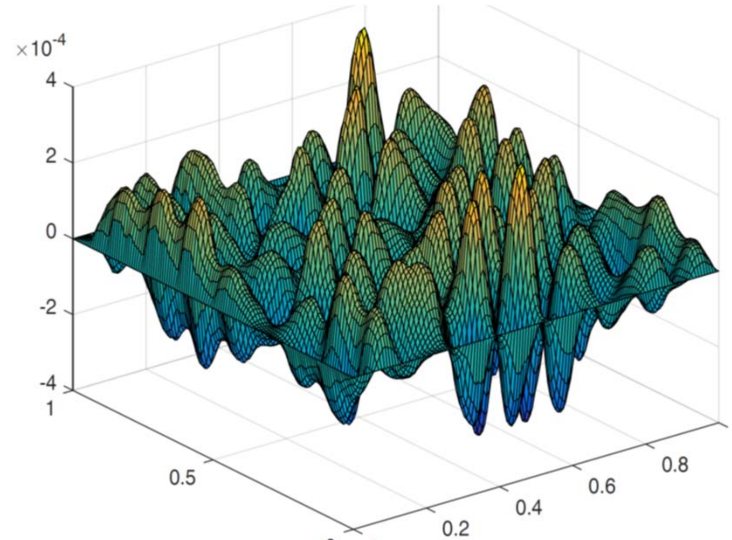



$$\chi_i^{(4)} := \sum_{j\in\mathcal{I}^{(4)}} W_{i,j}^{(4)}\,\psi_j^{(4)}$$






$$\chi_i^{(4)}$$

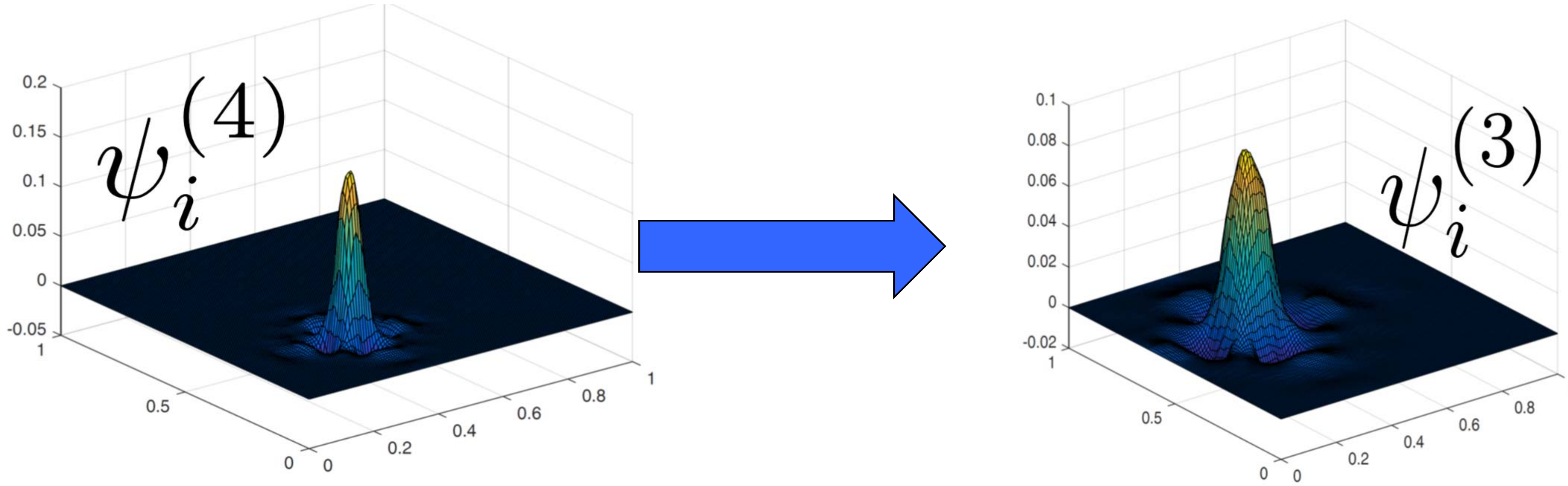



$$u^{(4)} - u^{(3)}$$

$$\boxed{u^{(4)} - u^{(3)} = \sum_i w_i^{(4)} \chi_i^{(4)}}$$

$$\boxed{B^{(4)} w^{(4)} = W^{(4),T} g^{(4)}} \qquad g^{(4)} = R^{(4,5)} g^{(5)}$$

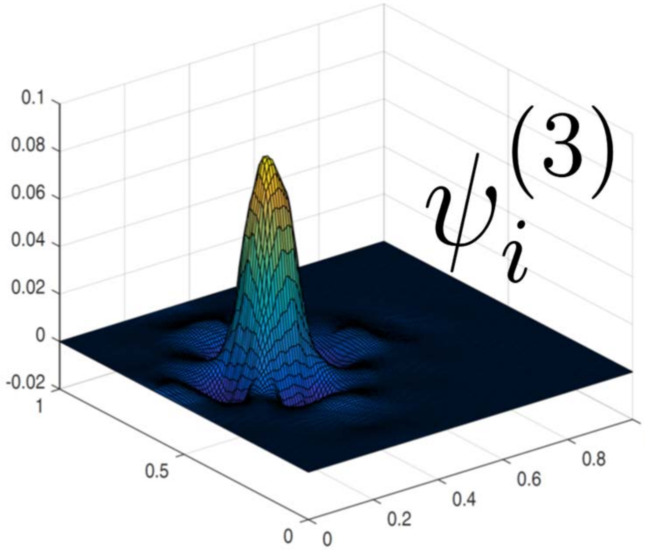
$$B^{(4)} = W^{(4)} A^{(4)} W^{(4),T}$$



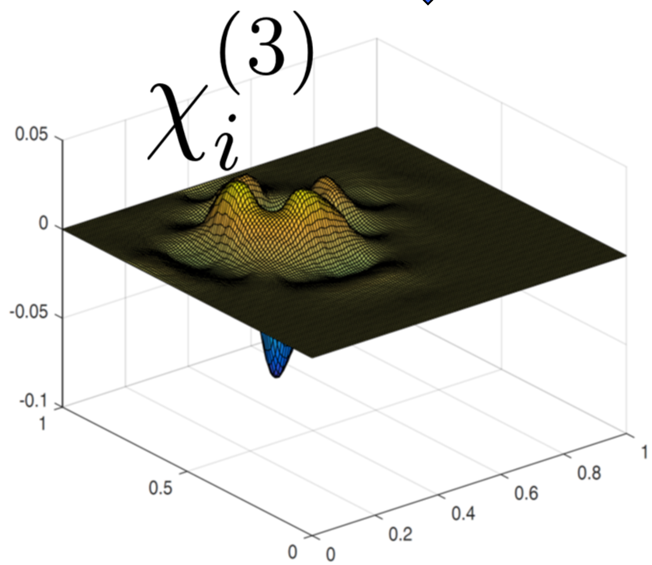

$$\boxed{\psi_i^{(3)} = R_{i,j}^{(3,4)} \psi_j^{(4)}}$$

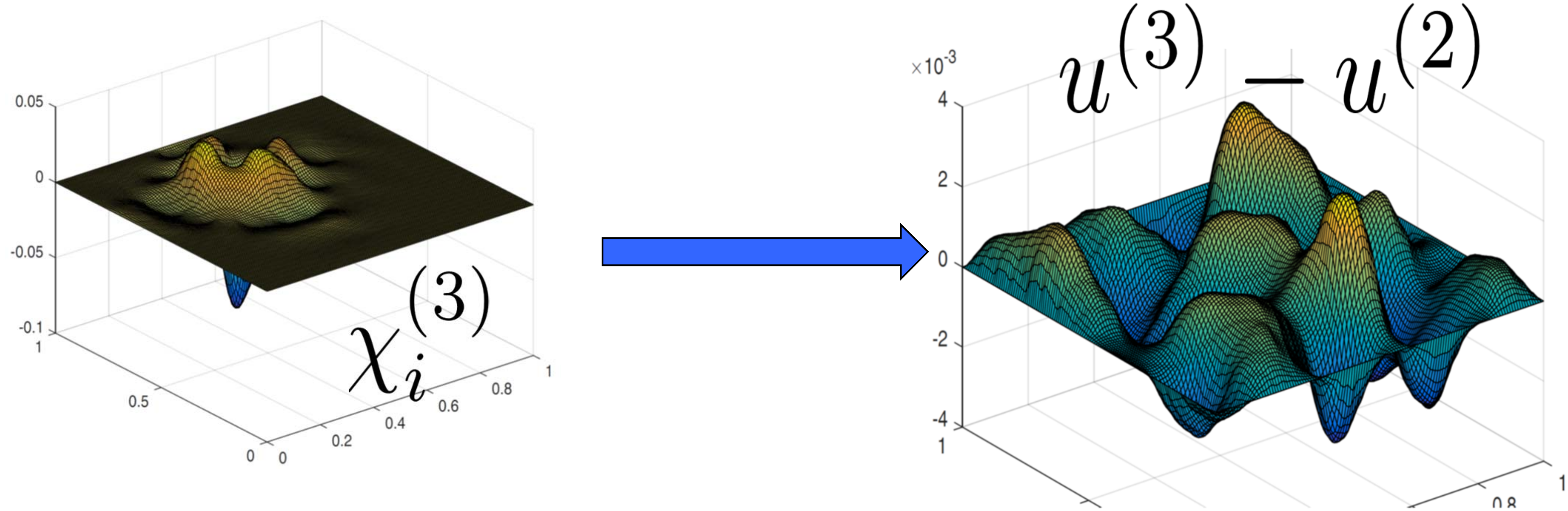
$$R^{(3,4)} = \pi^{(3,4)} \left( I^{(4)} - A^{(4)} W^{(4),T} B^{(4),-1} W^{(4)} \right)$$

$$\boxed{A^{(3)} = R^{(3,4)} A^{(4)} \left( R^{(3,4)} \right)^T}$$




$$\boxed{\chi_i^{(3)} := \sum_{j \in \mathcal{I}^{(3)}} W_{i,j}^{(3)} \psi_j^{(3)}}$$

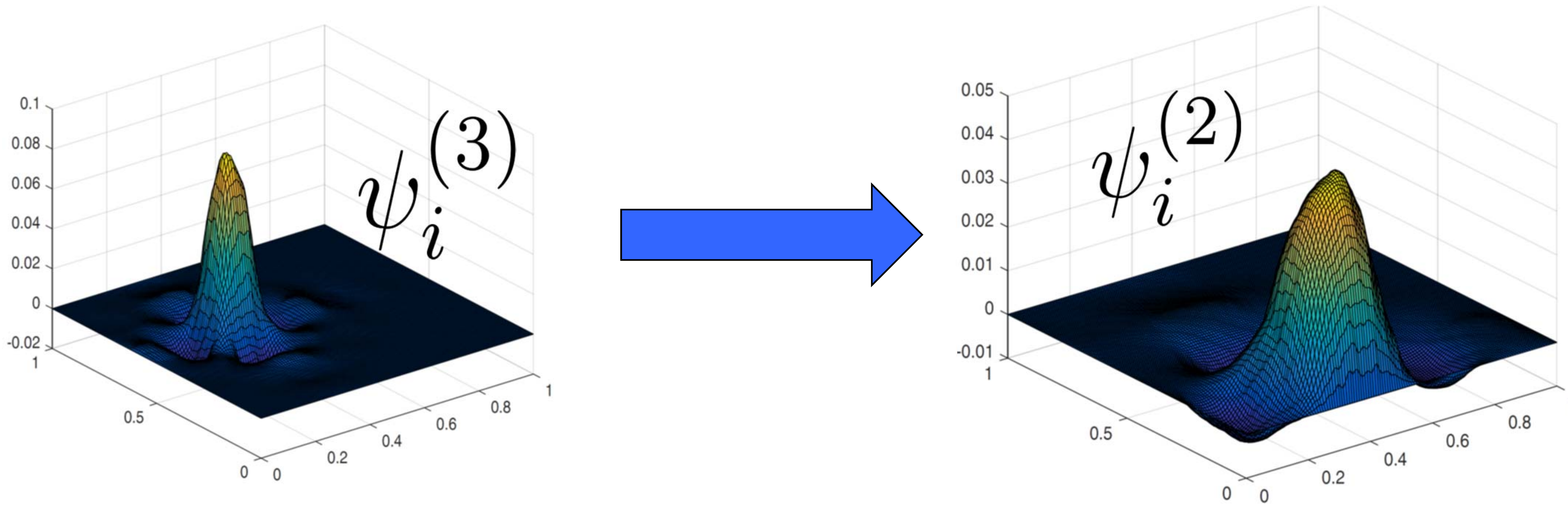





$$\boxed{u^{(3)} - u^{(2)} = \sum_i w_i^{(3)} \chi_i^{(3)}}$$

$$\boxed{B^{(3)} w^{(3)} = W^{(3),T} g^{(3)}} \qquad g^{(3)} = R^{(3,4)} g^{(4)}$$

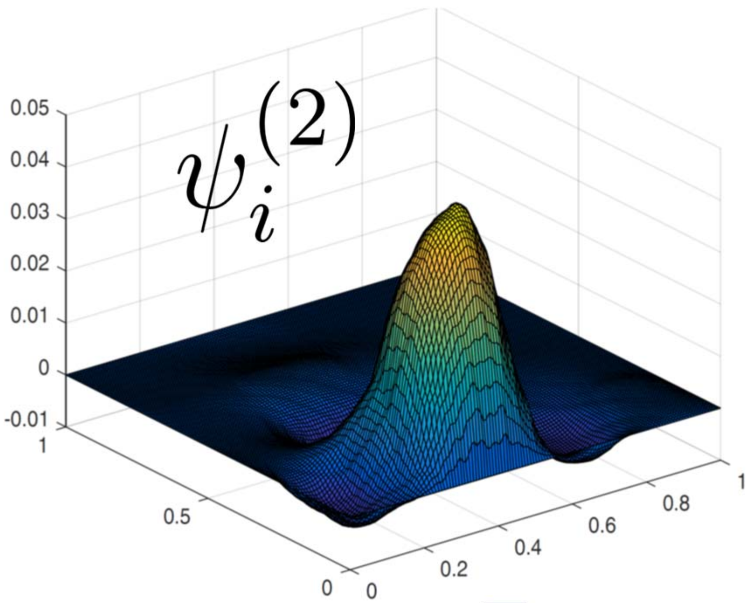
$$B^{(3)} = W^{(3)} A^{(3)} W^{(3),T}$$



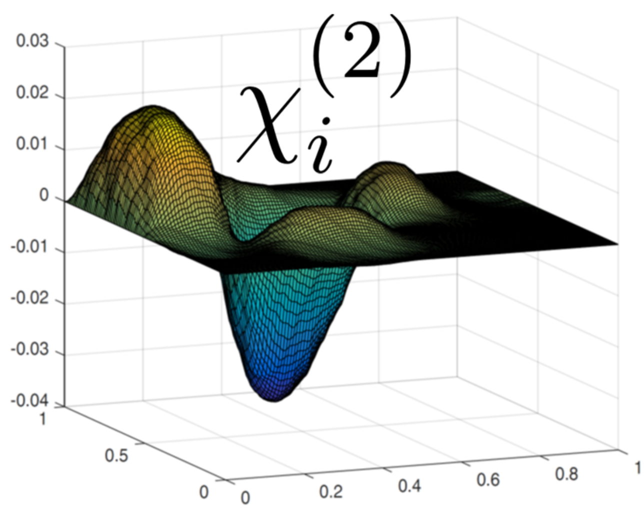

$$\boxed{\psi_i^{(2)} = R_{i,j}^{(2,3)} \psi_j^{(3)}}$$

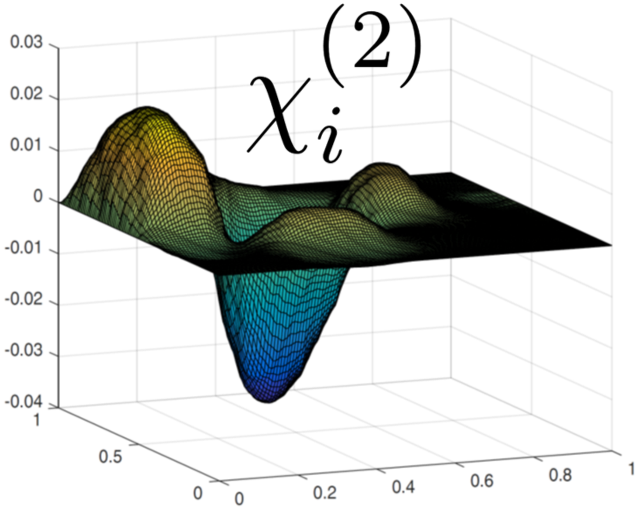
$$R^{(2,3)} = \pi^{(2,3)} \left( I^{(3)} - A^{(3)} W^{(3),T} B^{(3),-1} W^{(3)} \right)$$

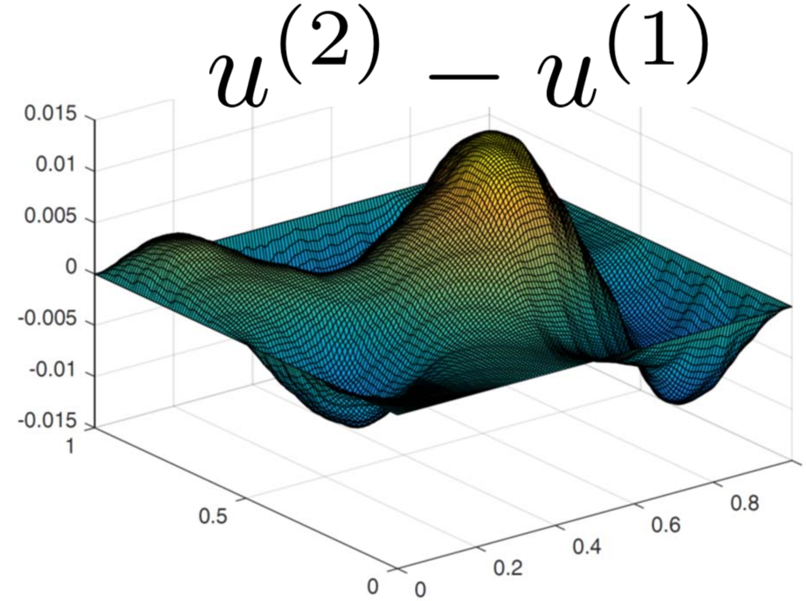
$$\boxed{A^{(2)} = R^{(2,3)} A^{(3)} \left( R^{(2,3)} \right)^T}$$




$$\boxed{\chi_i^{(2)} := \sum_{j \in \mathcal{I}^{(2)}} W_{i,j}^{(2)} \psi_j^{(2)}}$$



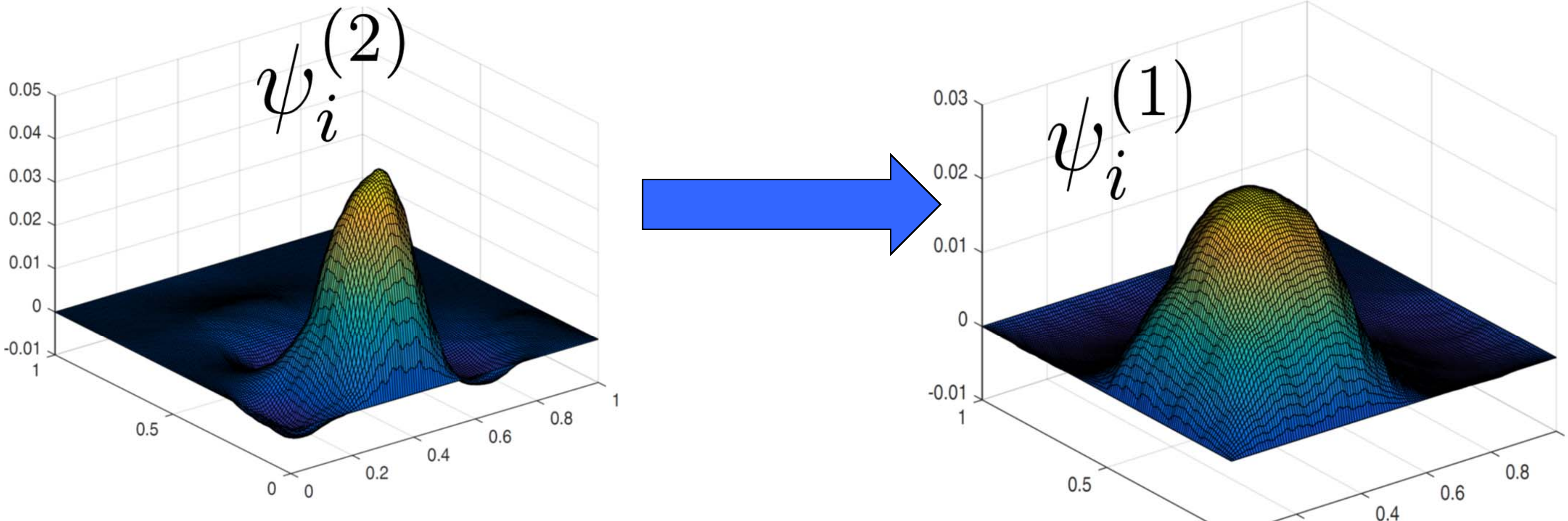



$$\boxed{u^{(2)} - u^{(1)} = \sum_i w_i^{(2)} \chi_i^{(2)}}$$

$$\boxed{B^{(2)} w^{(2)} = W^{(2),T} g^{(2)}} \qquad g^{(2)} = R^{(2,3)} g^{(3)}$$

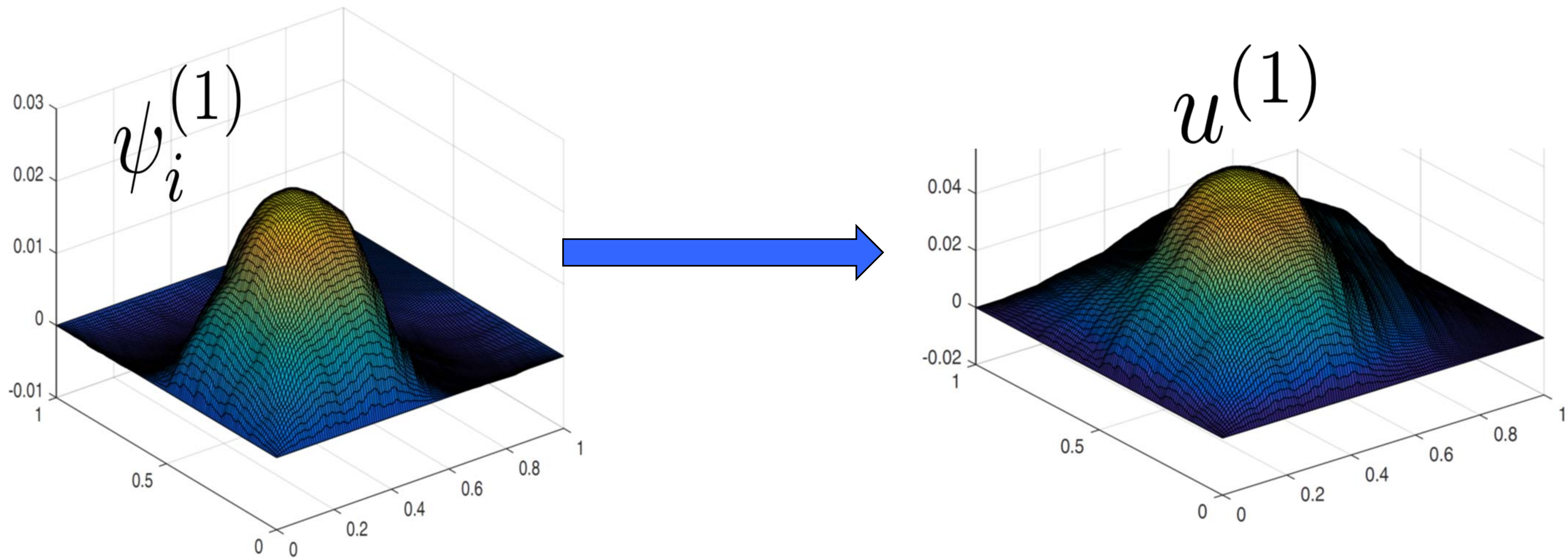
$$B^{(2)} = W^{(2)} A^{(2)} W^{(2),T}$$




$$\boxed{\psi_i^{(1)} = R_{i,j}^{(1,2)} \psi_j^{(2)}}$$

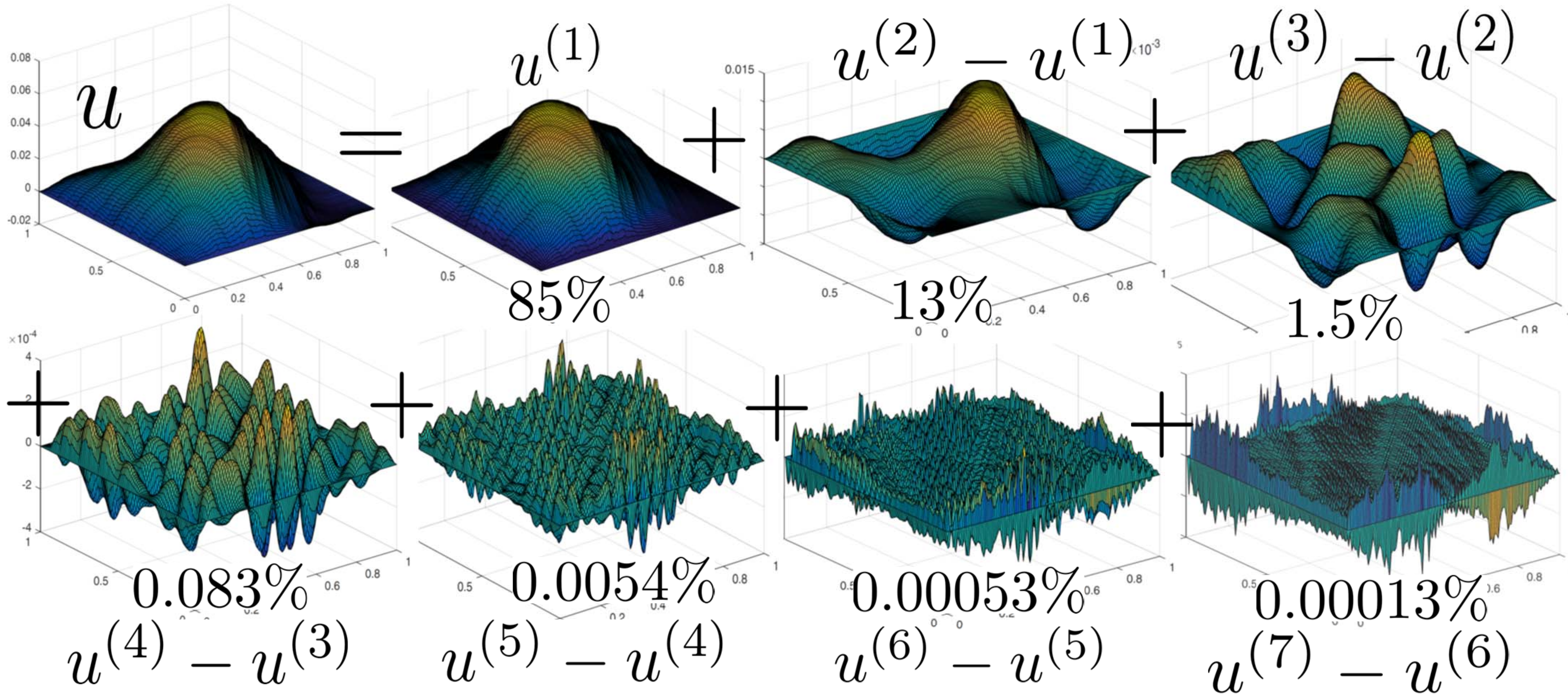
$$R^{(1,2)} = \pi^{(1,2)}(I^{(2)} - A^{(2)} W^{(2),T} B^{(2),-1} W^{(2)})$$

$$\boxed{A^{(1)} = R^{(1,2)} A^{(2)} (R^{(1,2)})^T}$$




$$\boxed{u^{(1)} = \sum_i v_i^{(1)} \psi_i^{(1)}}$$

$$\boxed{A^{(1)} v^{(1)} = g^{(1)}} \qquad g^{(1)} = R^{(1,2)} g^{(2)}$$




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

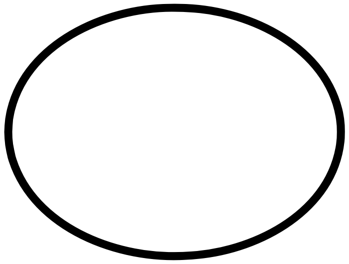
**Compression, inversion and and approximate PCA of dense kernel matrices at Near-linear computational complexity**

- Schäfer, Sullivan, Owhadi. 2017.
- arXiv:1706.02205

Florian Schäfer

Tim Sullivan

## Covariance function/kernel



$\Omega \subset \mathbb{R}^d$: Bounded domain

$\mathcal{L}$: arbitrary continuous positive symmetric linear bijection

$$\left(H_0^s(\Omega), \|\cdot\|_{H_0^s(\Omega)}\right) \xrightarrow{\quad\mathcal{L}\quad} \left(H^{-s}(\Omega), \|\cdot\|_{H^{-s}(\Omega)}\right)$$

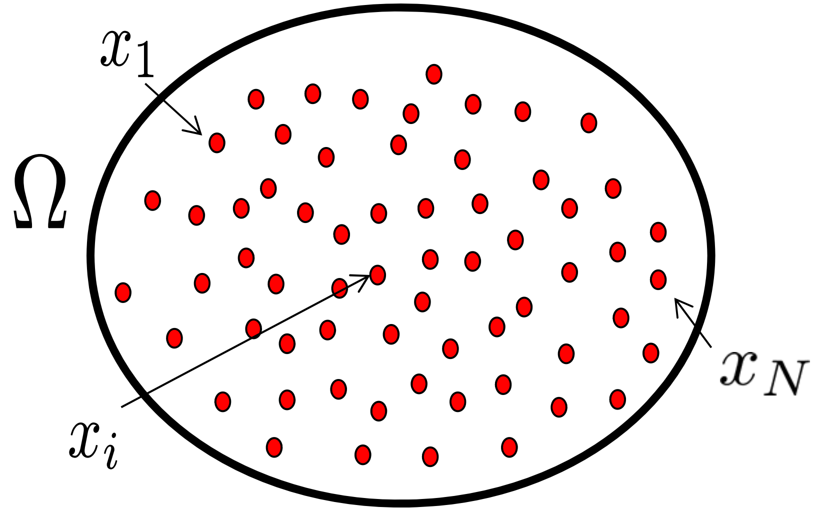
$\mathcal{L}$: is local $\quad \int_\Omega u\mathcal{L}v = 0$ if $u$ and $v$ have disjoint supports

## Covariance function = its Green functions

$$\boxed{G = \mathcal{L}^{-1}}$$

## The kernel/covariance matrix

$s > \frac{d}{2}$ ⟹ $G(x,y)$ is continuous




$x_1, \ldots, x_N$: Approximately homogeneous

$\Theta$: $N \times N$ symmetric positive definite matrix

$$\boxed{\Theta_{i,j} := G(x_i, x_j)}$$

## The kernel/covariance matrix

Important in

- Computational Physics
- (Gaussian process) statistics
- Kernel methods for machine learning (e.g. Support Vector Machines)

## Computational bottleneck

$\Theta$ is dense, naively we have

- Storage, $\mathcal{O}(N^2)$
- $\Theta v$, $\mathcal{O}(N^2)$
- $\Theta^{-1} v$, $\mathcal{O}(N^3)$
- $\det(\Theta)$, $\mathcal{O}(N^3)$
- $\mathrm{PCA}(\Theta)$, $\mathcal{O}(N^4)$

## Our algorithm

For $\epsilon > 0$ knowing only $\Omega$ and $\{x_i\}_{1\leq i\leq N}$, we will

- Select $\mathcal{O}(N \operatorname{polylog}(N) \operatorname{polylog}(\frac{1}{\epsilon}))$ entries of $\Theta$ and an ordering $P$ of $\{x_i\}_{1\leq i\leq N}$.
- From these entries compute a lower triangular matrix such that $nnz(n) = \mathcal{O}(N \operatorname{polylog}(N) \operatorname{polylog}(\frac{1}{\epsilon}))$.

## Theorem

The above can be done in complexity $N \operatorname{polylog}(N) \operatorname{polylog}(\frac{1}{\epsilon})$, in time and space, such that

$$\|\Theta - PLL^T P^T\| \leq \mathcal{O}(\epsilon)$$

Allows to approximate $\Theta v$, $\Theta^{-1} v$, $\det(\Theta)$, in $\mathcal{O}(N \operatorname{polylogN} \operatorname{polylog}\frac{1}{\epsilon})$ complexity

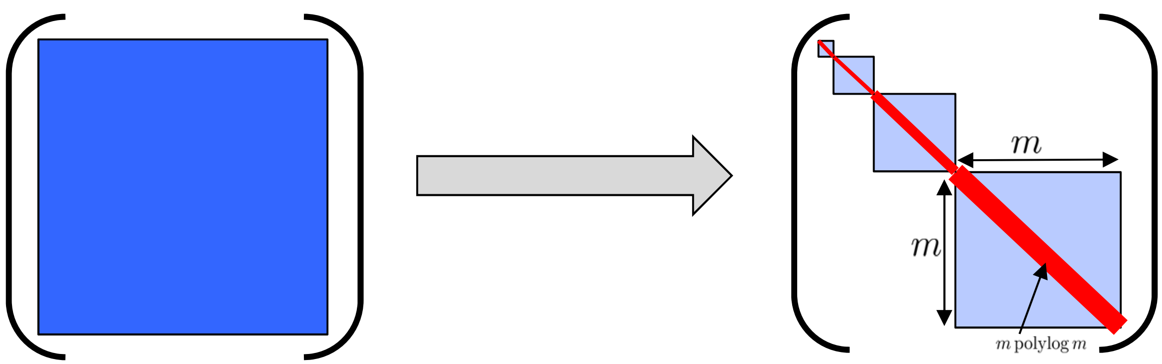
# Our algorithm

Furthermore,

- Empirically results also apply to Maérn kernel with possibly fractional smoothness.
- We obtain a near-linear complexity solver for elliptic PDE
- We obtain a sparse approximate PCA with near optimal low rank approximation.

$\boxed{\text{In complexity } N \text{ polylog} N \text{ polylog} \frac{1}{\epsilon}}$

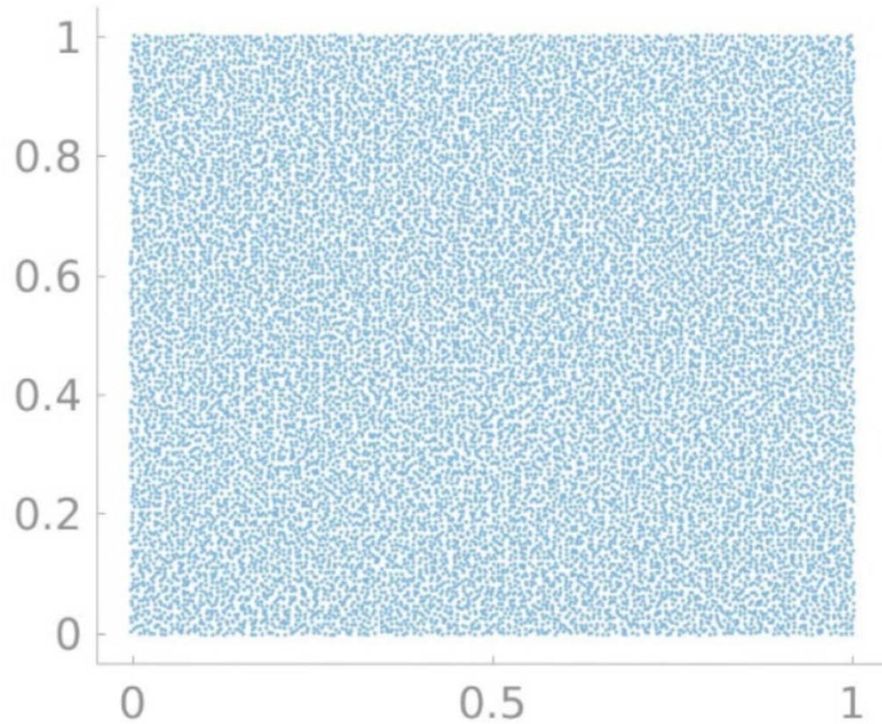
- Block diagonalize $\Theta$ into sparse well conditionned blocks



## A simple algorithm

- Decompose $\{x_i\}_{i \in \mathcal{I}}$ into a nested hierarchy:

$$\{x_i\}_{i \in \mathcal{I}^{(1)}} \subset \{x_i\}_{i \in \mathcal{I}^{(2)}} \subset \{x_i\}_{i \in \mathcal{I}^{(3)}} \subset \cdots \subset \{x_i\}_{i \in \mathcal{I}^{(q)}}$$



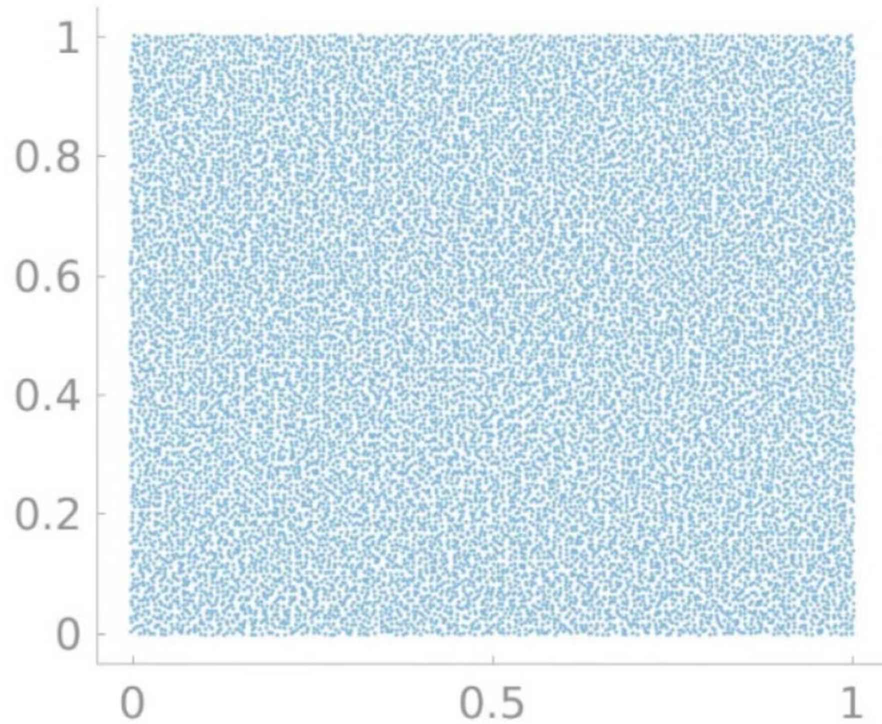
# A simple algorithm

- Decompose $\{x_i\}_{i \in \mathcal{I}}$ into a nested hierarchy:

$$\{x_i\}_{i \in \mathcal{I}^{(1)}} \subset \{x_i\}_{i \in \mathcal{I}^{(2)}} \subset \{x_i\}_{i \in \mathcal{I}^{(3)}} \subset \cdots \subset \{x_i\}_{i \in \mathcal{I}^{(q)}}$$

- Define

$$\mathcal{J}^{(k)} := \mathcal{I}^{(k)} / I^{(k-1)}$$



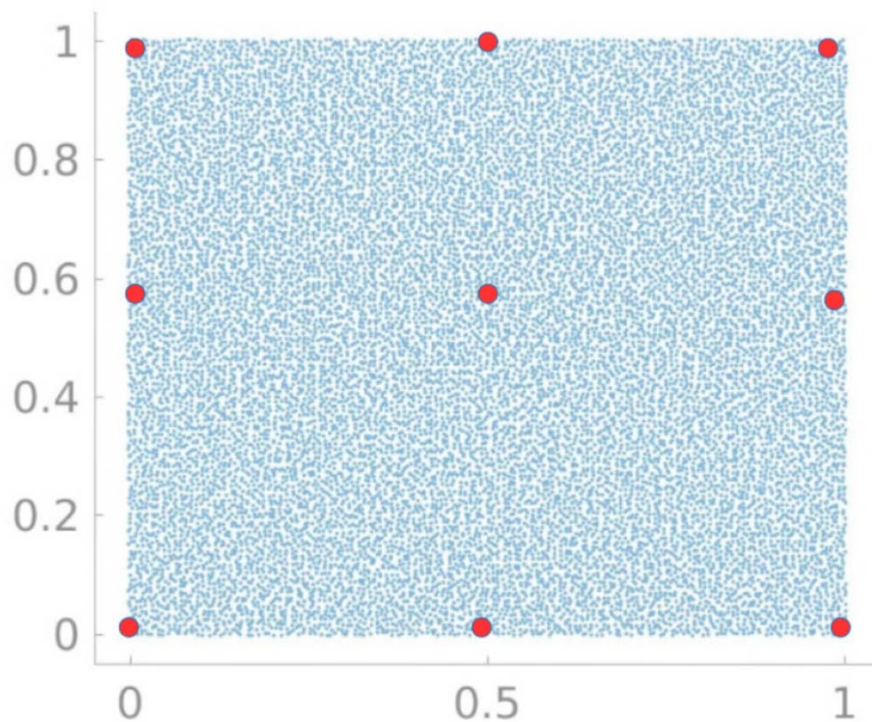
## A simple algorithm

- Decompose $\{x_i\}_{i\in\mathcal{I}}$ into a nested hierarchy:

$$\{x_i\}_{i\in\mathcal{I}^{(1)}} \subset \{x_i\}_{i\in\mathcal{I}^{(2)}} \subset \{x_i\}_{i\in\mathcal{I}^{(3)}} \subset \cdots \subset \{x_i\}_{i\in\mathcal{I}^{(q)}}$$

- Define

$$\mathcal{I}^{(1)} = \mathcal{J}^{(1)}$$



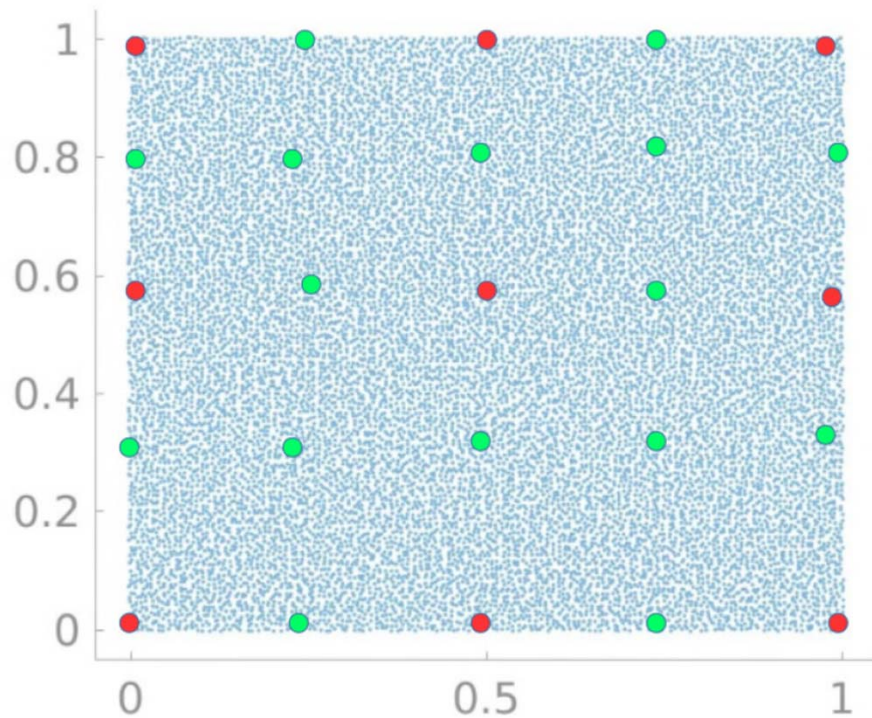
# A simple algorithm

- Decompose $\{x_i\}_{i\in\mathcal{I}}$ into a nested hierarchy:

$\{x_i\}_{i\in\mathcal{I}^{(1)}} \subset \{x_i\}_{i\in\mathcal{I}^{(2)}} \subset \{x_i\}_{i\in\mathcal{I}^{(3)}} \subset \cdots \subset \{x_i\}_{i\in\mathcal{I}^{(q)}}$

- Define

$$\mathcal{I}^{(2)} = \mathcal{J}^{(1)} \cup \mathcal{J}^{(2)}$$



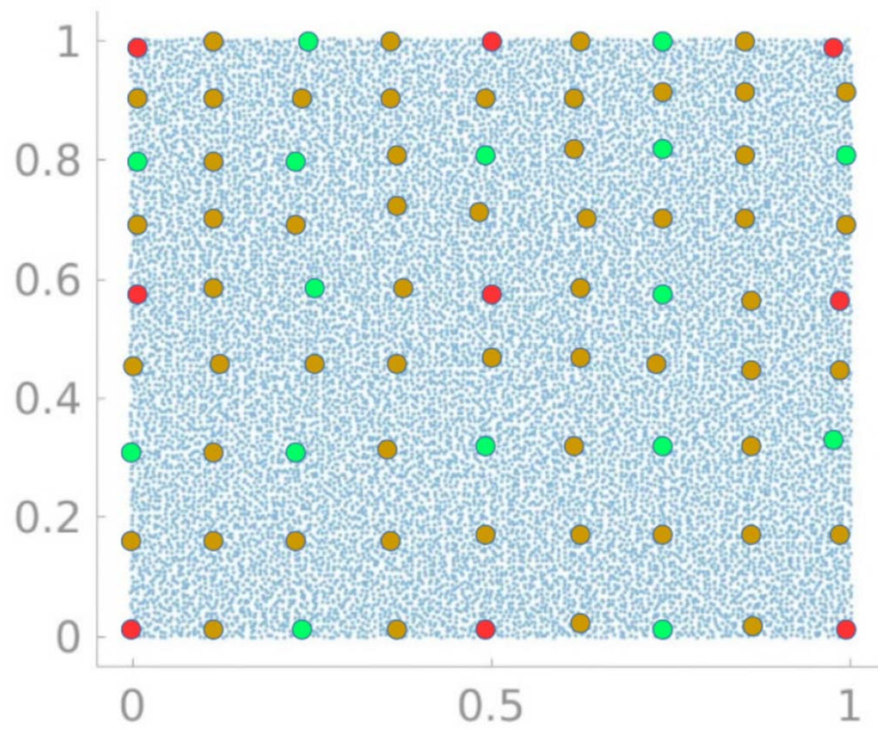
## A simple algorithm

- Decompose $\{x_i\}_{i\in\mathcal{I}}$ into a nested hierarchy:

$$\{x_i\}_{i\in\mathcal{I}^{(1)}} \subset \{x_i\}_{i\in\mathcal{I}^{(2)}} \subset \{x_i\}_{i\in\mathcal{I}^{(3)}} \subset \cdots \subset \{x_i\}_{i\in\mathcal{I}^{(q)}}$$

- Define

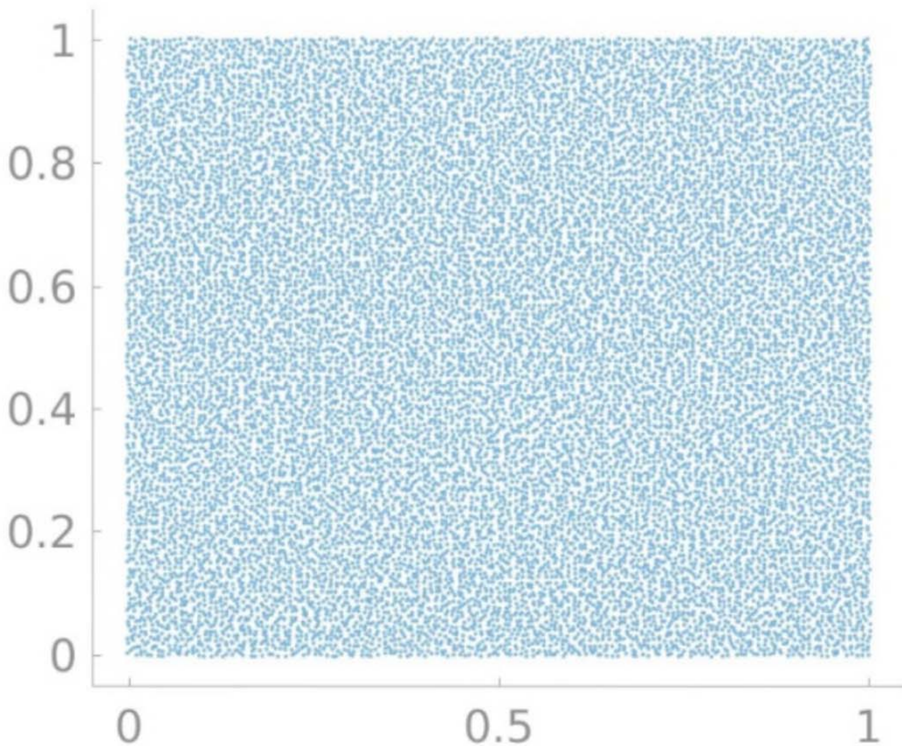
$$\mathcal{I}^{(3)} = \mathcal{J}^{(1)} \cup \mathcal{J}^{(2)} \cup \mathcal{J}^{(3)}$$

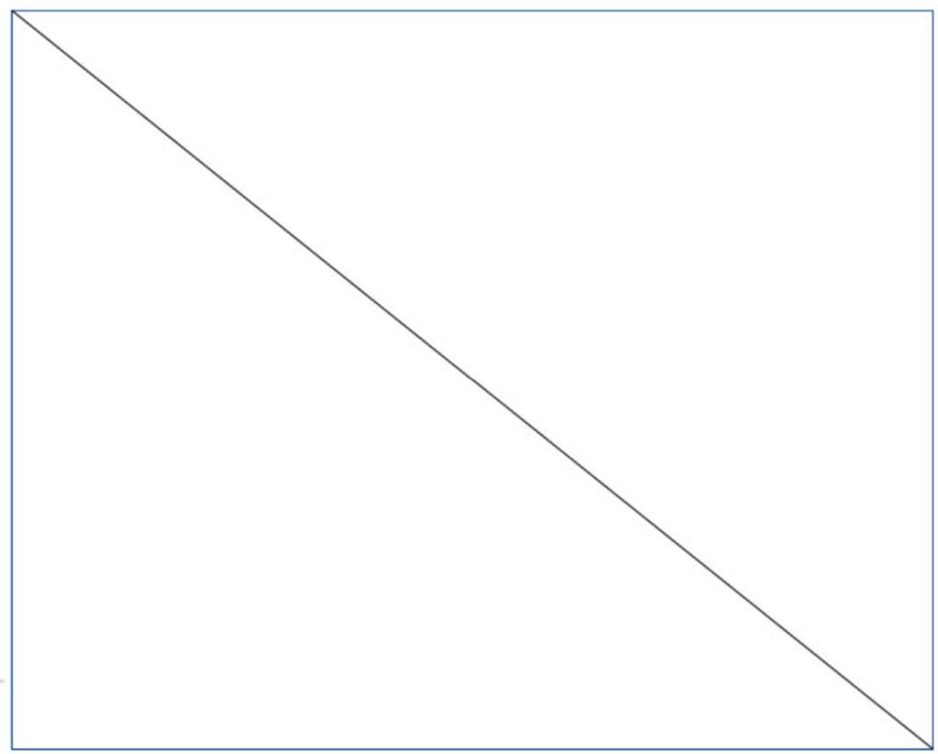


# A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\, j \in \mathcal{J}^{(l)},\, \mathrm{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

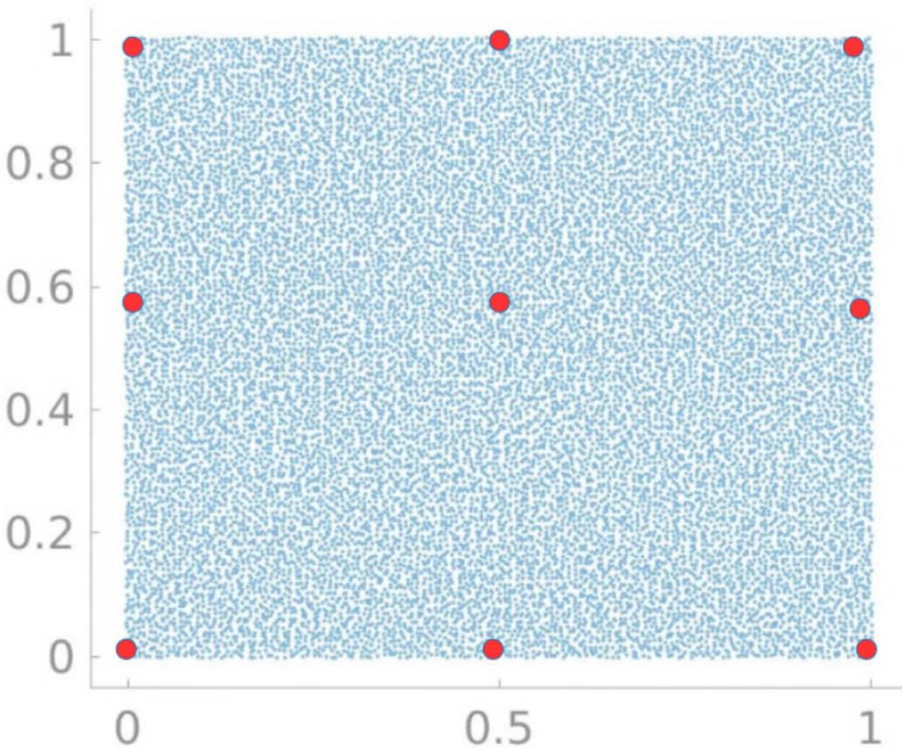


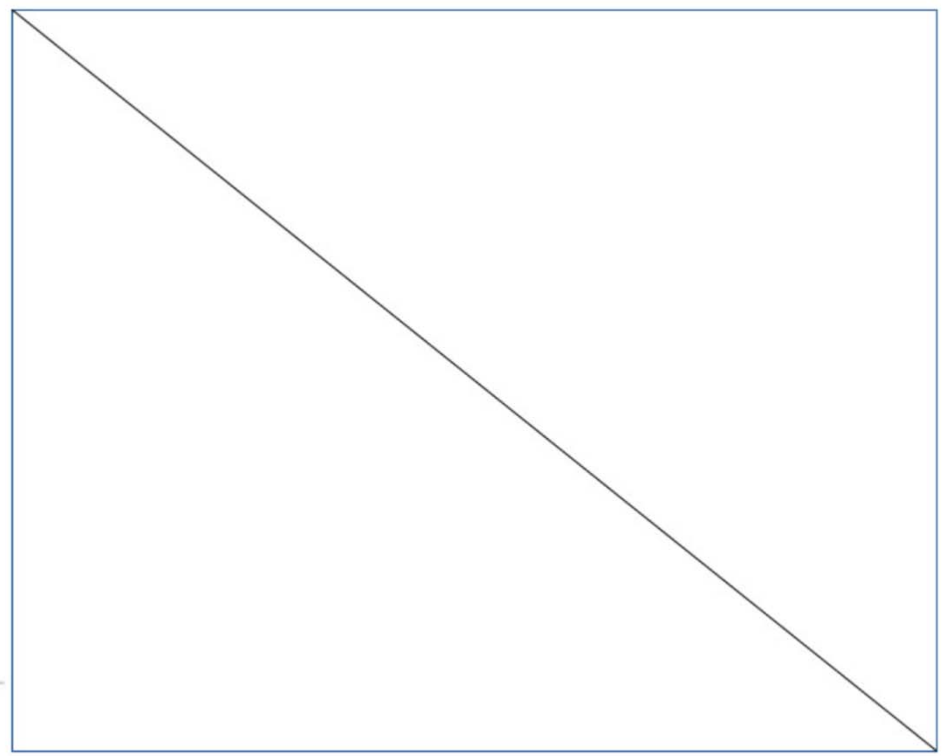



## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\ j \in \mathcal{J}^{(l)},\ \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

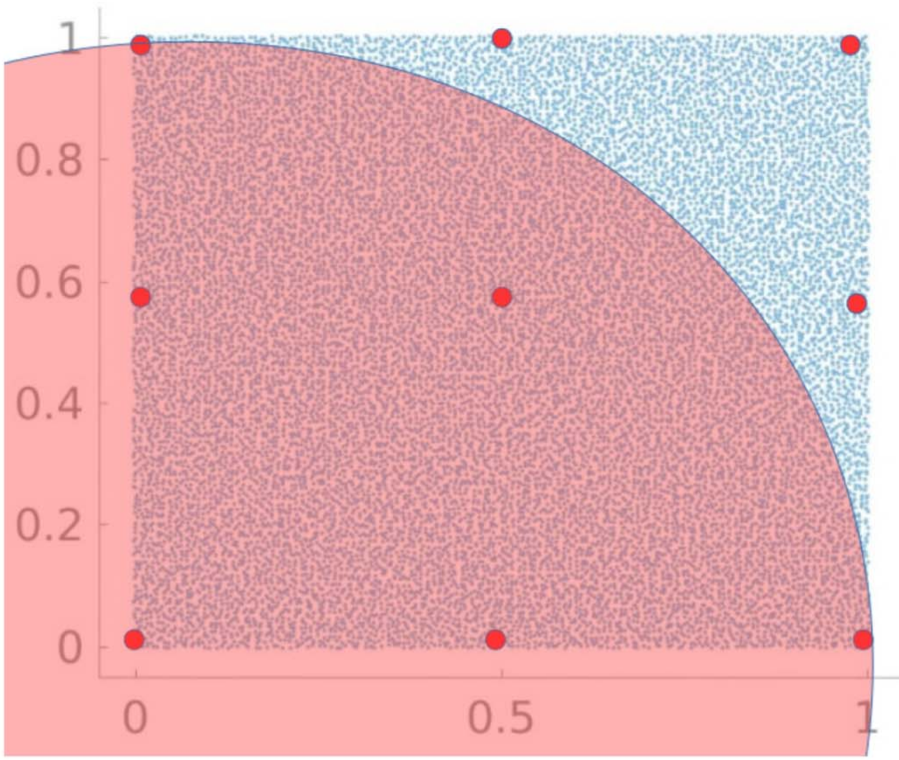


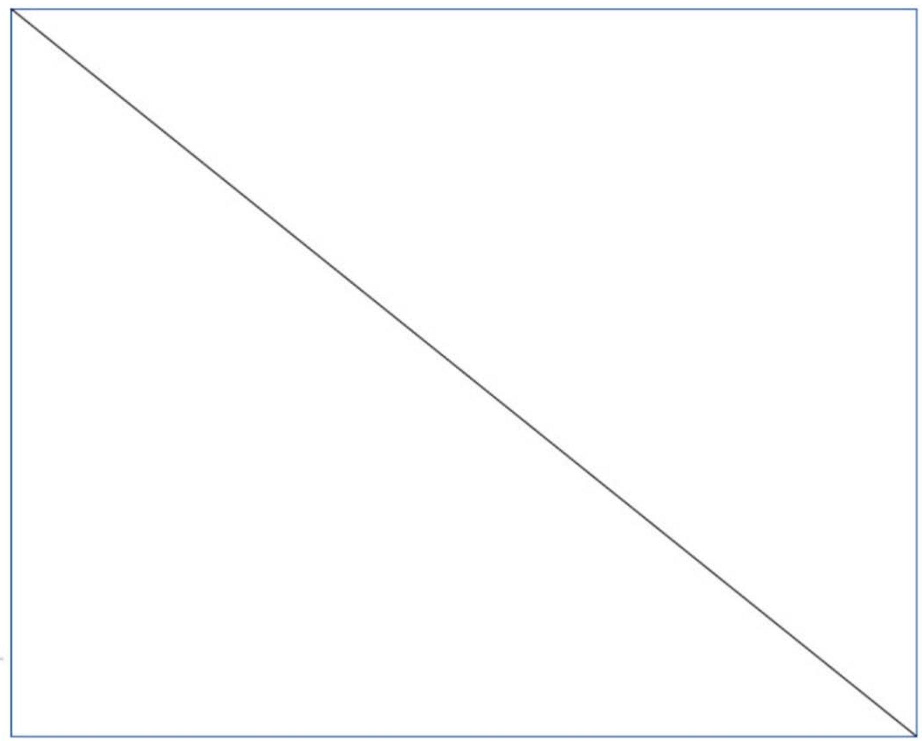



## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\ j \in \mathcal{J}^{(l)},\ \mathrm{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

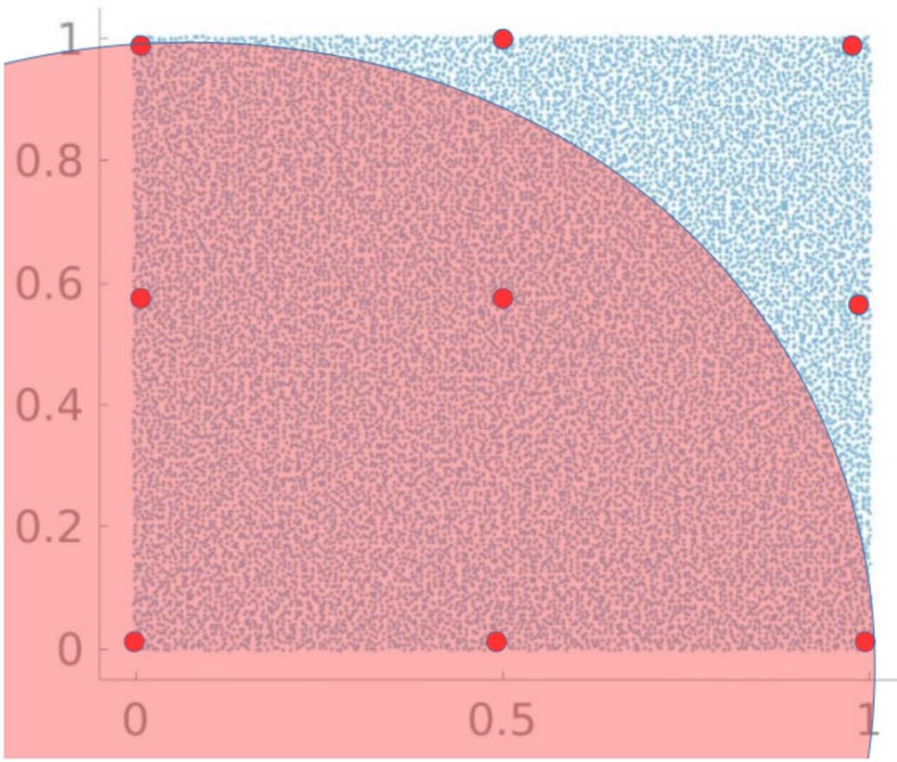


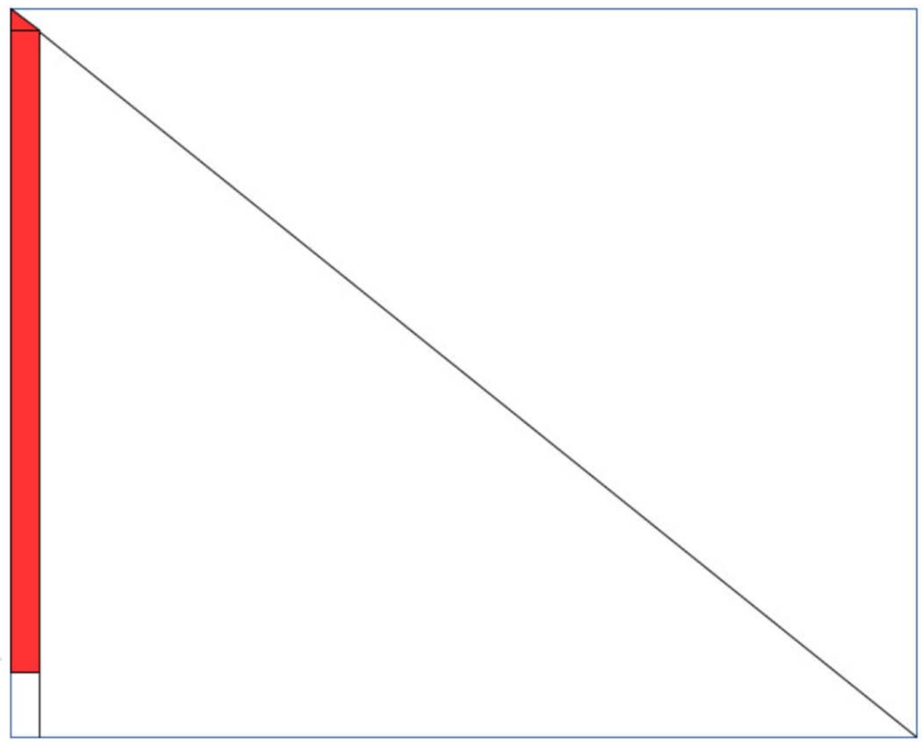



## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\, j \in \mathcal{J}^{(l)},\, \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

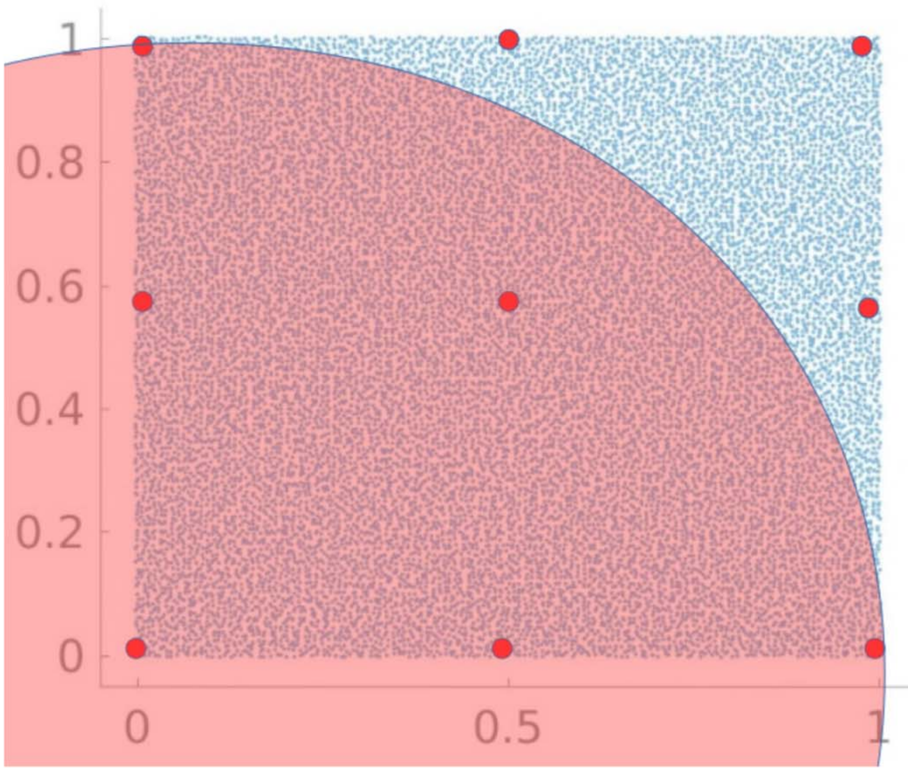


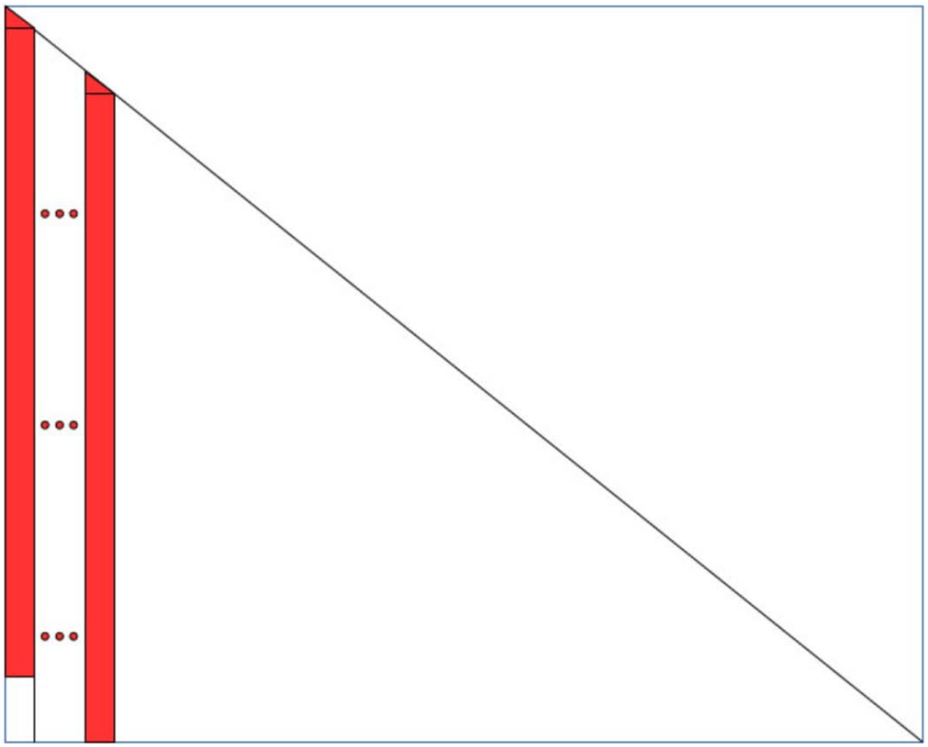



# A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\, j \in \mathcal{J}^{(l)},\, \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

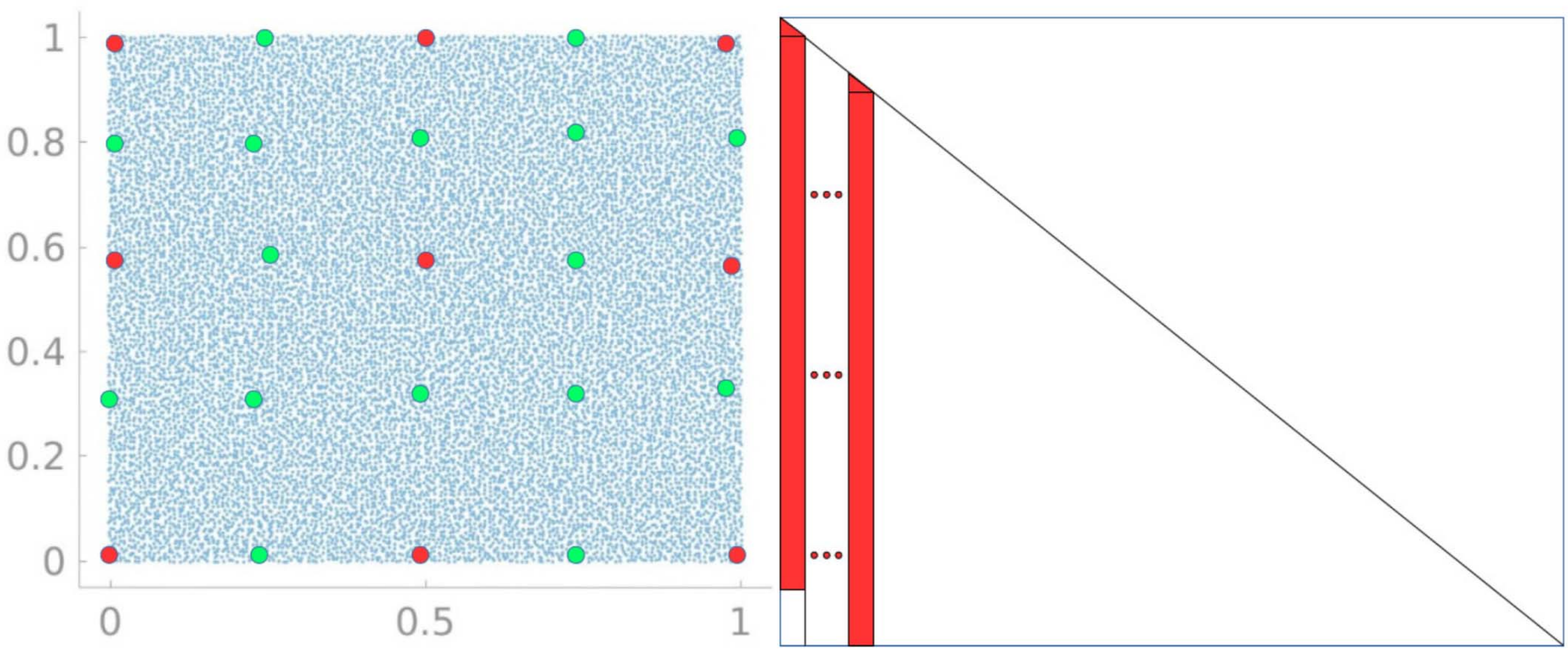


## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

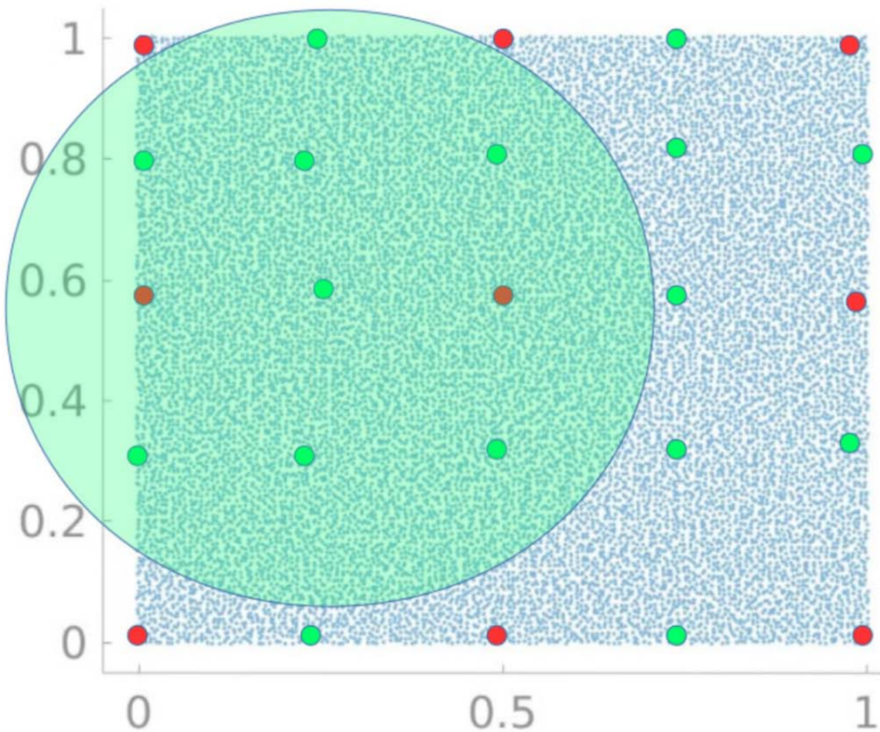
$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\ j \in \mathcal{J}^{(l)},\ \operatorname{dist}(x_i, x_j) \le 2 \times 2^{-\min(k,l)}\right\}$$

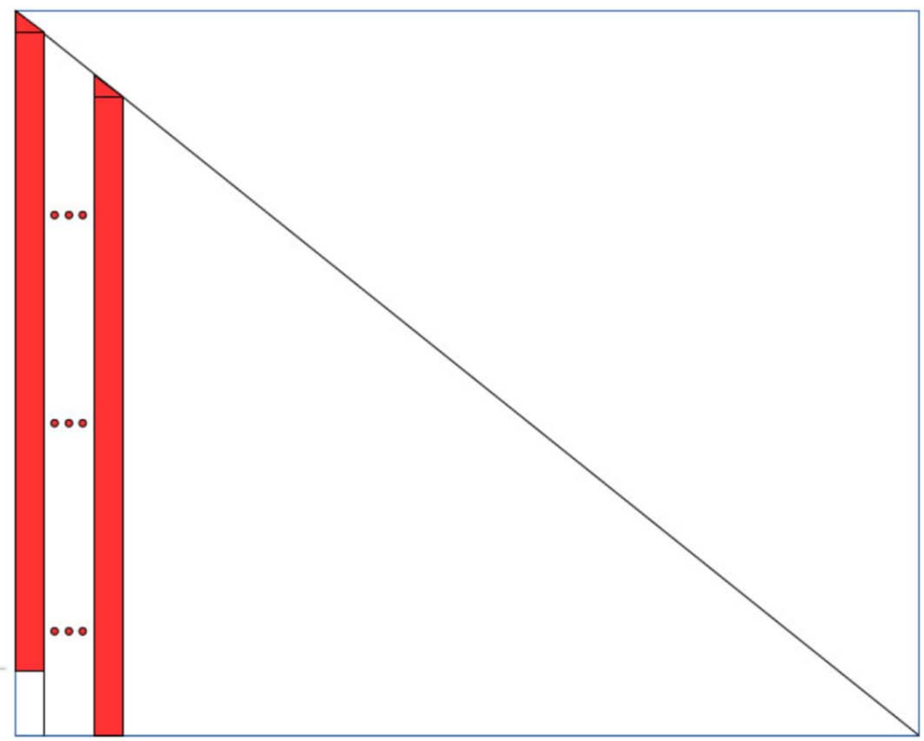


## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\, j \in \mathcal{J}^{(l)},\, \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

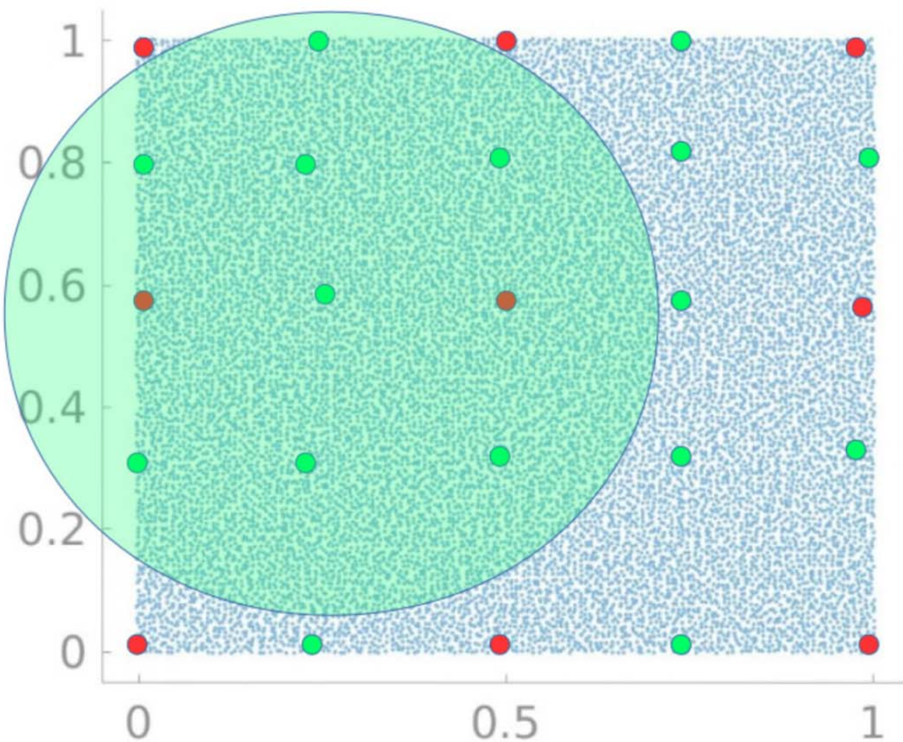


## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

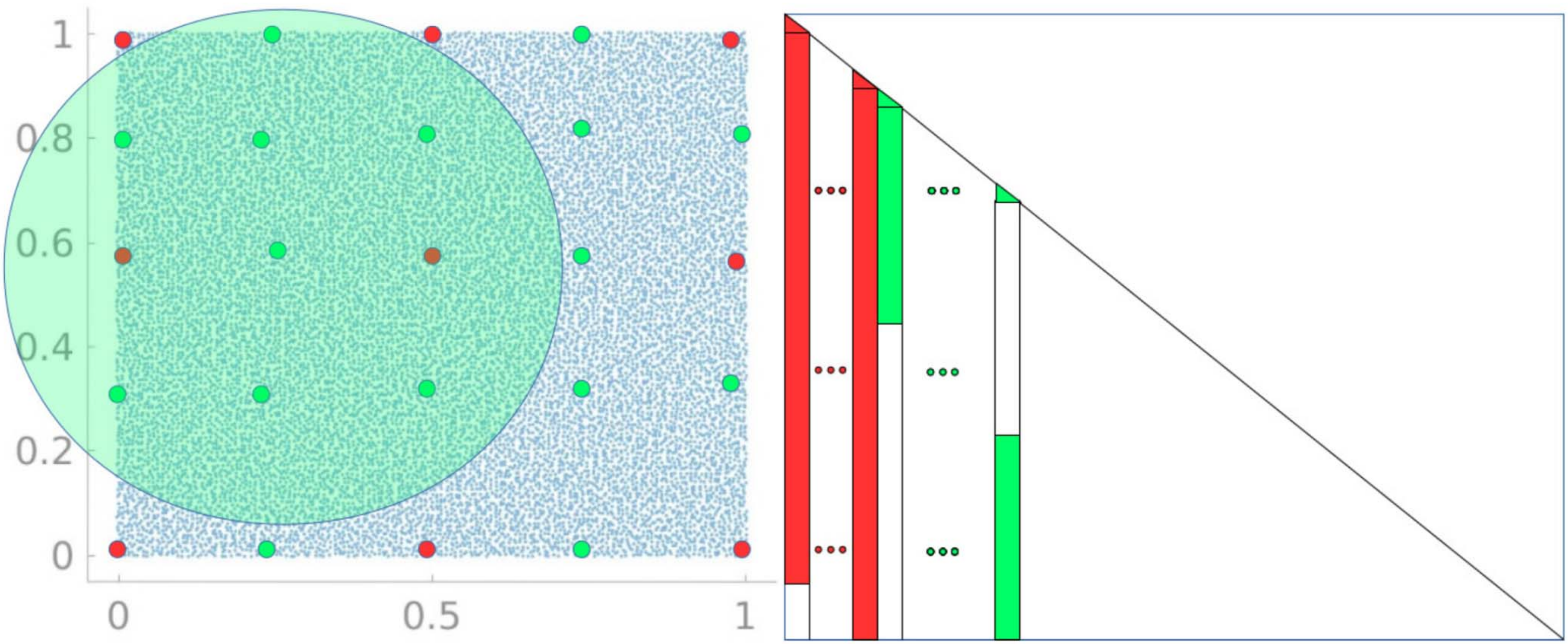
$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\ j \in \mathcal{J}^{(l)},\ \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$




# A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

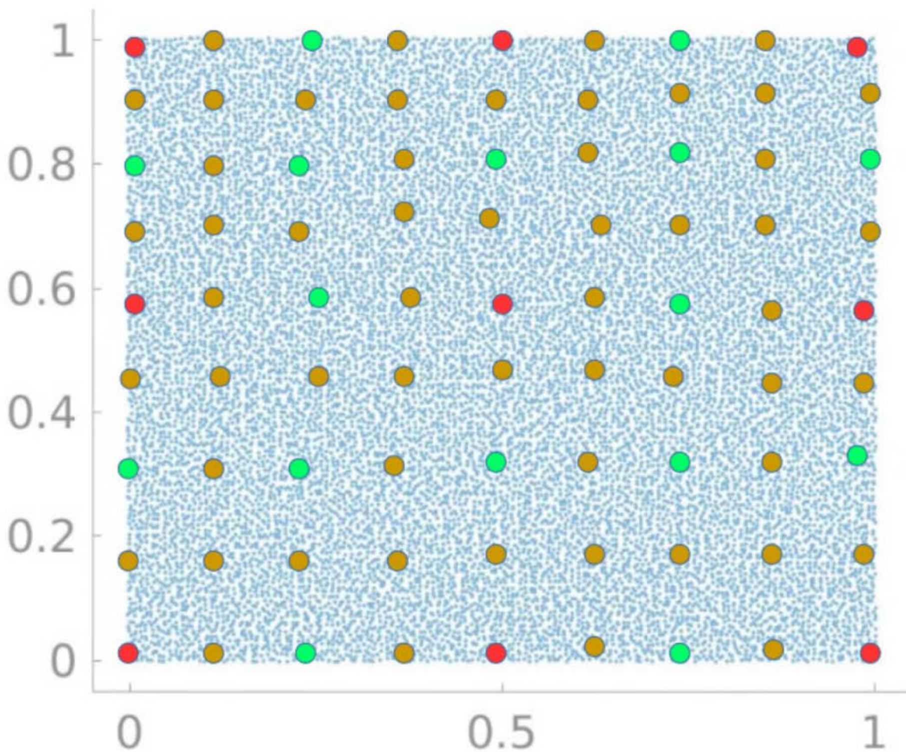
$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\ j \in \mathcal{J}^{(l)},\ \mathrm{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

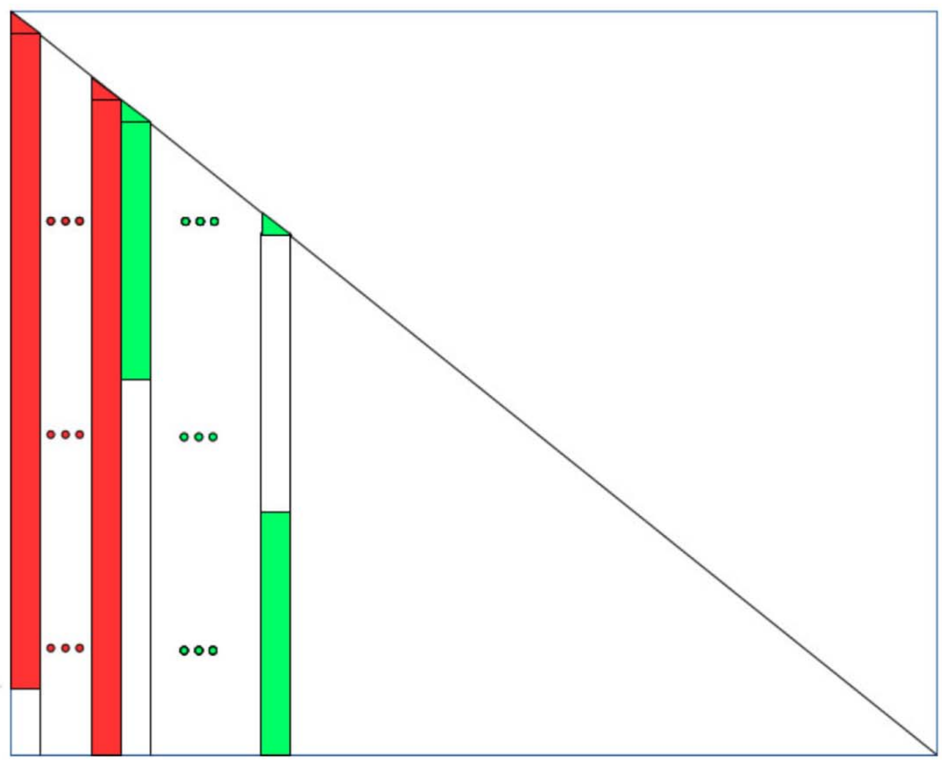


## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\, j \in \mathcal{J}^{(l)},\, \operatorname{dist}(x_i, x_j) \le 2 \times 2^{-\min(k,l)}\right\}$$

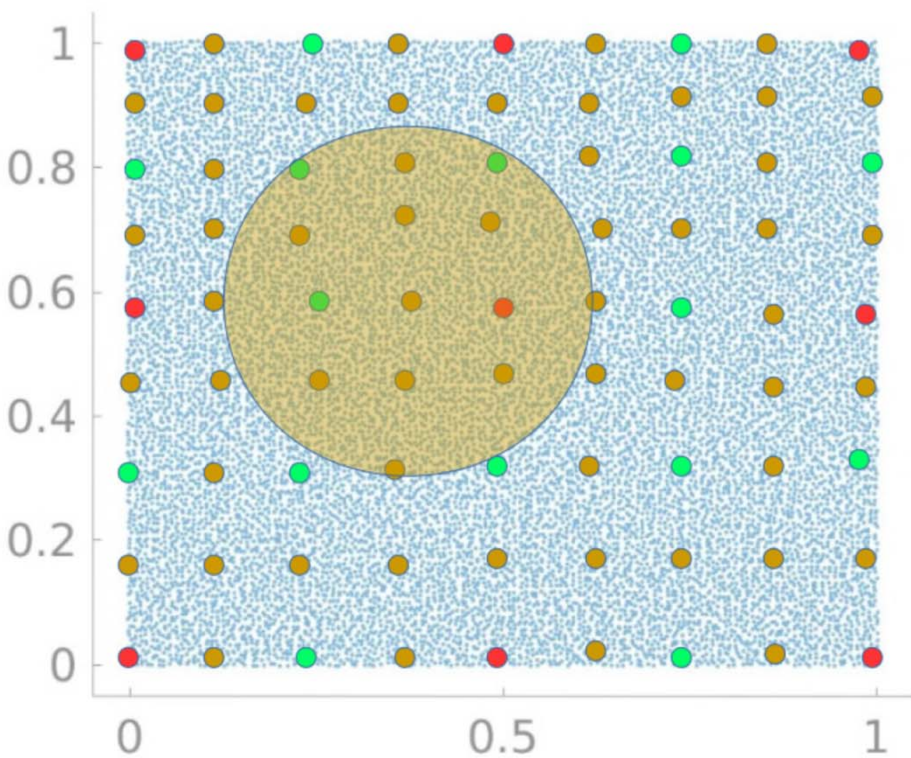


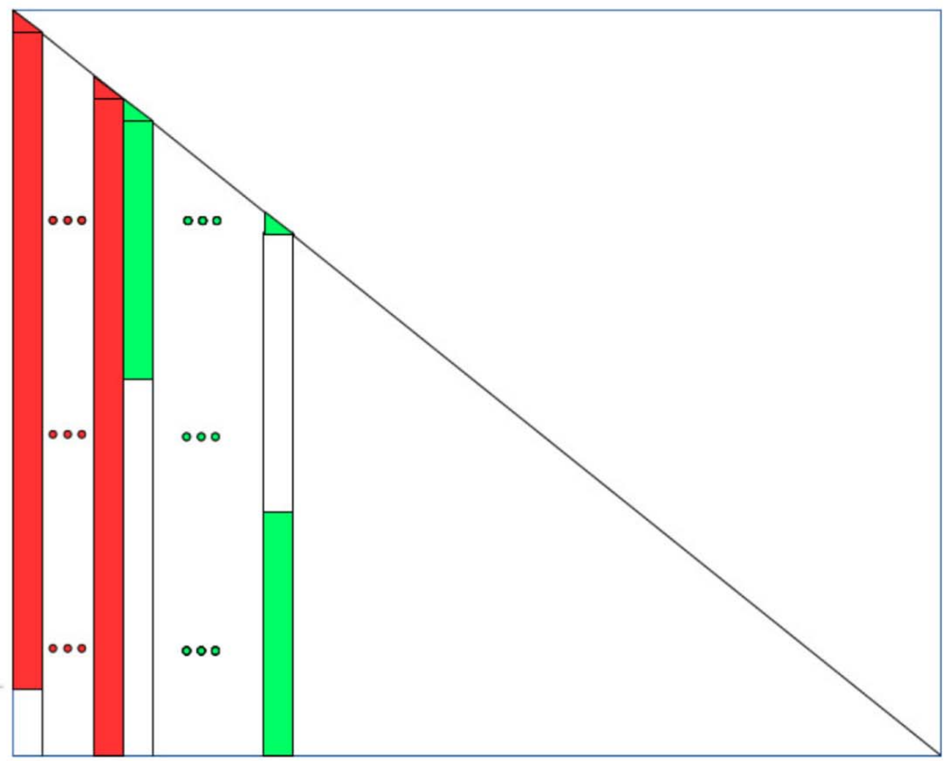



# A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\ j \in \mathcal{J}^{(l)},\ \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

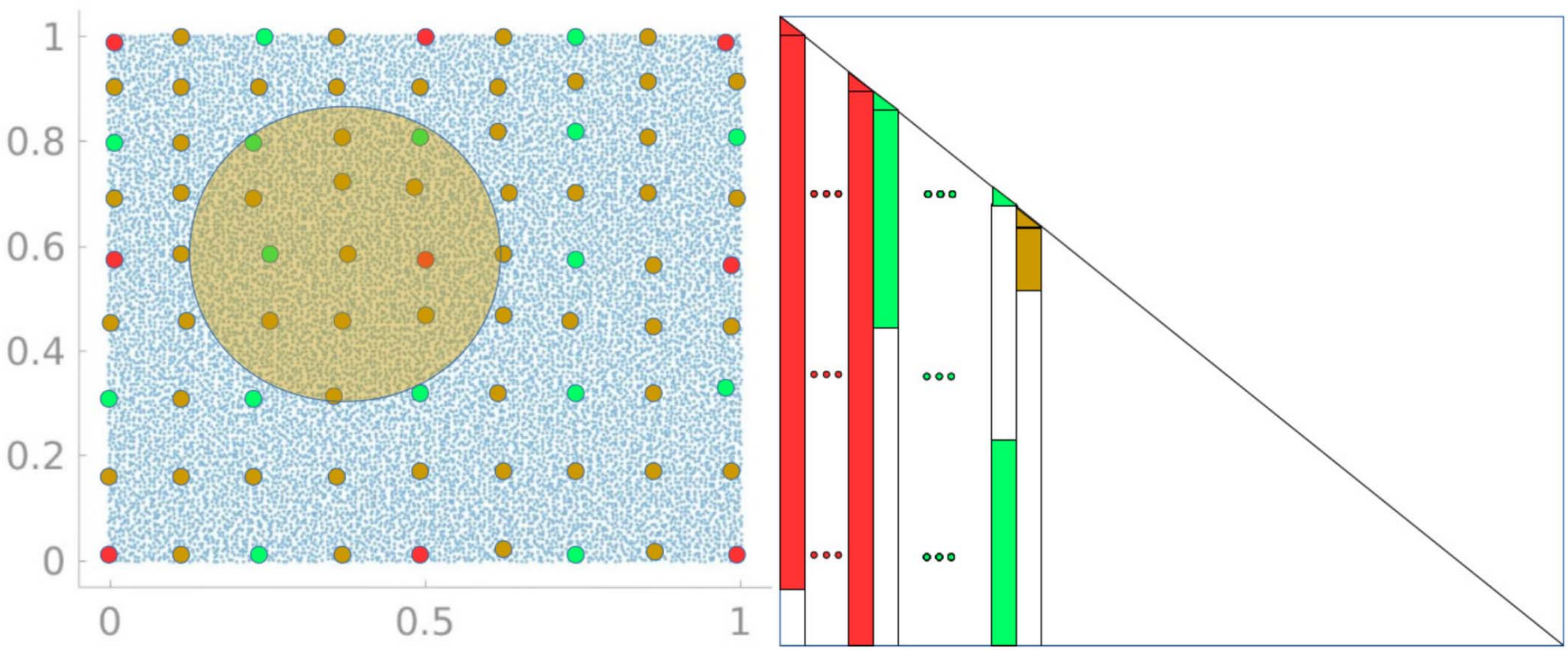


## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

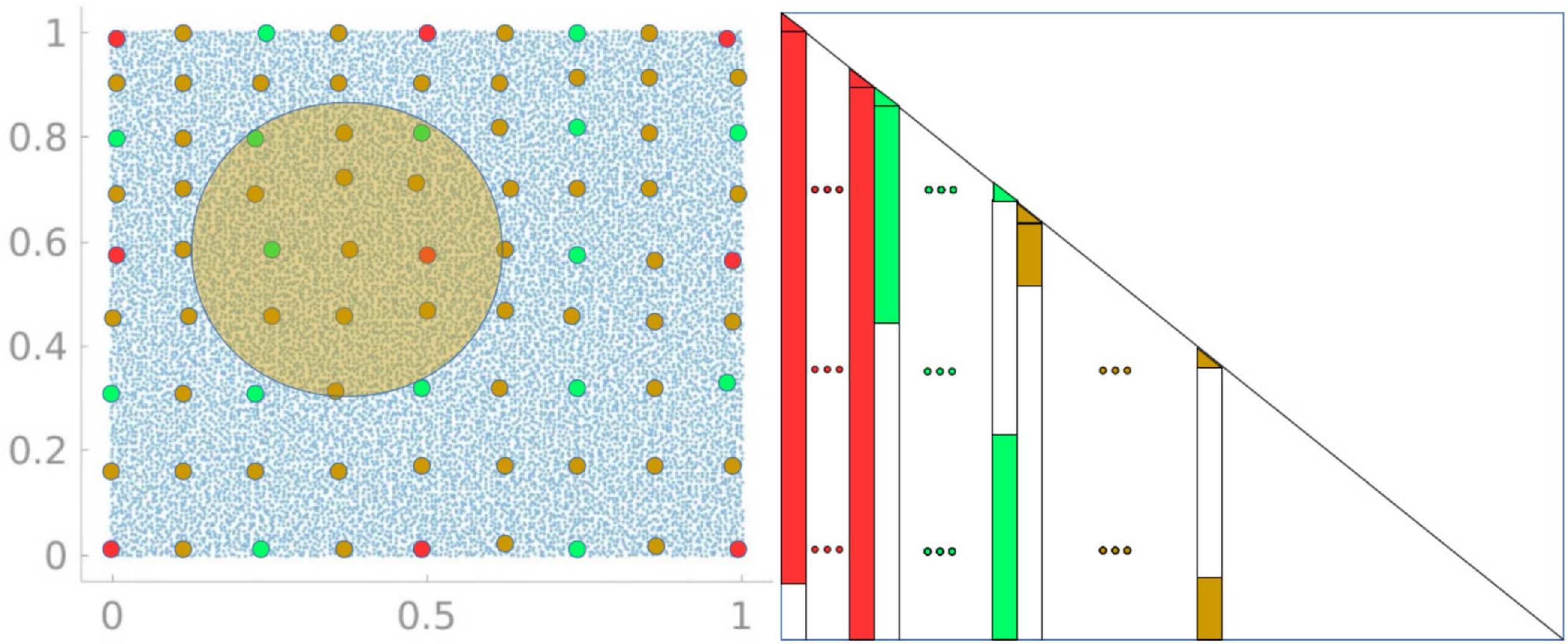
$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\, j \in \mathcal{J}^{(l)},\, \mathrm{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$



# A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

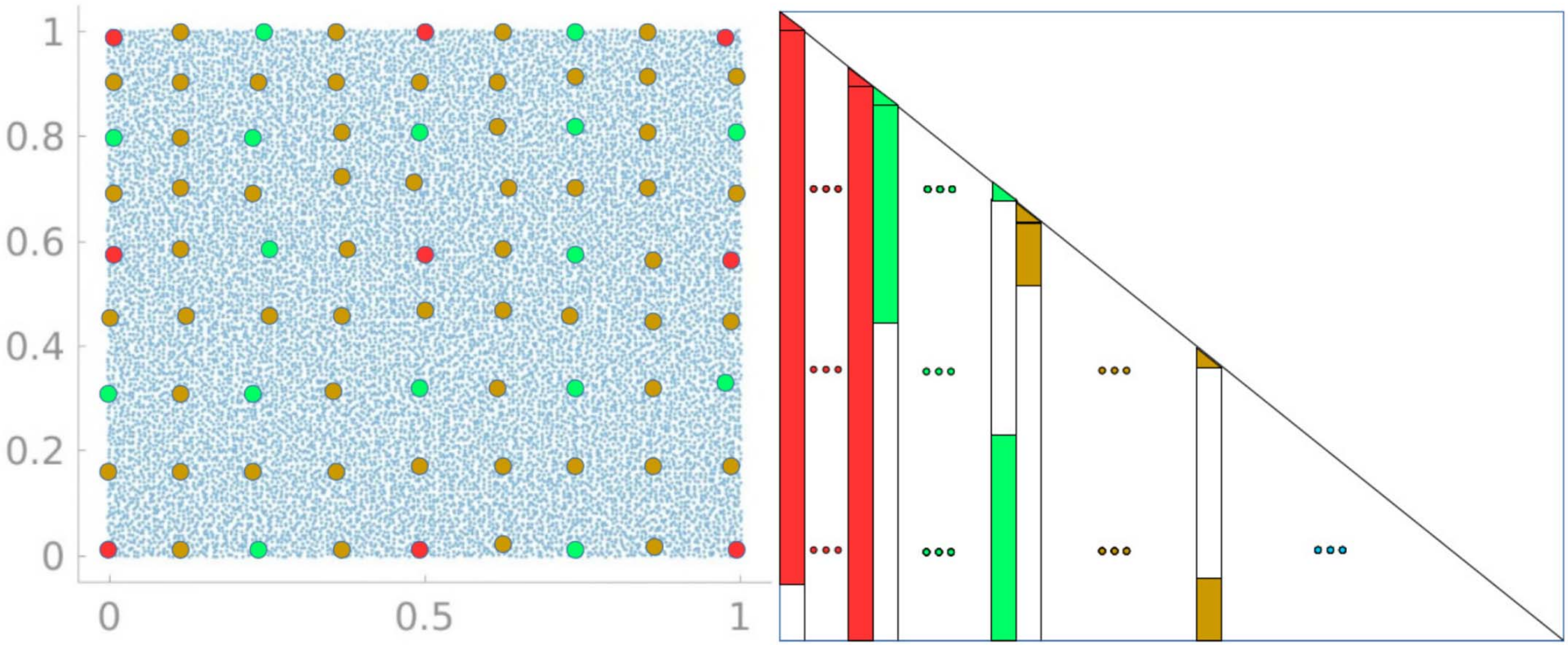
$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\, j \in \mathcal{J}^{(l)},\, \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$



## A simple algorithm

- We order the degrees of freedom from $\mathcal{J}^{(1)}$ to $\mathcal{J}^{(q)}$ and define the sparsity pattern:

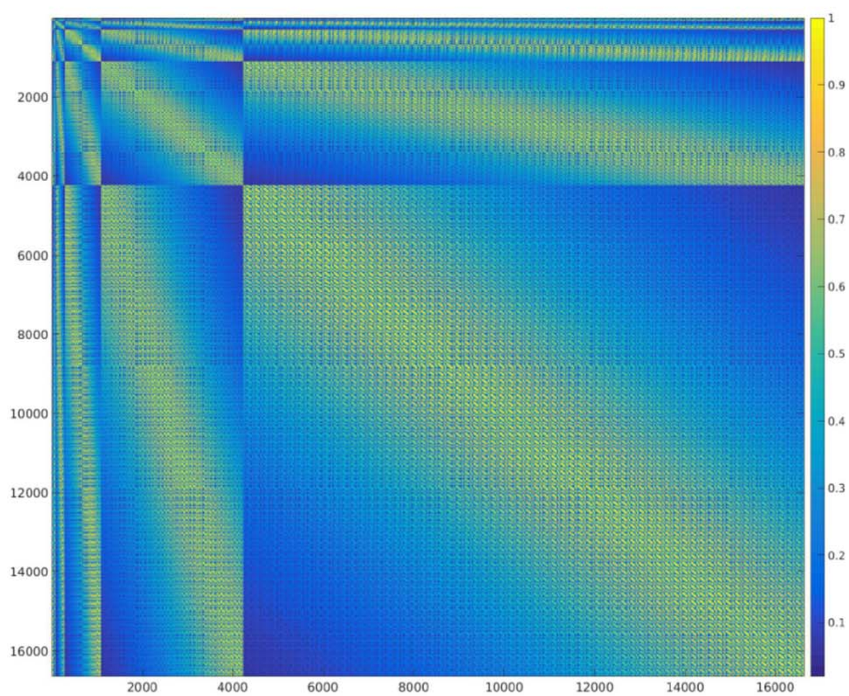
$$S := \left\{(i,j) \in \mathcal{I} \times \mathcal{I} \middle| i \in \mathcal{J}^{(k)},\ j \in \mathcal{J}^{(l)},\ \operatorname{dist}(x_i, x_j) \leq 2 \times 2^{-\min(k,l)}\right\}$$

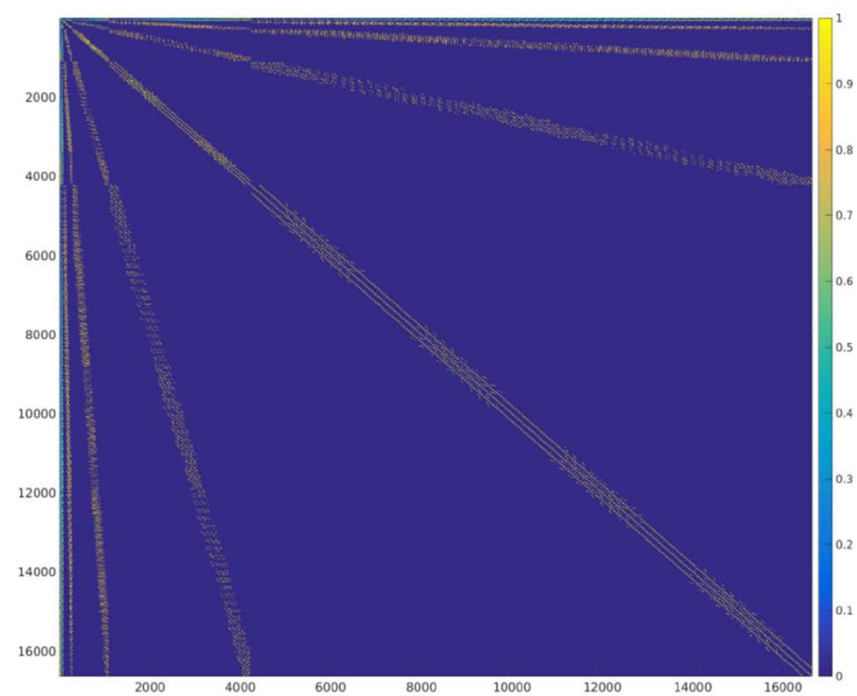


# A simple algorithm

- Write the entries in matrix:

$$\Gamma_{i,j} := \begin{cases} \Theta_{i,j}, & \text{for } (i,j) \in S \\ 0 & \text{else} \end{cases}$$



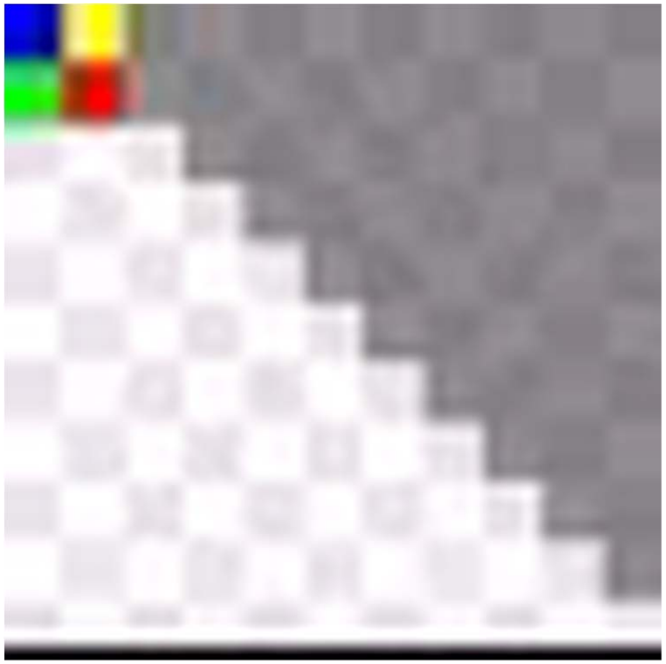



# A simple algorithm

Cholesky factorization $A = LL^T$ can be computed as

**Algorithm 1:** Cholesky factorisation

**for** $i \leftarrow 1$ *to* $N$ **do**
  $A_{i,i} \leftarrow \sqrt{A_{i,i}}$;
  **for** $j \leftarrow i+1$ *to* $N$ **do**
    **for** $k \leftarrow j$ *to* $N$ **do**
      $A_{k,j} \leftarrow A_{k,j} - A_{k,i} A_{j,i} / A_{i,i}$;
  $A_{:,i} \leftarrow A_{:,i} / \sqrt{A_{i,i}}$;
**return** LowerTriang ($A$)



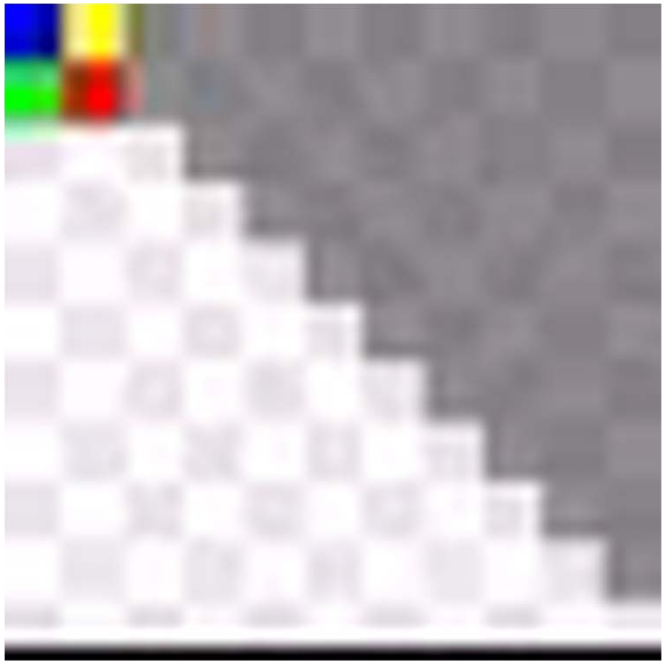
# A simple algorithm

Cholesky factorization $A = LL^T$ can be computed as

**Algorithm 1:** Cholesky factorisation

---

**for** $i \leftarrow 1$ *to* $N$ **do**
  $A_{i,i} \leftarrow \sqrt{A_{i,i}}$;
  **for** $j \leftarrow i+1$ *to* $N$ **do**
    **for** $k \leftarrow j$ *to* $N$ **do**
      $A_{k,j} \leftarrow A_{k,j} - A_{k,i}A_{j,i}/A_{i,i}$;
  $A_{:,i} \leftarrow A_{:,i}/\sqrt{A_{i,i}}$;
**return** LowerTriang $(A)$



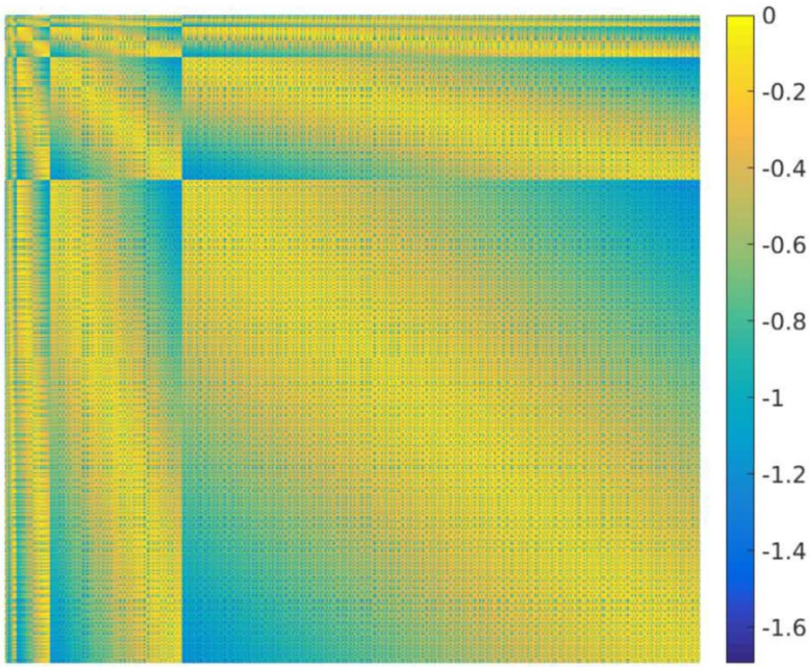
- One small Tweak: Skip all operations, for which $(k, j)$, $(k, i)$, or $(j, i)$ are outside of the sparsity pattern.

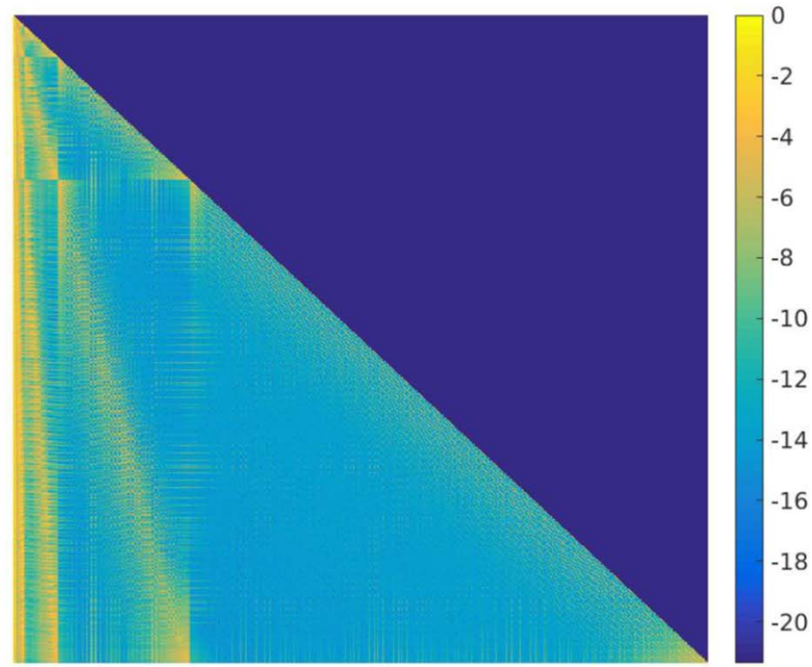
# A simple algorithm

The algorithm is oblivious to exact knowledge of the PDE and uses only the geometry of the discretisation.

## Why does it work?

- $\Theta$ has *almost* sparse Cholesky factors!



# Probabilistic interpretation of Gaussian elimination

- Assume that $X \sim \mathcal{N}(0, \Theta)$.
- Look at a single step of (Block-) Cholesky decomposition:

$$\begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix}$$
$$= \begin{pmatrix} \mathrm{Id} & 0 \\ \Theta_{21}\Theta_{11}^{-1} & \mathrm{Id} \end{pmatrix} \begin{pmatrix} \Theta_{11} & 0 \\ 0 & \Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & \Theta_{11}^{-1}\Theta_{12} \\ 0 & \mathrm{Id} \end{pmatrix}$$

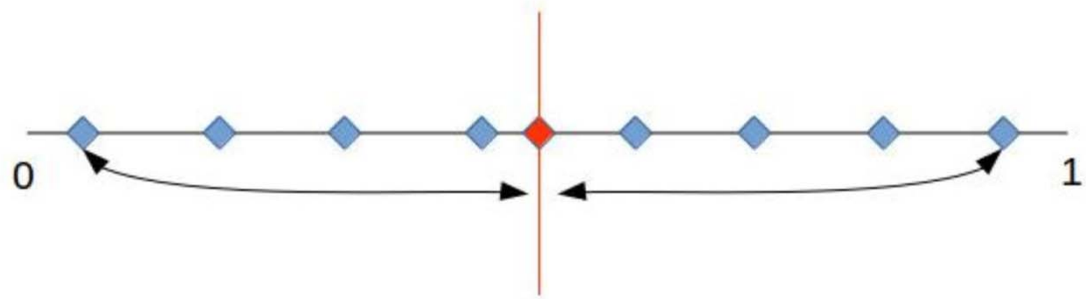
- But we know, that: $\Theta_{21}\Theta_{11}^{-1} b = \mathbb{E}[X_2 | X_1 = b]$, and $\Theta_{22} - \Theta_{21}\Theta_{11}^{-1}\Theta_{12} = \mathrm{Cov}[X_2 | X_1]$.
- Gaussian elimination $\Leftrightarrow$ conditioning of a Gaussian measure!

[ R. W. Cottle. Manifestations of the Schur complement.
Linear Algebra and its Applications , 8(3):189211, 1974. ]

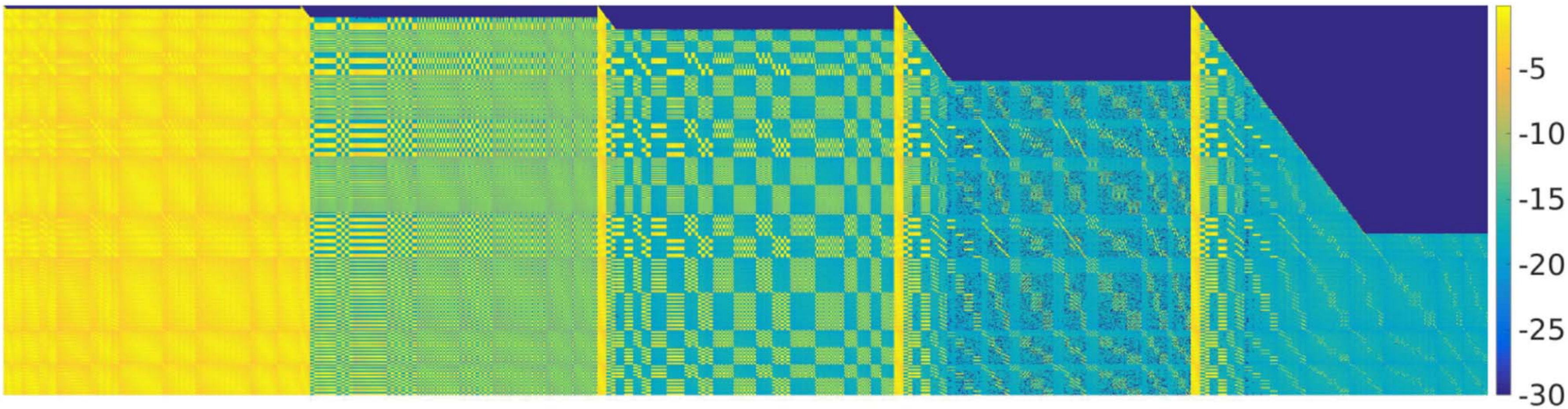
[H. Owhadi and C. Scovel.
Conditioning Gaussian measure on Hilbert space. 2015. arXiv:1506.04208 ]

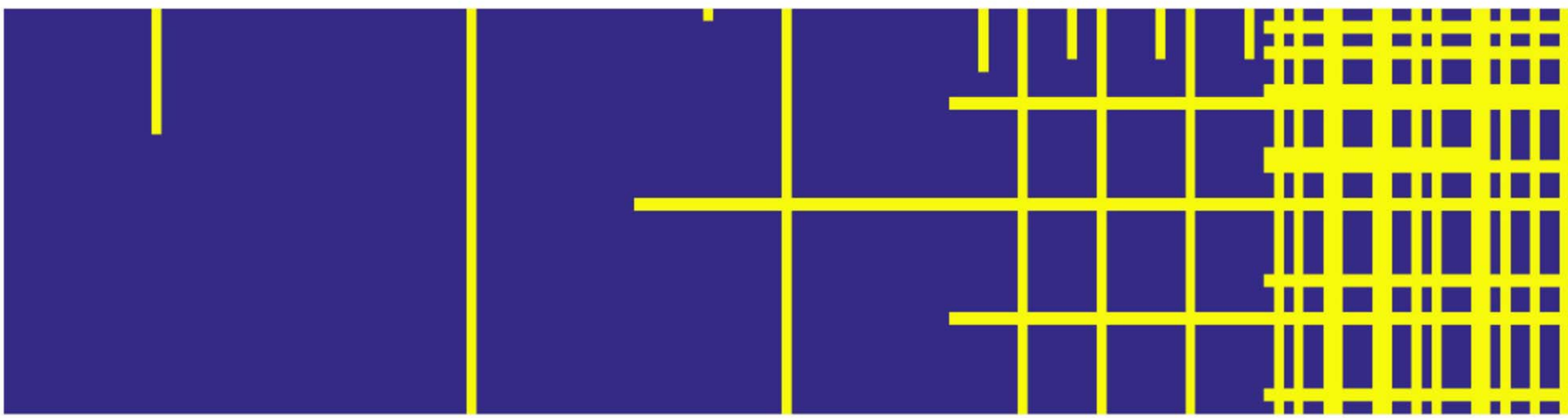
# The screening effect

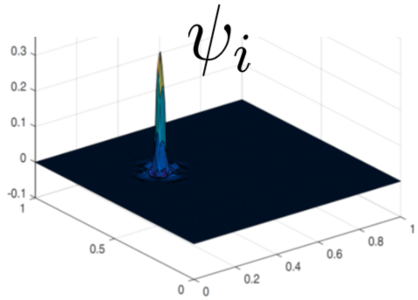
- Many *smooth* random fields exhibit the *screening effect*:
- $X_i$ and $X_j$ decorrelate upon conditioning on intermediate $X_k$.



# Fade-out vs Fill-in



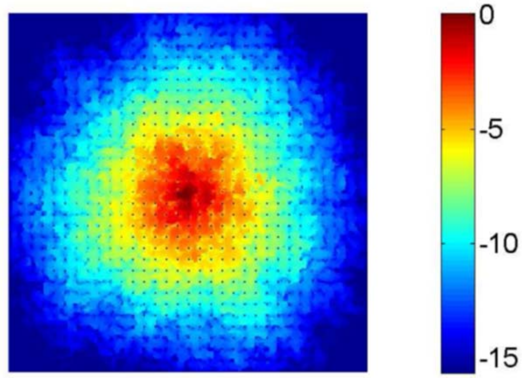



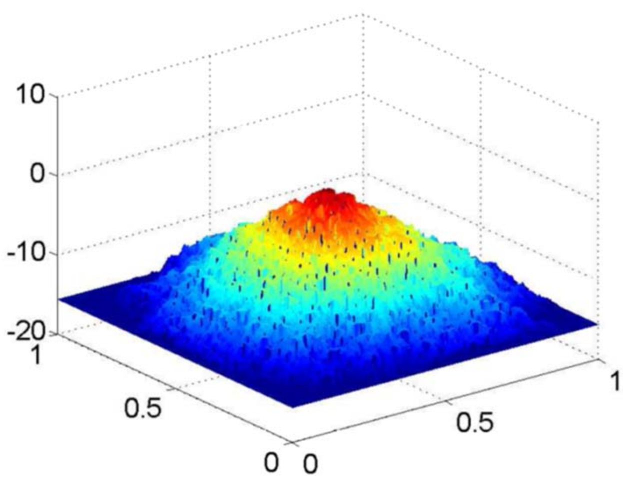


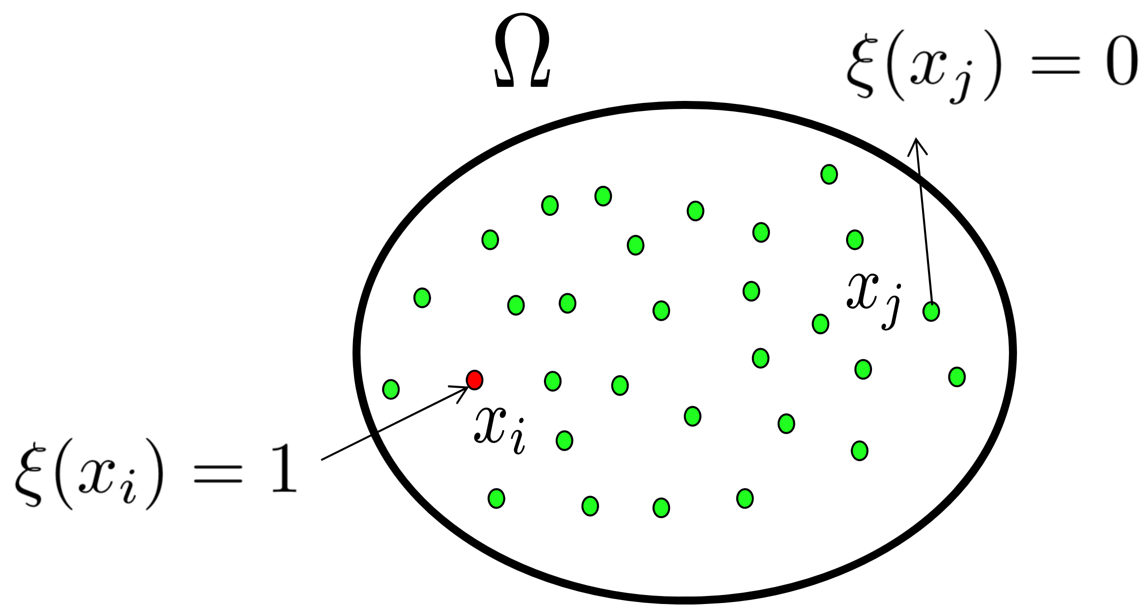

# Localization of Gamblets

$$\xi \sim \mathcal{N}(0, \mathcal{L}^{-1})$$

$$\psi_i(x) = \mathbb{E}\big[\xi(x) \mid \xi(x_j) = \delta_{i,j} \text{ for } j \in \mathcal{I}\big]$$



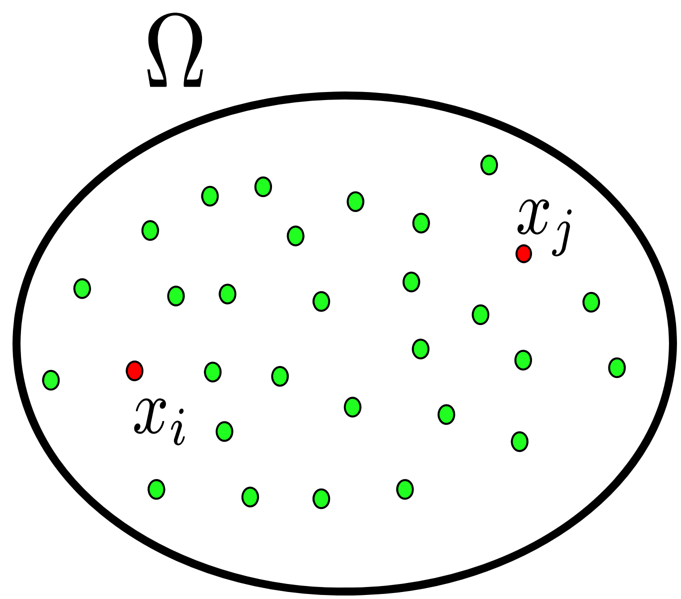



# Sparsity of the precision matrix

$$\Theta_{i,j} = \mathrm{Cov}\left(\xi(x_i), \xi(x_j)\right)$$

$$\Theta^{-1}_{i,j} = \left\langle \psi_i, \psi_j \right\rangle$$




$$\Theta^{-1}_{i,j} = 0$$

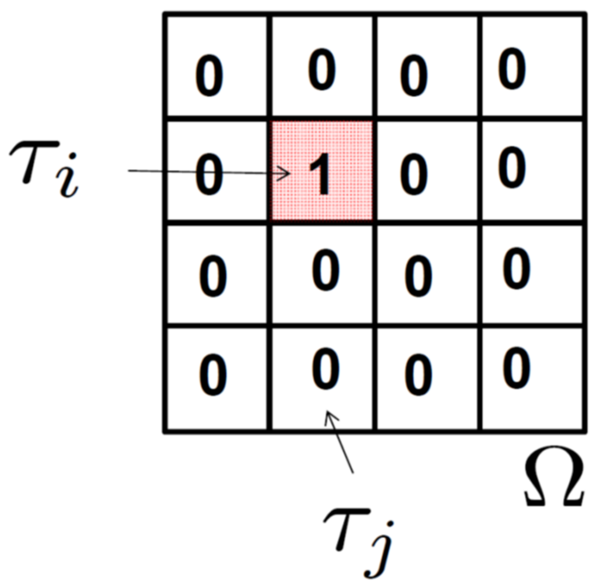
$$\Updownarrow$$

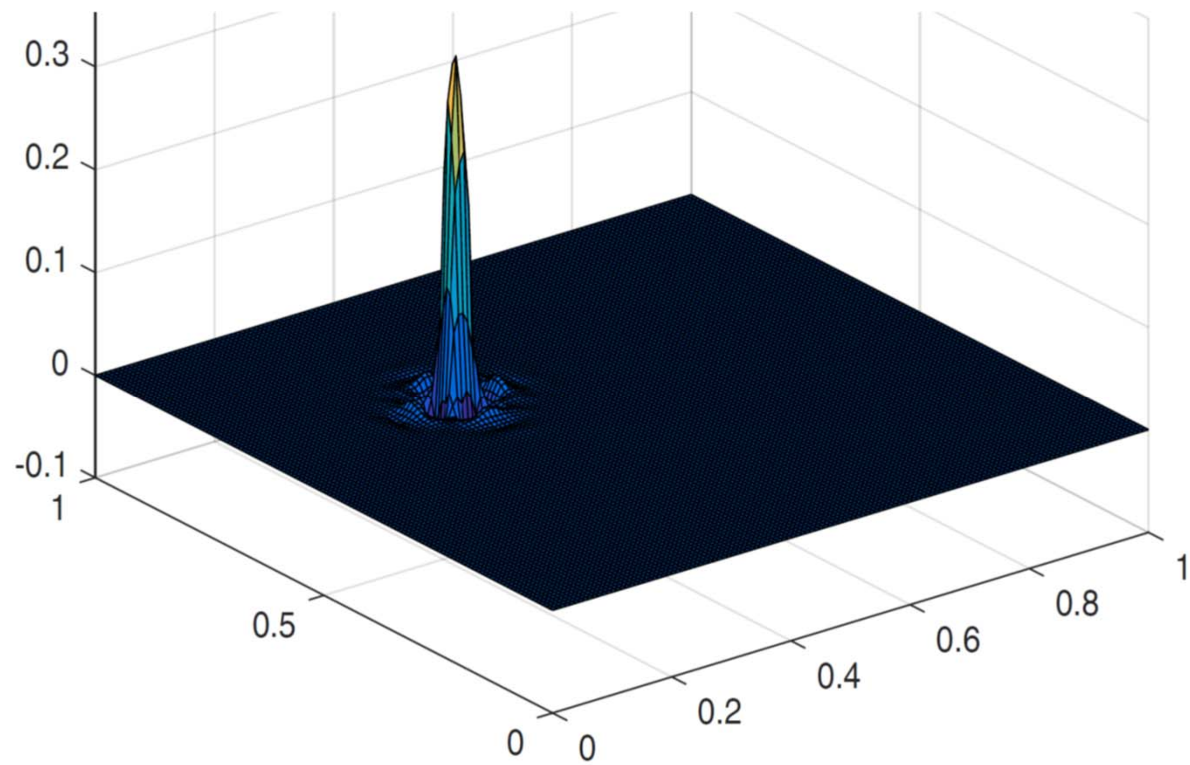
$$\mathrm{Cov}\left(\xi(x_i), \xi(x_j) \middle| \xi(x_l), l \neq i, j]\right) = 0$$

# Localization of Gamblets

$$\xi \sim \mathcal{N}(0, \mathcal{L}^{-1})$$

$$\psi_i = \mathbb{E}\Big[\xi \mid [\phi_j, \xi] = \delta_{i,j} \text{ for } j \in \mathcal{I}\Big]$$



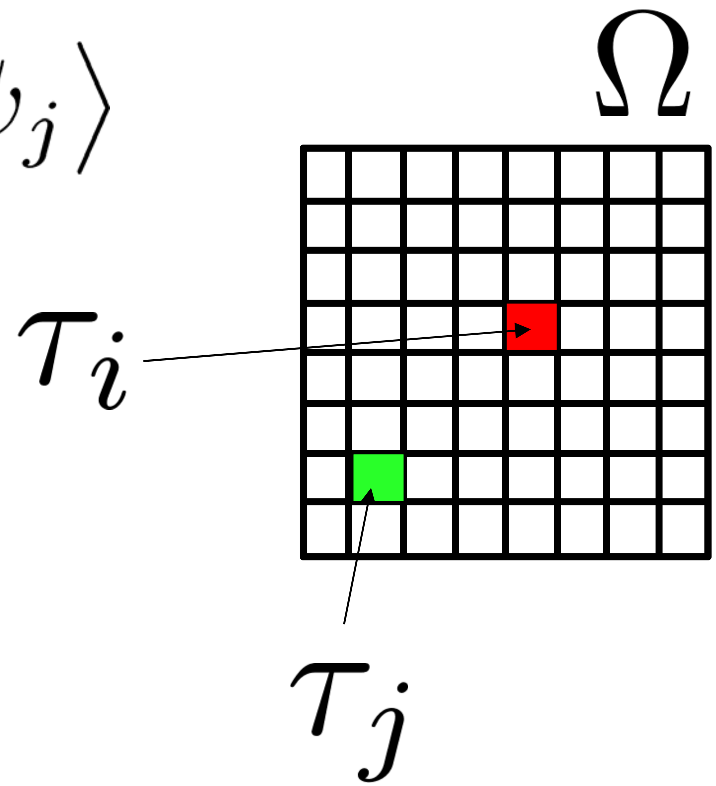




$$\phi_i = 1_{\tau_i}$$

# Sparsity of the precision matrix

$$\Theta_{i,j} = \mathrm{Cov}\left([\phi_i, \xi], [\phi_j, \xi]\right)$$

$$\Theta^{-1}_{i,j} = \langle \psi_i, \psi_j \rangle$$



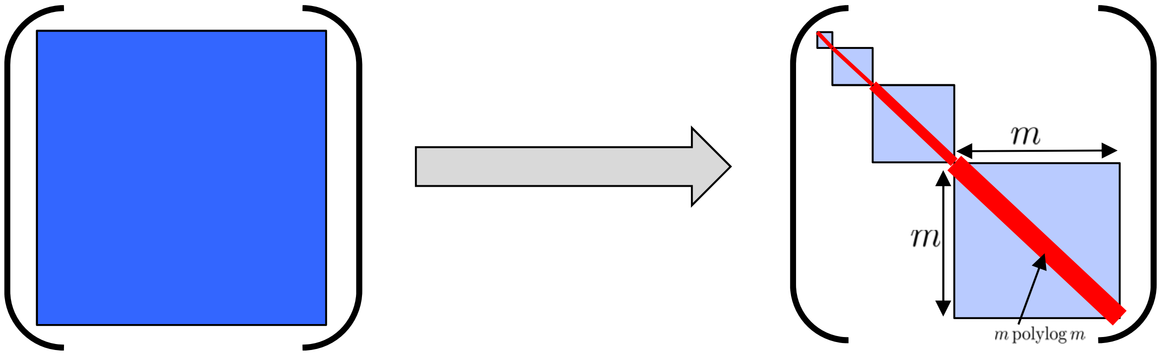

$$\Theta^{-1}_{i,j} = 0$$

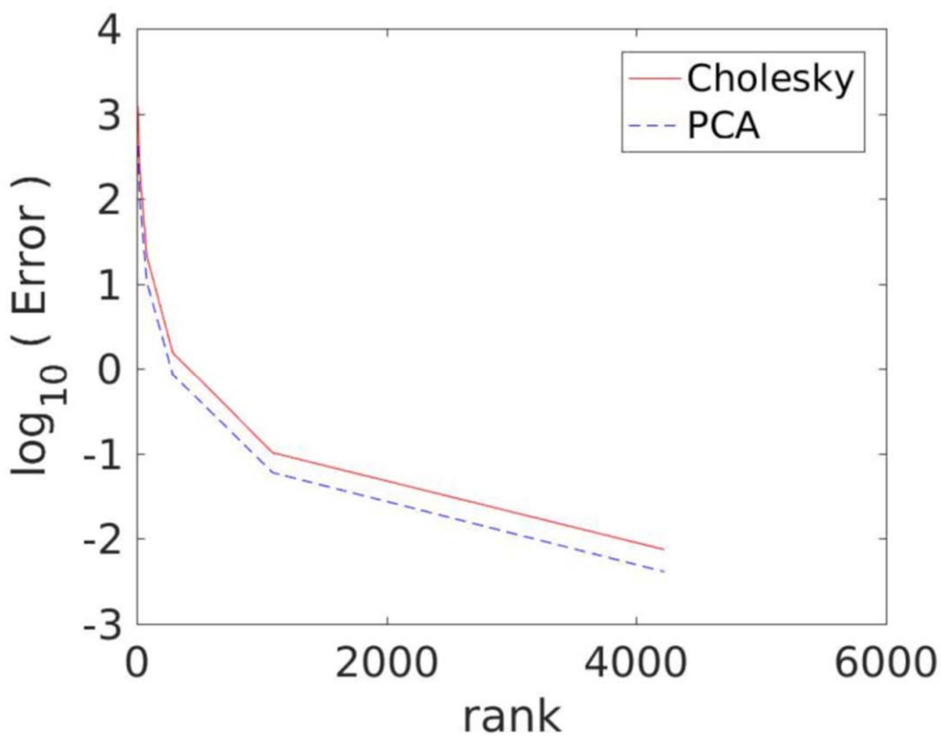
$$\Updownarrow$$

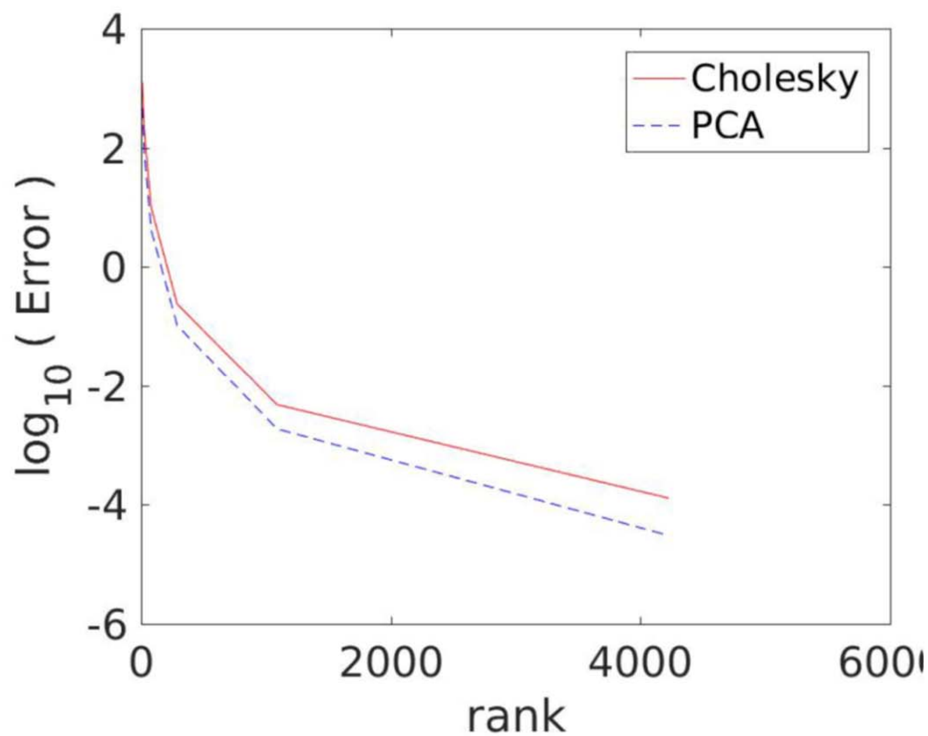
$$\mathrm{Cov}\left([\phi_i, \xi], [\phi_j, \xi] \middle| [\phi_l, \xi], \, l \neq i, j\right) = 0$$

# Sparse approximate PCA

- Block diagonalize $\Theta$ into sparse well conditionned blocks

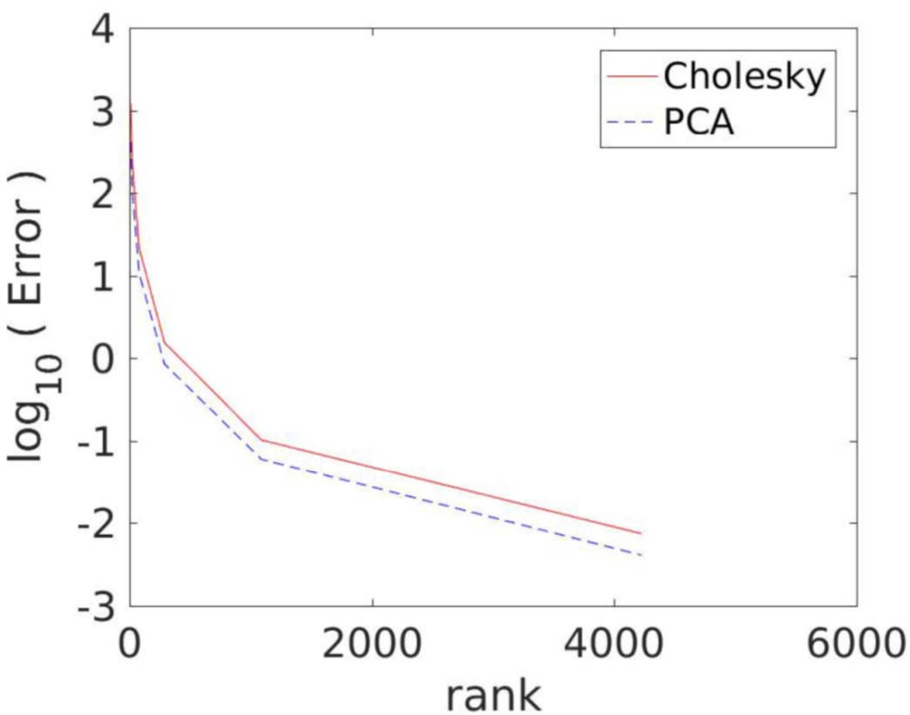


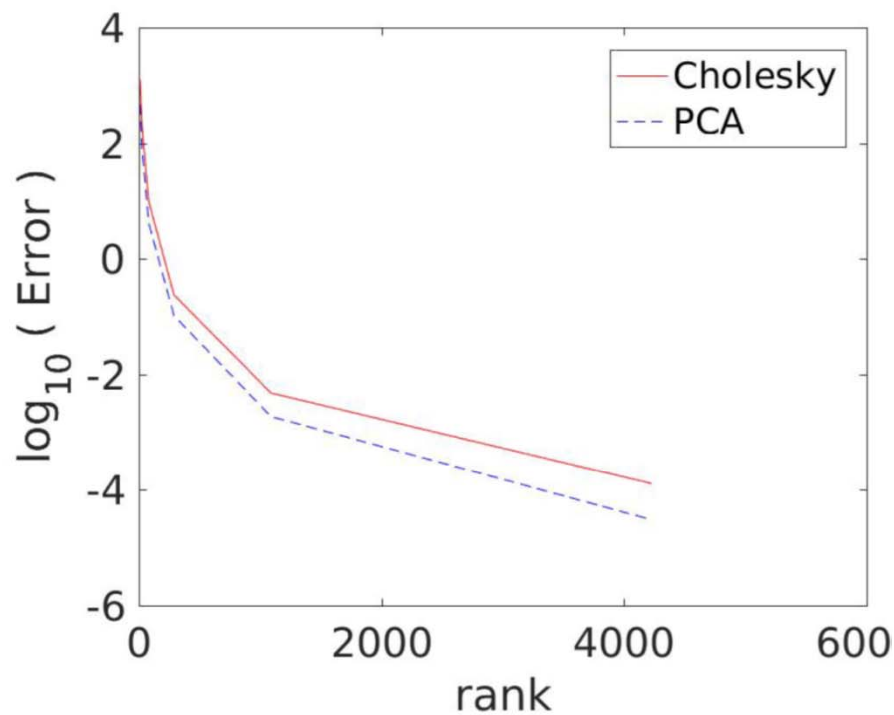



# Sparse approximate PCA

- Sparse approximate PCA inherited from Gamblet transform:
- $L(:,1:k)\,L(:,1:k)^T$ provides near optimal rank $k$ approximation.




- Simultaneous exploitation of sparsity and spectral decay.

# Problems at the boundary

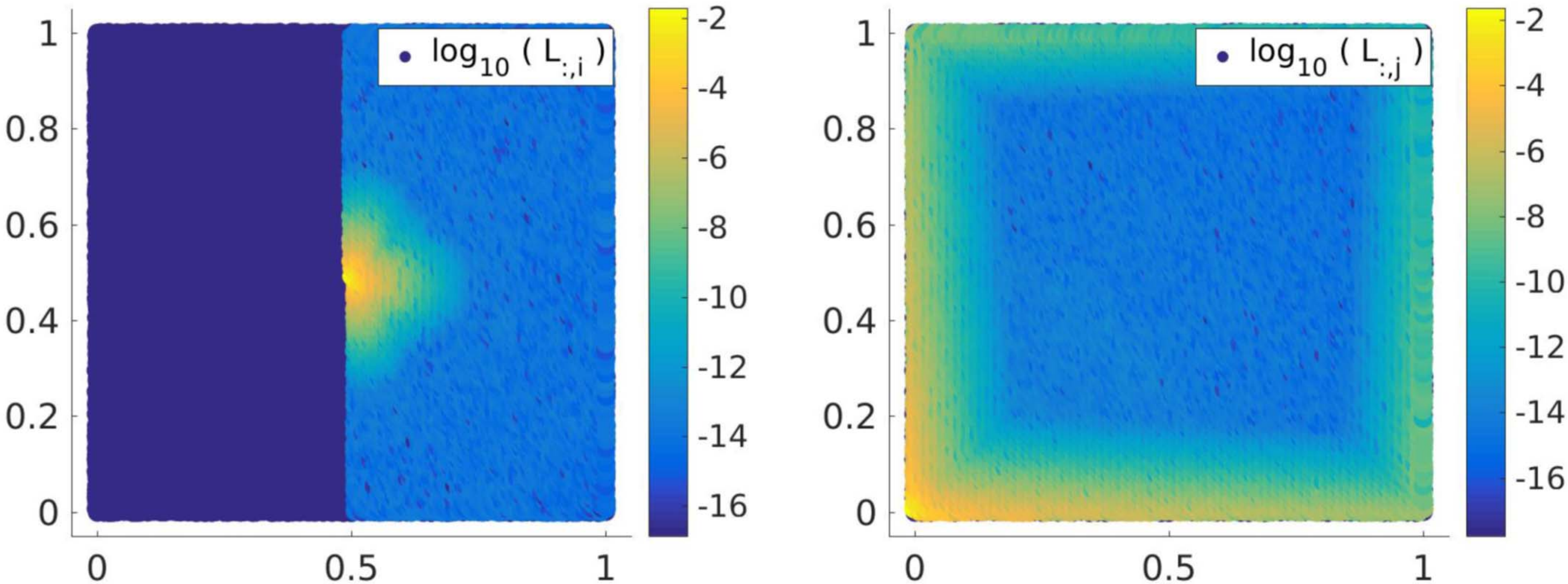
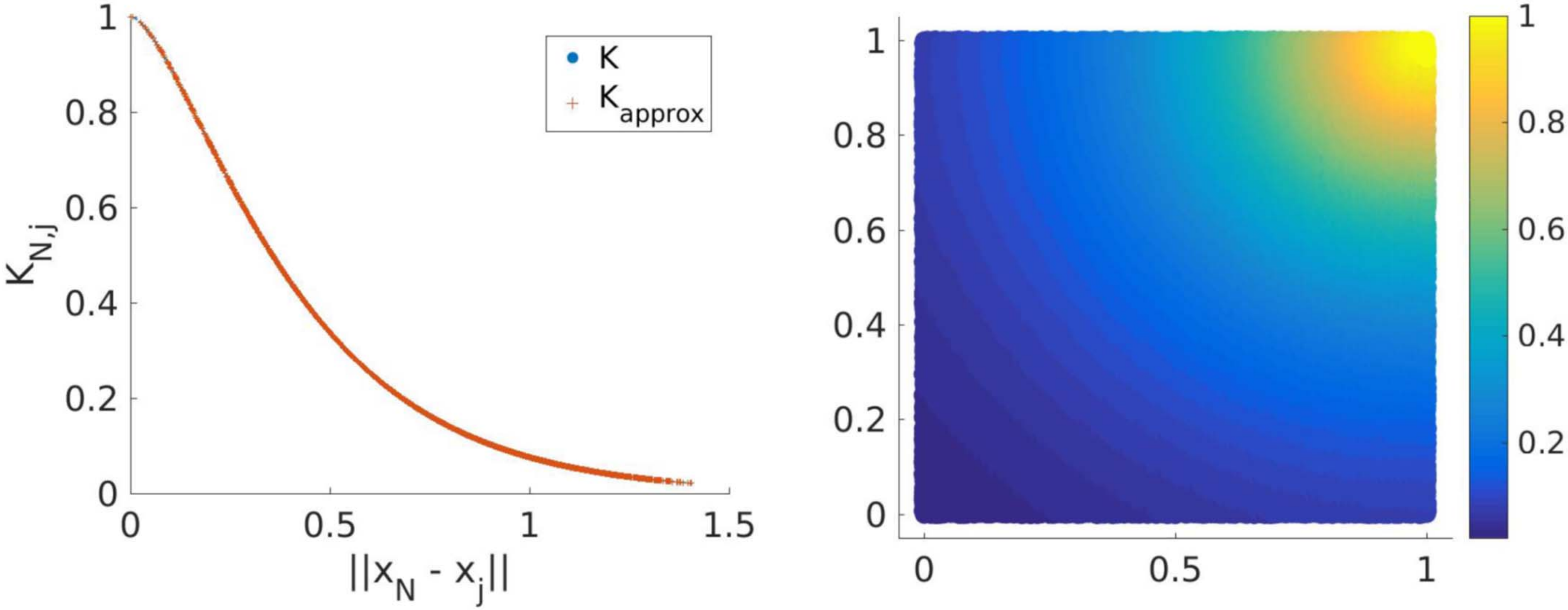
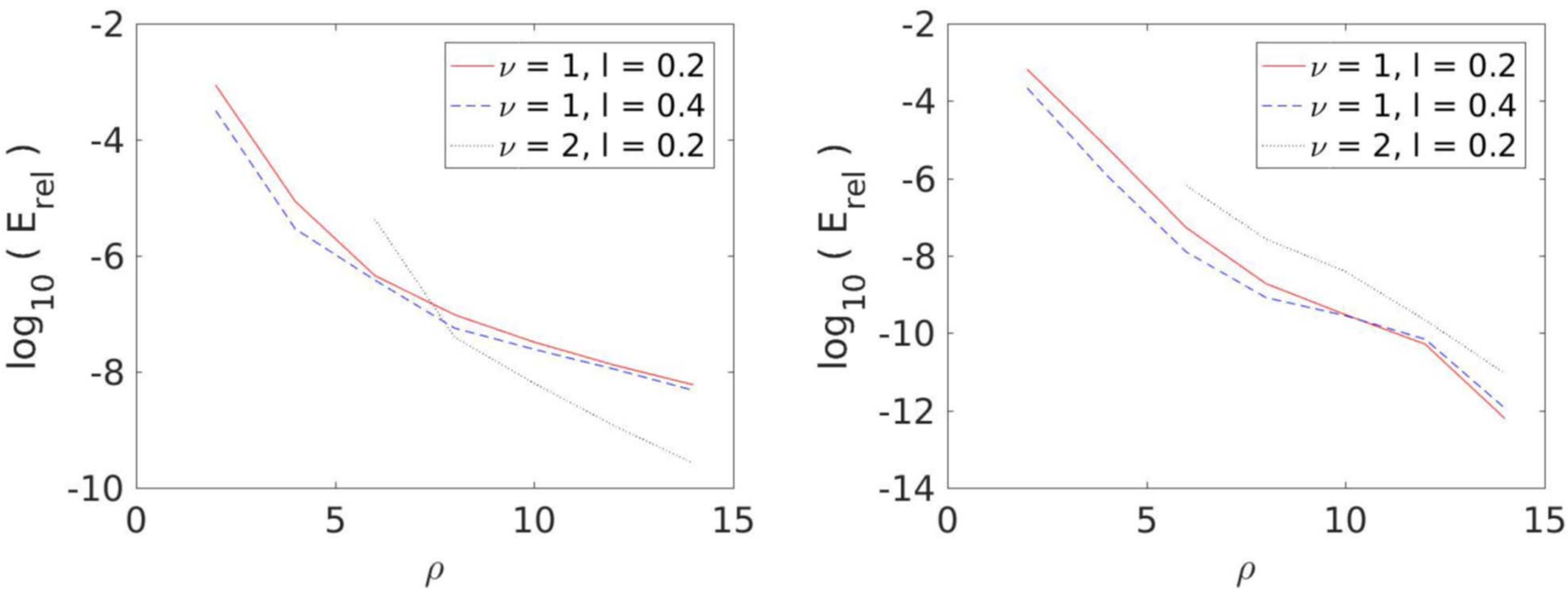



Figure: $\nu = 1$, $l = 0.4$

# Problems at the boundary



# Decay of the approximation error



# Sparse approximate PCA

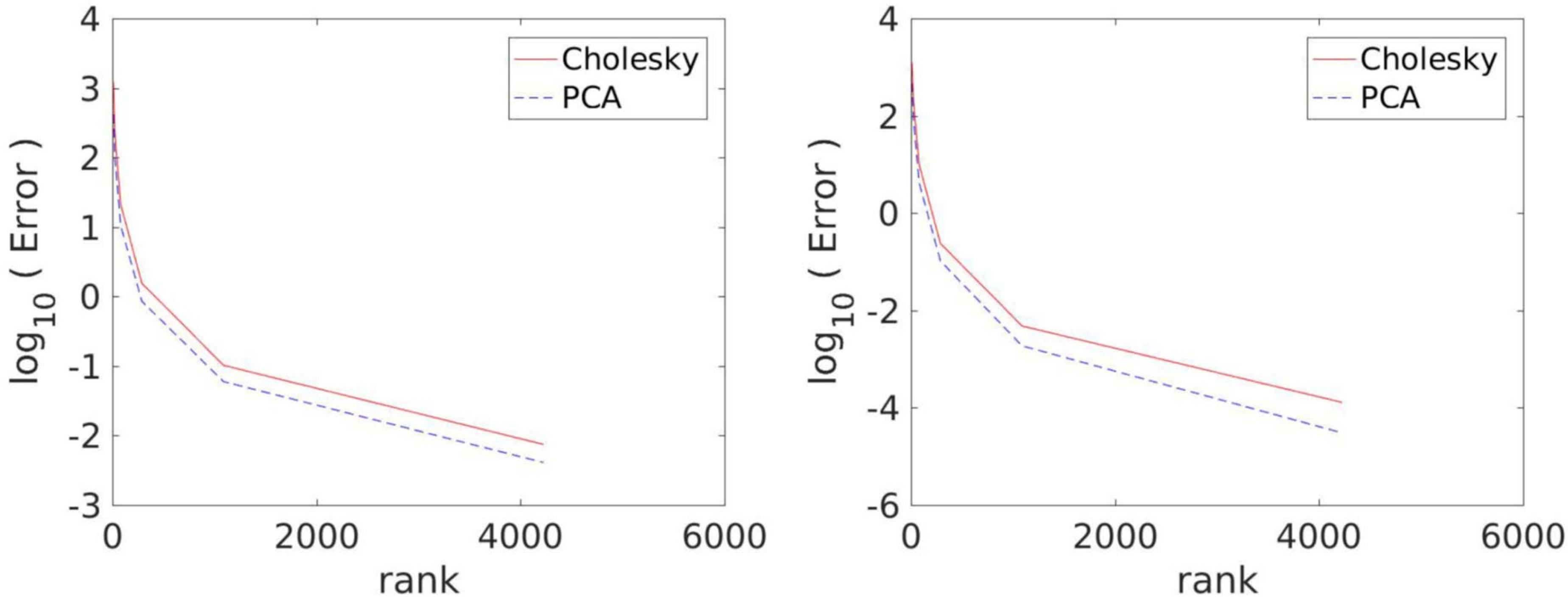
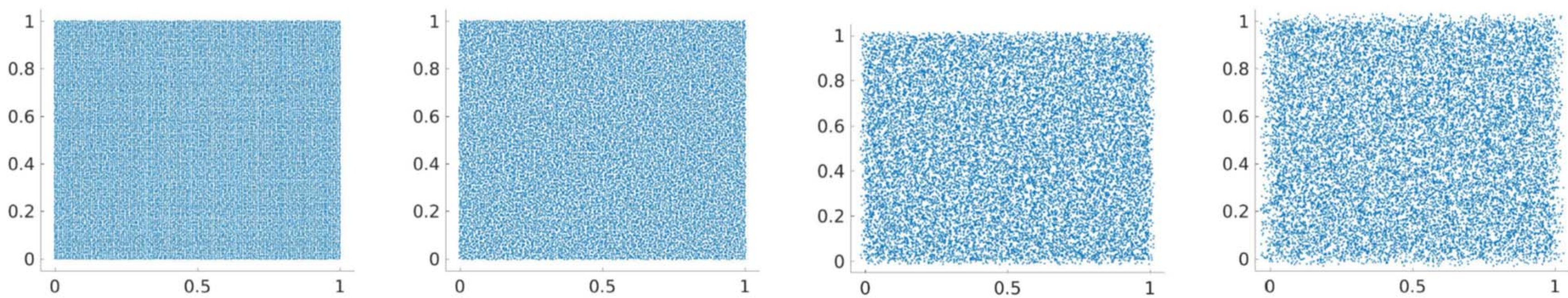



Figure: Near optimal sparse PCA: First panel: $\nu = 1$, $l = 0.2$, $\delta_x = 0.2$ and $\rho = 6$. Second panel: $\nu = 2$, $l = 0.2$ and $\delta_x = 0.2$ and $\rho = 8$.

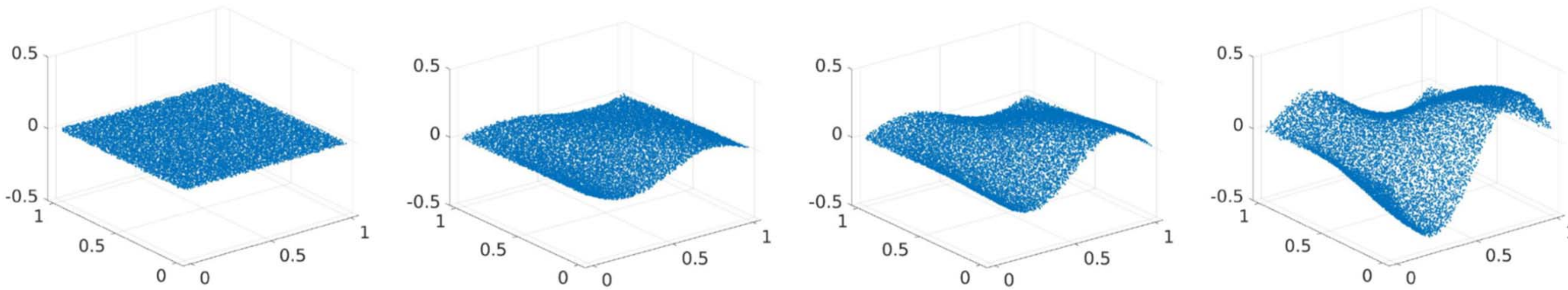
# Perturbation of the Mesh




| $\delta_x$ | $\|\Gamma^\rho - \Gamma\|$ | $\|\Gamma^\rho - \Gamma\|/\|\Gamma\|$ | $\|\Gamma^\rho - \Gamma\|_{\text{Fro}}$ | $\|\Gamma^\rho - \Gamma\|_{\text{Fro}}/\|\Gamma\|_{\text{Fro}}$ | $\#S$ | $\#S/N^2$ |
|---|---|---|---|---|---|---|
| 0.2 | 4.336e-03 | 1.560e-06 | 1.669e-02 | 1.026e-06 | 2.125e+07 | 7.675e-02 |
| 0.4 | 4.495e-03 | 1.617e-06 | 1.706e-02 | 1.063e-06 | 2.128e+07 | 7.683e-02 |
| 2.0 | 4.551e-03 | 1.638e-06 | 1.820e-02 | 1.077e-06 | 2.127e+07 | 7.682e-02 |
| 4.0 | 8.158e-03 | 2.940e-06 | 2.976e-02 | 1.933e-06 | 2.119e+07 | 7.652e-02 |

Table: Compression and accuracy for $q = 7,\ l = 0.2,\ \rho = 5,\ \nu = 1$ and different values of $\delta_x$.

# Data on low dimensional manifold




| $\delta_z$ | $\|\Gamma^\rho - \Gamma\|$ | $\|\Gamma^\rho - \Gamma\|/\|\Gamma\|$ | $\|\Gamma^\rho - \Gamma\|_{\text{Fro}}$ | $\|\Gamma^\rho - \Gamma\|_{\text{Fro}}/\|\Gamma\|_{\text{Fro}}$ | $\#S$ | $\#S/N^2$ |
|---|---|---|---|---|---|---|
| 0.0 | 5.049e-03 | 1.560e-06 | 1.885e-02 | 1.026e-06 | 2.126e+07 | 7.677e-02 |
| 0.1 | 6.341e-02 | 1.648e-06 | 1.232e-01 | 1.077e-06 | 2.083e+07 | 7.521e-02 |
| 0.2 | 1.204e-01 | 1.749e-06 | 2.203e-01 | 1.126e-06 | 1.976e+07 | 7.137e-02 |
| 0.4 | 1.954e-01 | 3.550e-06 | 5.098e-01 | 2.197e-06 | 1.722e+07 | 6.218e-02 |

Table: Compression and accuracy for $q = 7,\ l = 0.2,\ \rho = 5,\ \nu = 1,\ \delta_x = 2$ and different values of $\delta_z$.

# Fractional Operators

| $\nu$ | $\|\Gamma^{\rho} - \Gamma\|$ | $\|\Gamma^{\rho} - \Gamma\| / \|\Gamma\|$ | $\|\Gamma^{\rho} - \Gamma\|_{\text{Fro}}$ | $\|\Gamma^{\rho} - \Gamma\|_{\text{Fro}} / \|\Gamma\|_{\text{Fro}}$ | $\#S$ | $\#S/N^2$ |
|---|---|---|---|---|---|---|
| 1.0 | 1.266e-03 | 4.556e-07 | 4.987e-03 | 2.995e-07 | 2.776e+07 | 1.003e-01 |
| 1.1 | 1.813e-03 | 6.423e-07 | 6.216e-03 | 4.190e-07 | 2.776e+07 | 1.003e-01 |
| 1.3 | 3.235e-03 | 1.129e-06 | 1.039e-02 | 7.312e-07 | 2.776e+07 | 1.003e-01 |
| 1.5 | 5.245e-03 | 1.811e-06 | 1.652e-02 | 1.166e-06 | 2.776e+07 | 1.003e-01 |
| 1.6 | 6.800e-03 | 2.333e-06 | 2.148e-02 | 1.498e-06 | 2.776e+07 | 1.003e-01 |
| 1.8 | 9.891e-03 | 3.362e-06 | 3.088e-02 | 2.147e-06 | 2.776e+07 | 1.003e-01 |
| 2.0 | 1.238e-02 | 4.180e-06 | 3.892e-02 | 2.662e-06 | 2.776e+07 | 1.003e-01 |

Table: Compression and accuracy for $q = 7$, $l = 0.2$, $\rho = 6$, $\delta_x = 0.2$ and different values of $\nu$.

# Some references

## 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

[Beylkin, Coifman, Rokhlin, 1992] [Engquist, Osher, Zhong, 1992]

[Alpert, Beylkin, Coifman, Rokhlin, 1993]

[Cohen, Daubechies, Feauveau. 1992]

[Bacry, Mallat, Papanicolaou. 1993]

## 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]

# Low Rank Matrix Decomposition methods

Fast Multipole Method: [Greengard and Rokhlin, 1987]

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

[Bebendorf, 2008]:

# Hierarchical numerical homogenization method

[H. Owhadi, Multigrid with rough coefficients and Multiresolution operator decomposition from Hierarchical Information Games. SIAM Review, 2017]

H. Owhadi and C. Scovel. Universal Scalable Robust Solvers from Computational Information Games and fast eigenspace adapted Multiresolution Analysis, 2017. arXiv:1703.10761

Gamblets for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic ODEs/PDEs with rough coefficients, 2016. H. Owhadi and L. Zhang. arXiv:1606.07686

An adaptive fast solver for a general class of positive definite matrices via energy decomposition Thomas Y. Hou, D. Huang, K.C. Lam, P. Zhang. arXiv:1707.08277. 2017

## Sparse matrix Laplacians

Sparsified Cholesky and Multigrid Solvers for Connection Laplacians: [Kyng, Lee, Peng, Sachdeva, Spielman , 2016]

Approximate Gaussian Elimination: [Kyng and Sachdeva, 2016]

## Structured sparse matrices (SDD matrices)

Graph sparsification: [Spielman and Teng , 2004]

Diagonally dominant linear systems: [Spielman and Teng , 2014]

[Koutis, Miller, Gary and Peng , 2014]

[Cohen, Kyng, Miller, Pachocki, Peng, Rao, and Xu, 2014]

[Kelner, Orecchia, Sidford, Zhu, 2013]

## Localization problem in Numerical Homogenization

[Chu-Graham-Hou-2010] (limited inclusions)

[Efendiev-Galvis-Wu-2010] (limited inclusions or mask)

[Babuska-Lipton 2010] (local boundary eigenvectors)

[Owhadi-Zhang 2011] (localized transfer property)

[Malqvist-Peterseim 2012] Local Orthogonal Decomposition

[Owhadi-Zhang-Berlyand 2013] (Rough Polyharmonic Splines)

[ A. Gloria, S. Neukamm, and F. Otto, 2015] (quantification of ergodicity)

[Hou and Liu,DCDS-A, 2016] [Chung-Efendiev-Hou, JCP 2016]

[Owhadi, Multiresolution operator decomposition, SIREV 2017]

[Owhadi, Zhang, gamblets for hyperbolic and parabolic PDEs, 2016]

[Hou, Qin, Zhang, 2016] [Hou, Zhang, 2017]

[Hou and Zhang, 2017]: Higher order PDEs (localization under strong ellipticity, $h$ sufficiently small, and higher order polynomials as measurement functions)

[Kornhuber, Peterseim, Yserentant, 2016]: Subspace decomposition

## Subspace decomposition/correction and Schwarz iterative methods

[J. Xu, 1992]: Iterative methods by space decomposition and subspace correction

[Griebel-Oswald, 1995]: Schwarz algorithms

# Numerical Homogenization

**Harmonic Coordinates** Babuska, Caloz, Osborn, 1994

Kozlov, 1979 Allaire Brizzi 2005; Owhadi, Zhang 2005

**MsFEM** [Hou, Wu: 1997]; [Efendiev, Hou, Wu: 1999]

[Fish - Wagiman, 1993] [Chung-Efendiev-Hou, JCP 2016]

**Variational Multiscale Method, Orthogonal decomposition**

[Hughes, Feijóo, Mazzei, Quincy. 1998]

[Malqvist-Peterseim 2012] Local Orthogonal Decomposition

**Projection based method** Nolen, Papanicolaou, Pironneau, 2008

**HMM** Engquist, E, Abdulle, Runborg, Schwab, et Al. 2003-...

**Flux norm** Berlyand, Owhadi 2010; Symes 2012

**Bayesian Numerical Homogenization** Owhadi 2014

**Gamblets – Operator compression** [Owhadi, SIREV 2017]

[Owhadi, Zhang, 2016] [Hou, Qin, Zhang, 2016]

[Hou, Zhang, 2017]

# Thank you


- **Universal Scalable Robust Solvers from Computational Information Games and fast eigenspace adapted Multiresolution Analysis, 2017. arXiv:1703.10761. H. Owhadi and C. Scovel.**
- Compression, inversion, and approximate PCA of dense kernel matrices at near-linear computational complexity, Schäfer, Sullivan, Owhadi. 2017.
- Multigrid with gamblets. L. Zhang and H. Owhadi, 2017
- Gamblets for opening the complexity-bottleneck of implicit schemes for hyperbolic and parabolic ODEs/PDEs with rough coefficients, 2016. H. Owhadi and L. Zhang. arXiv:1606.07686
- Multigrid with rough coefficients and Multiresolution operator decomposition from Hierarchical Information Games. H. Owhadi. SIAM Review, 59(1), 99149, 2017. arXiv:1503.03467
- Towards Machine Wald (book chapter). Houman Owhadi and Clint Scovel. Springer Handbook of Uncertainty Quantification, 2016, arXiv:1508.02449.
- Bayesian Numerical Homogenization. H. Owhadi. SIAM Multiscale Modeling & Simulation, 13(3), 812828, 2015. arXiv:1406.6668


Florian Schäfer

Clint Scovel

Tim Sullivan

Lei Zhang


DARPA EQUiPS / AFOSR award no FA9550-16-1-0054 (Computational Information Games)