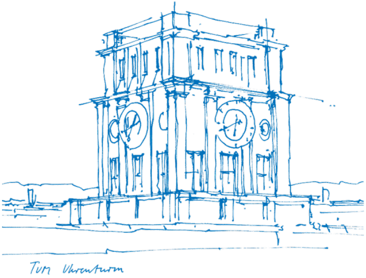
Algorithms for Scientific Computing

Hierarchical Methods and Sparse Grids
– *d*-Dimensional Hierarchical Basis –

Michael Bader
Technical University of Munich

Summer 2022

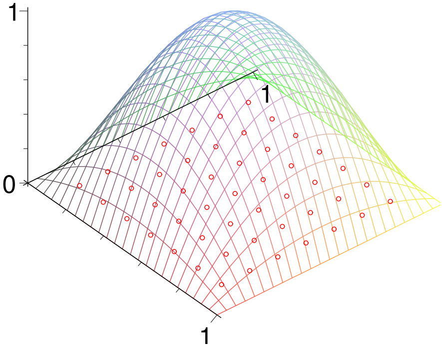


Intermezzo/"Big Picture": Archimedes' Quadrature

- Start with $2d$ example (compare tutorials):

$$f := 16x_1(x_1-1)x_2(x_2-1), \qquad \Omega = [0,1]^2 \qquad \Rightarrow f|_{\partial\Omega} = 0$$

- Consider hierarchical surplus at grid points with $n = 3$, $h_3 = 2^{-3}$

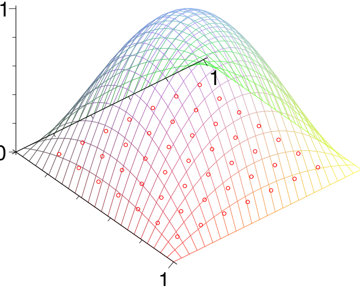


$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{4}$	$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	1	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$
$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{4}$	$\frac{1}{64}$	$\frac{1}{16}$	$\frac{1}{64}$
$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{256}$	$\frac{1}{64}$	$\frac{1}{256}$

"Big Picture": Archimedes' Quadrature (2)

$$\int_\Omega f\, d\vec{x} = 4/9 = 0.\bar{4} \qquad \sum = \frac{441}{1024} = 0.4306640625$$

- Consider volume of subvolumes (pagodas) for quadrature

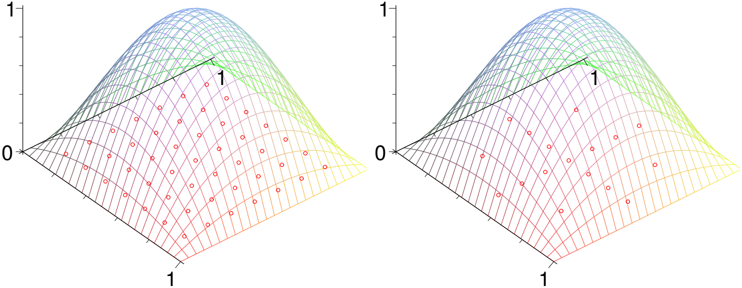


$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$	$\frac{1}{256}$	$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$
$\frac{1}{2048}$	$\frac{1}{256}$	$\frac{1}{2048}$	$\frac{1}{32}$	$\frac{1}{2048}$	$\frac{1}{256}$	$\frac{1}{2048}$
$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$	$\frac{1}{256}$	$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$
$\frac{1}{256}$	$\frac{1}{32}$	$\frac{1}{256}$	$\frac{1}{4}$	$\frac{1}{256}$	$\frac{1}{32}$	$\frac{1}{256}$
$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$	$\frac{1}{256}$	$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$
$\frac{1}{2048}$	$\frac{1}{256}$	$\frac{1}{2048}$	$\frac{1}{32}$	$\frac{1}{2048}$	$\frac{1}{256}$	$\frac{1}{2048}$
$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$	$\frac{1}{256}$	$\frac{1}{16384}$	$\frac{1}{2048}$	$\frac{1}{16384}$

"Big Picture": Archimedes' Quadrature (3)

What, if we leave out (adaptively) all subvolumes with volume $< \epsilon = \frac{1}{256}$?

- 49 grid points (full grid) $\Rightarrow$ 17 grid points (*sparse grid*)



- Approximation of volume:

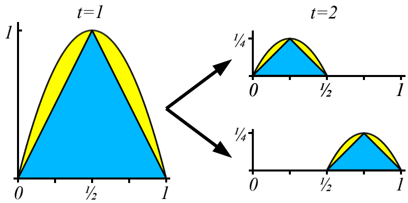
$$\frac{441}{1024} = 0.4306640625 \qquad \Rightarrow \qquad \frac{27}{64} = 0.421875$$

Part I

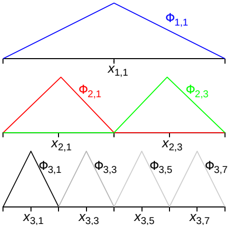
Hierarchical Decomposition, *d*-Dimensional

Recall: Archimedes and Hierarchical Basis (in 1D)

Archimedes Quadrature:



Hierarchical Basis:



- use nodal basis functions $\phi_{l,i}$ with $\mathcal{I}_l := \{i : 1 \leq i < 2^l,\ i \text{ odd}\}$
  $\rightsquigarrow$ hierarchical basis $\Psi_n := \bigcup_{l=1}^{n} \{\phi_{l,i} : i \in \mathcal{I}_l\}$
- hierachical function spaces $W_l := \text{span}\,\{\phi_{l,i} :\ i \in \mathcal{I}_l\}$ and $V_l = V_{l-1} \oplus W_l$
- unique hierachical representation $u = \sum_{l=1}^{n} w_l = \sum_{l=1}^{n} \sum_{i \in \mathcal{I}_l} v_{l,i} \phi_{l,i}$
- size of surpluses $v_{l,i}$ roughly decays with 4^{-n} for smooth functions

Hierarchical Decomposition – Step by Step

Now (and more formally), starting with d-dimensional hierarchical decompositions ...

Transfer from $d = 1$ to $d > 1$

- Functions in multiple variables $\vec{x} = (x_1, \dots, x_d)$
- Domain $\Omega := [0, 1]^d$
- We consider only functions u which are 0 on $\partial\Omega$
  (on the edges of the square, faces of the cube, ...)
- Each hierarchical grid described by multi-index

$$\vec{l} = (l_1, \dots, l_d) \in \mathbb{N}^d$$

- Grids have different mesh-widths in different dimensions:

$$\vec{h}_{\vec{l}} := (h_1, \dots, h_d) := (2^{-l_1}, \dots, 2^{-l_d}) =: 2^{-\vec{l}}$$

Hierarchical Decomposition, $d > 1$

Introducing further notation (which we'll need later on):

- Grid points (for function evaluations):

$$\vec{x}_{\vec{l},\vec{i}} = (i_1 \cdot h_{l_1}, \ldots, i_d \cdot h_{l_d})$$

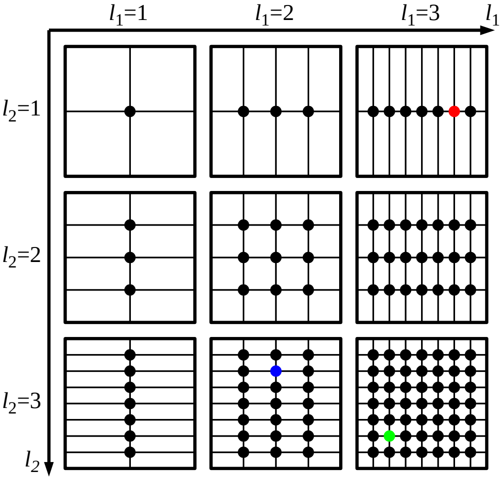
- Comparisons of multi-indices component-wise:

$$\vec{l} \leq \vec{i} \qquad \Longleftrightarrow \qquad l_k \leq i_k,\ k = 1, \ldots, d$$

- Two norms for multi-indices $\vec{l}$
  - index sum: $|\vec{l}|_1 := |l_1| + \ldots + |l_d|$
  - maximum index: $|\vec{l}|_\infty := \max\{|l_1|, \ldots, |l_d|\}$

Note: taking the absolute values, $|\cdot|$, for $l_k \in \mathbb{N}$ is not necessary, but is part of the usual definition of $|\cdot|_1$ and $|\cdot|_\infty$

Practicing Identifiers $\vec{l}$, $\vec{h}_{\vec{l}}$, $\vec{x}_{\vec{l},\vec{i}}$



Piecewise *d*-linear Functions

Suitable generalization of piecewise linear functions

- Piecewise d-linear functions w.r.t. $\vec{h}_{\vec{l}}$ grid
  $\rightarrow$ If you fix $d-1$ coordinates, they are linear in remaining x_j
- $V_{\vec{l}}$: space of all functions for given $\vec{l}$

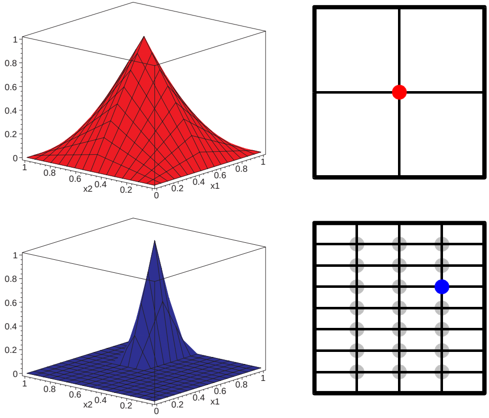
Alternative point of view:

- Define suitable basis $\Phi_{\vec{l}}$
- Regard $V_{\vec{l}}$ as span of $\Phi_{\vec{l}}$
- d-dimensional basis functions:
  products of one-dimensional hat functions:

$$\phi_{\vec{l},\vec{i}}(\vec{x}) = \prod_{j=1}^{d} \phi_{l_j,i_j}(x_j) = \phi_{l_1,i_1}(x_1) \cdot \phi_{l_2,i_2}(x_2) \cdot \ldots \cdot \phi_{l_d,i_d}(x_d)$$

d-dimensional Basis Functions

- Basis functions are *pagoda functions* (not pyramids!)
- Examples: $\phi_{(1,1),(1,1)}$, and $\phi_{(2,3),(3,5)}$:



Function Spaces $V_{\vec{l}}$ and V_n

- Basis for space of piecewise linear functions w.r.t. $h_{\vec{l}}$ grid

$$\Phi_{\vec{l}} := \{\phi_{\vec{l},\vec{i}}, \vec{1} \leq \vec{i} < 2^{\vec{l}}\}$$

- Function space

$$V_{\vec{l}} := \mathrm{span}\{\Phi_{\vec{l}}\}$$

with

$$\dim V_{\vec{l}} = (2^{l_1} - 1) \cdot \ldots \cdot (2^{l_d} - 1) \in O(2^{|\vec{l}|_1})$$

- Special case $l_1 = \ldots = l_d \rightsquigarrow$ function space denoted as V_n:

$$V_n := V_{(n,\ldots,n)}$$

Hierarchical Increments $W_{\vec{l}}$

Analogous to 1d:

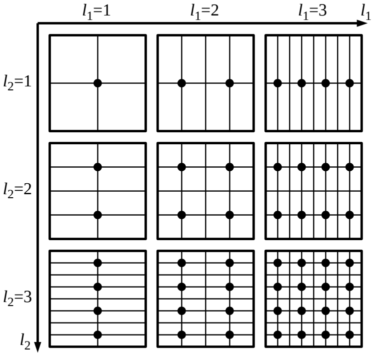
- Omit grid points with even index (exist on coarser grid)
- Now in all directions

$$\mathcal{I}_{\vec{l}} := \{\vec{i} : \vec{1} \le \vec{i} < 2^{\vec{l}},\ \text{all } i_j \text{ odd}\}$$

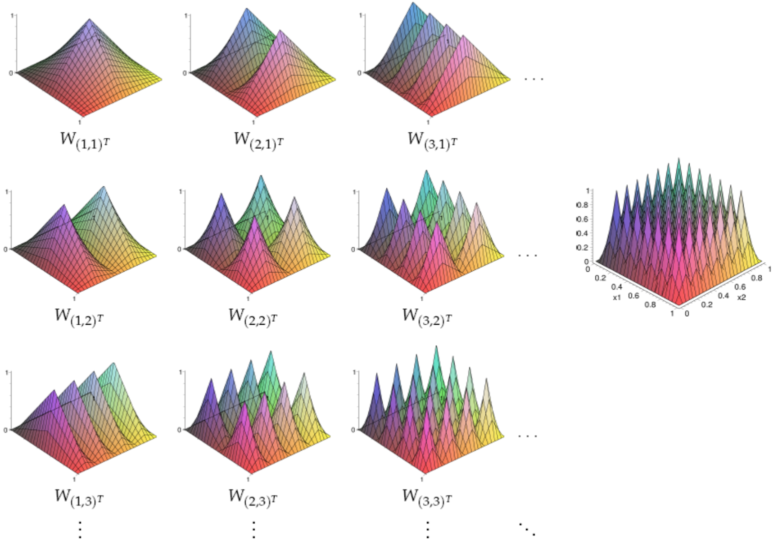
$\Rightarrow$ Hierarchical increment spaces

$$W_{\vec{l}} := \text{span}\{\phi_{\vec{l},\vec{i}}\}_{\vec{i} \in \mathcal{I}_{\vec{l}}}$$

contain all functions of $V_{\vec{l}}$ that vanish at all grid points of all coarser grids



Hierarchical Increments $W_{\vec{l}}$ vs. Nodal Basis



Hierarchical Subspace Decomposition

- For $\vec{l}' \in \mathbb{N}^d$ we obtain a unique representation of each $u \in V_{\vec{l}'}$ as

$$u = \sum_{\vec{l} \leq \vec{l}'} w_{\vec{l}}$$

with $w_{\vec{l}} \in W_{\vec{l}}$

⇒ Representation in the *hierarchical basis*

$$u = \sum_{\vec{l} \leq \vec{l}'} w_{\vec{l}} = \sum_{\vec{l} \leq \vec{l}'} \sum_{\vec{i} \in \mathcal{I}_{\vec{l}}} v_{\vec{l},\vec{i}} \, \phi_{\vec{l},\vec{i}}$$

with *d-dimensional hierarchical surpluses* $v_{\vec{l},\vec{i}}$

Determining the Hierarchical Surpluses in 2D

We now compute the hierarchical surpluses $v_{\vec{l},\vec{i}}$ for some $u \in V_n = V_{(n,n)}$:

$$u(\vec{x}) = \sum_{\phi_{\vec{l},\vec{i}} \in \Phi_{(n,n)}} u(x_{\vec{l},\vec{i}}) \cdot \phi_{\vec{l},\vec{i}}(\vec{x}) = \sum_{i_1=1}^{2^n-1} \sum_{i_2=1}^{2^n-1} u(x_{\vec{n},\vec{i}}) \cdot \phi_{n,i_1}(x_1)\phi_{n,i_2}(x_2)$$

First step

- Hierarchization in x_1-direction
  (fix x_2 and employ $1d$ hierarchization in x_1-direction):

$$u(\vec{x}) = \sum_{l_1=1}^{n} \sum_{i_1 \in \mathcal{I}_{l_1}} \sum_{i_2=1}^{2^n-1} v_{l_1,i_1}(x_{n,i_2}) \cdot \phi_{l_1,i_1}(x_1) \cdot \phi_{n,i_2}(x_2)$$

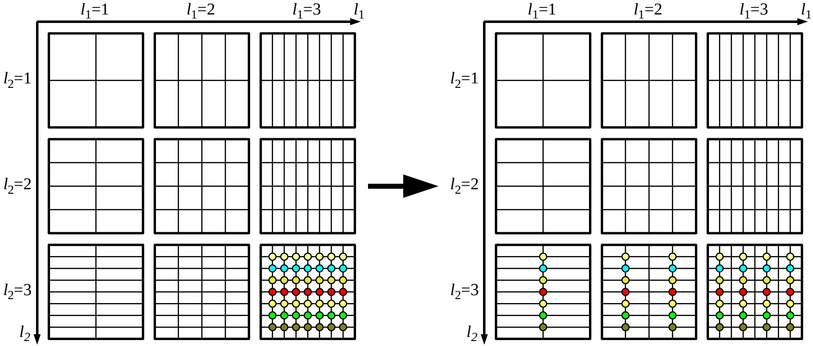
with $1d$ surplus (still depending on x_2, evaluated at all $x_2 = x_{n,i_2}$)

$$v_{l_1,i_1}(x_2) = u(x_{l_1,i_1}, x_2) - \frac{u(x_{l_1,i_1-1}, x_2) + u(x_{l_1,i_1+1}, x_2)}{2}$$

Note: the indices $i \pm 1$ of the grid points x_{l_1,i_1-1} and x_{l_1,i_1+1} are even, such that the corresponding hierarchical basis functions belong to a parent/ancestor level.

Determining the Hierarchical Surpluses in 2D (2)

A bit more intuitive:
We mark the grid points of the corresponding ansatz functions we use (before and after)

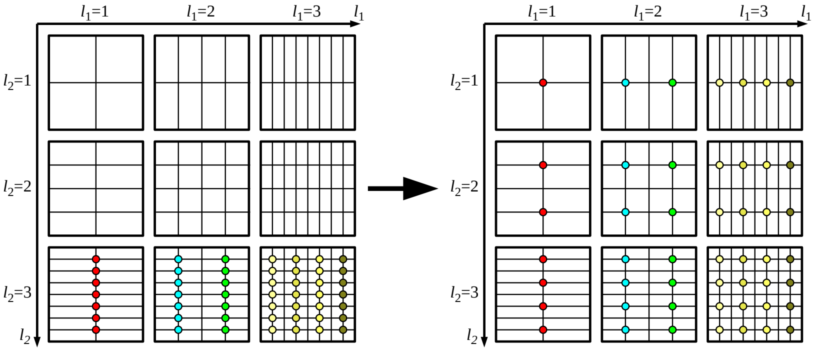


Determining the Hierarchical Surpluses in 2D (3)

Second step

- Hierarchize every $v_{l_1,i_1}(x_2)$ (separately) in x_2 dimension::

$$u(\vec{x}) = \sum_{l_1=1}^{n} \sum_{i_1 \in \mathcal{I}_{l_1}} \sum_{l_2=1}^{n} \sum_{i_2 \in \mathcal{I}_{l_2}} v_{(l_1,l_2),(i_1,i_2)} \cdot \phi_{l_1,i_1}(x_1) \cdot \phi_{l_2,i_2}(x_2)$$



Determining the Hierarchical Surpluses

(the general d-dimensional case)

Now: compute the d-dim. hierarchical surpluses $v_{\vec{l},\vec{i}}$ for some $u(\vec{x}) \in V_n$:

$$u(\vec{x}) = \sum_{\phi_{\vec{l},\vec{i}} \in \Phi_{(n,\ldots,n)}} u(x_{\vec{l},\vec{i}}) \cdot \phi_{\vec{l},\vec{i}}(\vec{x}) = \sum_{\phi_{\vec{l},\vec{i}} \in \Phi_{(n,\ldots,n)}} u(x_{\vec{l},\vec{i}}) \cdot \phi_{l_1,i_1}(x_1) \cdot \ldots \cdot \phi_{l_d,i_d}(x_d)$$

First step

- Hierarchization in x_d-direction
  (fix $x_1, \ldots, x_{d-1}$ and employ $1d$ hierarchization):

$$u = \sum_{l_d=1}^{n} \sum_{i_d \in \mathcal{I}_{l_d}} \sum_{\phi_{\vec{l},\vec{i}} \in \Phi_{(n,\ldots,n)}} v_{l_d,i_d}(x_{\vec{n},(i_1,\ldots,i_{d-1})}) \cdot \phi_{l_d,i_d}(x_d) \cdot \phi_{l_1,i_1}(x_1) \cdot \ldots \cdot \phi_{l_{d-1},i_{d-1}}(x_{d-1})$$

with $1d$ surplus – evaluted at $(x_1, \ldots, x_{d-1}) = x_{\vec{n},(i_1,\ldots,i_{d-1})}$:

$$v_{l_d,i_d}(x_1, \ldots, x_{d-1}) = u(x_1, \ldots, x_{d-1}, x_{l_d,i_d}) - \frac{u(x_1, \ldots, x_{d-1}, x_{l_d,i_d-1}) + u(x_1, \ldots, x_{d-1}, x_{l_d,i_d+1})}{2}$$

Determining the Hierarchical Surpluses

(the general *d*-dimensional case, second step)

Second step

- Hierarchize every $v_{l_d,i_d} : \mathbb{R}^{d-1} \to \mathbb{R}$ (separately) in its first argument:

$$u(\vec{x}) = \sum_{l_d=1}^{n} \sum_{i_d \in \mathcal{I}_{l_d}} \sum_{l_{d-1}=1}^{n} \sum_{i_{d-1} \in \mathcal{I}_{l_{d-1}}} \sum_{\phi_{\vec{l},\vec{i}} \in \Phi_{(n,\ldots,n)}} \Bigg(v_{l_d,i_d}(x_{\vec{n},(i_1,\ldots,i_{d-2})}) \cdot \phi_{l_d,i_d}(x_d) \cdot \phi_{l_{d-1},i_{d-1}}(x_{d-1}) \cdot \phi_{l_1,i_1}(x_1) \cdot \ldots \cdot \phi_{l_{d-2},i_{d-2}}(x_{d-2}) \Bigg)$$

Steps 3 to *d*

- All steps correspondingly for each remaining dimension
- Afterwards we have computed surpluses $v_{\vec{l},\vec{i}}$ (functions in zero parameters / scalar values)

$$u(\vec{x}) = \sum_{\vec{l}} \sum_{\vec{i} \in \mathcal{I}_{\vec{l}}} v_{\vec{l},\vec{i}} \cdot \phi_{l_1,i_1}(x_1) \cdot \ldots \cdot \phi_{l_d,i_d}(x_d) = \sum_{\vec{l}} \sum_{\vec{i} \in \mathcal{I}_{\vec{l}}} v_{\vec{l},\vec{i}} \phi_{\vec{l},\vec{i}}(\vec{x}) = \sum_{\vec{l}} w_{\vec{l}}(\vec{x})$$

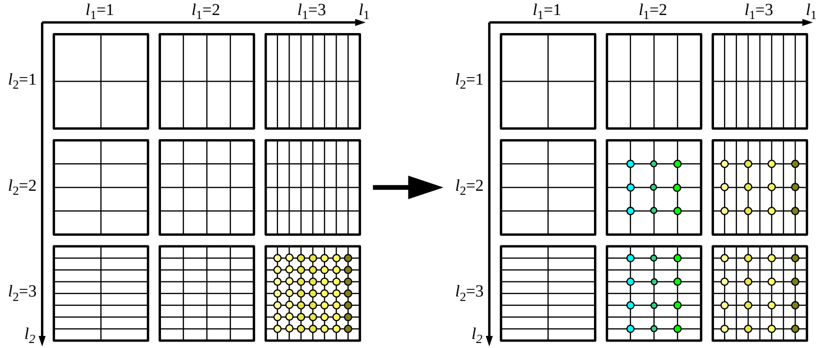
where $\sum_{\vec{l}} \sum_{\vec{i} \in \mathcal{I}_{\vec{l}}}$ is short for $\sum_{l_d=1}^{n} \sum_{i_d \in \mathcal{I}_{l_d}} \cdots \sum_{l_1=1}^{n} \sum_{i_1 \in \mathcal{I}_{l_1}}$

Comparison 2D with Wavelet Transform

Apply Level-Wise Hierarchisation

First level:

- First step: split into nodal basis and hierarchical surpluses (1st argument)
- Second step: split into nodal basis and hierarchical surpluses (2nd argument)

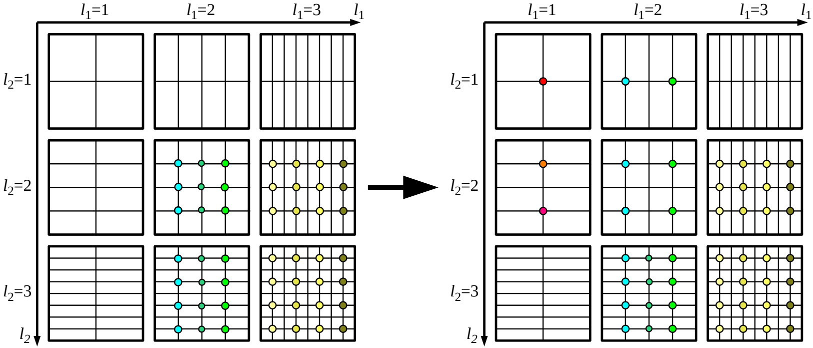


Comparison 2D with Wavelet Transform (2)

Apply Level-Wise Hierarchisation

Second level:

- First step: split into nodal basis and hierarchical surpluses (1st argument)
- Second step: split into nodal basis and hierarchical surpluses (2nd argument)



Part II

Hierarchical Decomposition – Outlook on Cost and Accuracy

Analysis of Hierarchical Decomposition

- Contribution of summands in hierarchical decomposition
  $\rightarrow$ in 1D:

$$u = \sum_{l=1}^{n} w_l = \sum_{l=1}^{n} \sum_{i \in \mathcal{I}_l} v_{l,i} \phi_{l,i}$$

  $\rightarrow$ in dD:

$$u = \sum_{\vec{l}} w_{\vec{l}} = \sum_{\vec{l}} \sum_{\vec{i} \in \mathcal{I}_{\vec{l}}} v_{\vec{l},\vec{i}} \phi_{\vec{l},\vec{i}}(\vec{x})$$

- start analysis in univariate setting
- and port to mulitvariate setting
  $\rightarrow$ Cost/benefit analysis quantifies reduction of effort
- Need several norms to measure w_l

Norms of Functions

As always, we assume sufficiently smooth functions $u : [0,1] \to \mathbb{R}$, then:

- Maximum norm

$$\|u\|_\infty := \max_{x\in[0,1]} |u(x)|$$

- L^2 norm

$$\|u\|_2 := \sqrt{\int_0^1 u(x)^2\,dx},$$

  for the L^2 scalar product

$$(u,v)_2 := \int_0^1 u(x)v(x)\,dx$$

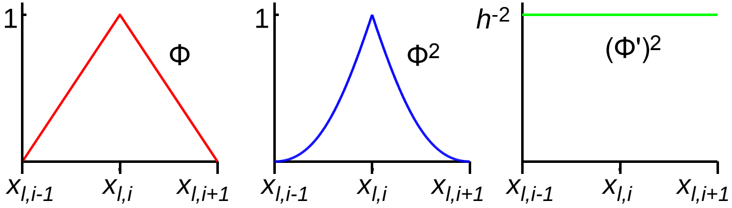
- Energy norm

$$\|u\|_E := \|u'\|_2$$

Norms of Basis Functions

For the basis functions $\phi_{l,i}$, we obtain

$$
\begin{aligned}
\|\phi_{l,i}\|_\infty &= 1 \\
\|\phi_{l,i}\|_2 &= \sqrt{\frac{2h_l}{3}} \\
\|\phi_{l,i}\|_E &= \sqrt{\frac{2}{h_l}}
\end{aligned}
$$



Estimation of Surpluses

- Consider surplus $v_{l,i}$ of basis function $\phi_{l,i}$:

$$v_{l,i} := u(x_{l,i}) - \tfrac{1}{2}\left(u(x_{l,i-1}) + u(x_{l,i+1})\right)$$

- u two times differentiable

$\Rightarrow$ We can then write $v_{l,i}$ as (see separate proof)

$$v_{l,i} = \int_0^1 \psi_{l,i}(x) u''(x)\, dx \qquad \text{with} \quad \psi_{l,i} := -\frac{h_l}{2}\phi_{l,i}$$

- $v_{l,i}$ depends on u'', thus we define for future use

$$\mu_2(u) := \|u''\|_2 \quad \textit{and} \quad \mu_\infty(u) := \|u''\|_\infty.$$

$\rightarrow$ note: $\mu_2(u)$ and $\mu_\infty(u)$ are properties of the function u

Estimation of Surplusses (2)

- With integral representation $v_{l,i} = \int_0^1 -\frac{h_l}{2}\phi_{l,i}(x)u''(x)\,dx$, we can bound

$$|v_{l,i}| \leq \frac{h_l}{2} \cdot \left(\int_0^1 \phi_{l,i}\,dx \right) \cdot \mu_\infty(u) = \frac{h_l^2}{2} \cdot \mu_\infty(u) \in \mathcal{O}(h_l^2)$$

- and, via Cauchy-Schwartz inequality $|(u,v)| \leq \|u\| \cdot \|v\|$,

$$|v_{l,i}| \leq \frac{h_l}{2} \|\phi_{l,i}\|_2 \cdot \mu_2(u|_{T_i}) = \sqrt{\frac{h_l^3}{6}} \cdot \mu_2(u|_{T_i}),$$

where $u|_{T_i}$ restricts u to the support $T_i = [x_{l,i-1}, x_{l,i+1}]$ of $\phi_{l,i}$

Estimation of w_l

- Estimate contribution of entire level l in hierarchical decomposition of u, i.e.

$$w_l = \sum_{i \in \mathcal{I}_l} v_{l,i} \phi_{l,i}$$

- Use that supports of $\phi_{l,i}$ are pairwise disjoint
- Maximum norm

$$\|w_l\|_\infty \leq \frac{h_l^2}{2} \cdot \mu_\infty(u) \in \mathcal{O}(h_l^2)$$

- L^2 norm

$$\|w_l\|_2^2 = \sum_{i \in \mathcal{I}_l} |v_{l,i}|^2 \cdot \|\phi_{l,i}\|_2^2 \leq \frac{h_l^3}{6} \cdot \frac{2h_l}{3} \cdot \sum_{i \in \mathcal{I}_l} \mu_2(u|_{T_i})^2 = \frac{h_l^4}{9} \mu_2(u)^2$$

$\Rightarrow \|w_l\|_2 \in \mathcal{O}(h_l^2)$

Estimation of w_l (2)

- Energy norm

$$\begin{aligned} \|w_l\|_E^2 &= \sum_{i \in \mathcal{I}_l} |v_{l,i}|^2 \cdot \|\phi_{l,i}\|_E^2 = \sum_{i \in \mathcal{I}_l} |v_{l,i}|^2 \frac{2}{h_l} \\ &\leq \frac{2}{h_l} \cdot \frac{h_l^4}{4} \cdot \frac{1}{2h_l} \mu_\infty(u)^2 = \frac{h_l^2}{4} \mu_\infty(u)^2 \end{aligned}$$

($2^{l-1} = 1/(2h_l)$ summands)

$\Rightarrow \|w_l\|_E \in \mathcal{O}(h_l)$

Estimation of w_l (3)

- We can write u (twice differentiable) as infinite series

$$u = \sum_{l=1}^{\infty} w_l$$

- Convergent in all three norms
- Approximation error given as

$$u - u_n := u - \sum_{l=1}^{n} w_l = \sum_{l=n+1}^{\infty} w_l$$

$\Rightarrow$ in maximum and L^2 norm: $\mathcal{O}(h_n^2)$
$\Rightarrow$ in energy norm: $\mathcal{O}(h_n)$

Towards d Dimensions: Norms of $\phi_{\vec{l},\vec{i}}$

- Estimating the $w_{\vec{l}}$ will enable us to select those subspaces that contribute most to overall solution (best cost-benefit ratios)
- Same procedure as for $d = 1$,
  but slightly more complicated functions

Start with norms

- Maximum norm:

$$\|\phi_{\vec{l},\vec{i}}\|_\infty := \max_{\vec{x}\in[0,1]^d} |\phi_{\vec{l},\vec{i}}(\vec{x})| = 1$$

- L^2 norm:

$$\|\phi_{\vec{l},\vec{i}}\|_2 := \sqrt{\int_{[0,1]^d} \phi_{\vec{l},\vec{i}}(\vec{x})^2 \, d\vec{x}} = \prod_{j=1}^{d} \|\phi_{l_j,i_j}\|_2 = \sqrt{\left(\frac{2}{3}\right)^d \prod_{j=1}^{d} h_{l_j}} = \sqrt{\left(\frac{2}{3}\right)^d 2^{-|\vec{l}|_1}}$$

Norms of $\phi_{\vec{l},\vec{i}}$ (2)

- Energy norm
  (defined as L^2 norm of the Euclidean norm of the gradient $\nabla\phi_{\vec{l},\vec{i}}$):

$$
\begin{aligned}
\|\phi_{\vec{l},\vec{i}}\|_E &:= \sqrt{\int_{[0,1]^d} \nabla\phi_{\vec{l},\vec{i}}(\vec{x}) \cdot \nabla\phi_{\vec{l},\vec{i}}(\vec{x})\, d\vec{x}} = \ldots = \\
&= \sqrt{2\left(\frac{2}{3}\right)^{d-1} \sum_{j=1}^{d} \frac{h_1 \cdot \ldots \cdot h_d}{h_j^2}} \qquad \text{(here always: } h_j := h_{l_j}\text{)} \\
&= \sqrt{2\left(\frac{2}{3}\right)^{d-1} 2^{-|\vec{l}|_1} \sum_{j=1}^{d} 2^{2l_j}}
\end{aligned}
$$

- For the two-dimensional settings ($d = 2$), we obtain

$$
\|\phi_{\vec{l},\vec{i}}\|_E = \sqrt{\frac{4}{3}\left(\frac{h_1}{h_2} + \frac{h_2}{h_1}\right)}
$$

Estimation of Surpluses

- Hierarchical surpluses now depend on mixed 2nd derivatives

$$\partial^{2d} u := \frac{\partial^{2d} u}{\partial x_1^2 \cdot \ldots \cdot \partial x_d^2}$$

- If we define

$$\psi_{\vec{l},\vec{i}} := \prod_{j=1}^{d} \psi_{l_j,i_j} = \left(\prod_{j=1}^{d} \frac{-h_j}{2} \right) \phi_{\vec{l},\vec{i}} = (-1)^d 2^{-|\vec{l}|_1 - d} \phi_{\vec{l},\vec{i}}$$

we can derive an integral representation similar to 1D:

$$v_{\vec{l},\vec{i}} = \int_{[0,1]^d} \psi_{\vec{l},\vec{i}} \cdot \partial^{2d} u \, d\vec{x}$$

(Proof: Fubini's theorem and 1d integral representation)

Estimation of Surpluses (2)

- We define (correspondingly to $1d$)

$$\mu_2(u) := \|\partial^{2d} u\|_2 \quad \text{and} \quad \mu_\infty(u) := \|\partial^{2d} u\|_\infty$$

- We can thus bound $v_{\vec{l},\vec{i}}$ as

$$|v_{\vec{l},\vec{i}}| \leq \left(\prod_{j=1}^{d} \frac{h_j}{2}\right) \cdot \left(\int_{[0,1]^d} \phi_{\vec{l},\vec{i}} \, d\vec{x}\right) \cdot \mu_\infty(u) = \left(\prod_{j=1}^{d} \frac{h_j^2}{2}\right) \cdot \mu_\infty(u) = 2^{-2|\vec{l}|_1 - d} \mu_\infty(u)$$

and

$$\begin{aligned} |v_{\vec{l},\vec{i}}| &\leq \left(\prod_{j=1}^{d} \frac{h_j}{2}\right) \|\phi_{\vec{l},\vec{i}}\|_2 \cdot \mu_2(u|_{T_{\vec{i}}}) = \sqrt{\frac{h_1^3 \cdot \ldots \cdot h_d^3}{6^d}} \cdot \mu_2(u|_{T_{\vec{i}}}) \\ &= \left(\tfrac{1}{6}\right)^{d/2} 2^{-3|\vec{l}|_1/2} \mu_2(u|_{T_{\vec{i}}}) \end{aligned}$$

Estimation of $w_{\vec{l}}$

- Obtain estimates for $w_{\vec{l}}$ in subspace $W_{\vec{l}}$ analogously as in $1d$:
  $\rightarrow$ Make use of the fact that supports of basis functions for a grid are disjoint (apart from the boundaries)
- Maximum norm

$$\|w_{\vec{l}}\|_\infty \leq \left(\prod_{j=1}^{d} \frac{h_j^2}{2}\right) \cdot \mu_\infty(u) = 2^{-2|\vec{l}|_1 - d} \mu_\infty(u),$$

- L^2 norm

$$\|w_{\vec{l}}\|_2 \leq \left(\prod_{j=1}^{d} \frac{h_j^2}{3}\right) \cdot \mu_2(u) = 3^{-d} \cdot 2^{-2|\vec{l}|_1} \mu_2(u),$$

- Energy norm

$$\|w_{\vec{l}}\|_E \leq \sqrt{\frac{1}{4}\left(\frac{1}{12}\right)^{d-1} \sum_{j=1}^{d} \frac{h_1^4 \cdot \ldots \cdot h_d^4}{h_j^2}} \cdot \mu_\infty(u) = \sqrt{\frac{1}{4}\left(\frac{1}{12}\right)^{d-1} 2^{-4|\vec{l}|_1} \sum_{j=1}^{d} 2^{2l_j}} \cdot \mu_\infty(u)$$

Analysis of Cost-Benefit Ratio

- Consider not individual basis functions, but whole hierarchical increments
- From the tableau of subspaces, select those subspaces that minimize the cost, or maximize the benefit respectively,
  for $u : [0,1]^d \to \mathbb{R}$ (u sufficiently often differentiable)

Cost

- Measure cost in number of grid points ("coefficients")

$$c(\vec{l}) = |\mathcal{I}_{\vec{l}}| = 2^{|\vec{l}|_1 - d}$$

Benefit

- How to measure benefit? $\rightsquigarrow$ interpolation error
- Let $L \subset \mathbb{N}^d$ be the set of indices of all selected grids, then

$$u_L := \sum_{\vec{l} \in L} w_{\vec{l}} \quad \text{and} \quad u - u_L = \sum_{\vec{l} \notin L} w_{\vec{l}}$$

Analysis of Cost-Benefit Ratio (2)

- For each component $w_{\vec{l}}$, we have derived bounds of the type

$$\|w_{\vec{l}}\| \leq s(\vec{l}) \cdot \mu(u)$$

with $s(\vec{l}) = 2^{-d} \cdot 2^{-2|\vec{l}|_1}$ or $s(\vec{l}) = 3^{-d} \cdot 2^{-2|\vec{l}|_1}$ and appropriate indices for norm and μ

- We obtain

$$\begin{aligned} \|u - u_L\| &\leq \sum_{\vec{l} \notin L} \|w_{\vec{l}}\| \leq \left(\sum_{\vec{l} \notin L} s(\vec{l}) \right) \mu(u) \\ &= \left[\left(\sum_{\vec{l} \in \mathbb{N}^d} s(\vec{l}) \right) - \left(\sum_{\vec{l} \in L} s(\vec{l}) \right) \right] \mu(u) \end{aligned}$$

- 1st factor depends only on selected subspaces, 2nd factor only on u
- Justifies to interpret $s(\vec{l})$ as benefit/contribution of subspace $W_{\vec{l}}$

Quality of Approximation of Full Grid V_n

Examine cost $c(\vec{l})$ and benefit $s(\vec{l})$ for full grid

- Regular grid with mesh-width $h = 2^{-n}$ in each direction (*full grid*) for function space $V_n \rightsquigarrow$ total cost $\dim V_n \in \mathcal{O}(2^{dn})$, or $\dim V_n \in \mathcal{O}(h^{-d})$
- Considered subset of hierarchical increments:

$$L_n := \{\vec{l} : |\vec{l}|_\infty \leq n\}.$$

- Bounds in L^2 and maximum norm involve factor

$$s(\vec{l}) = C \cdot 2^{-2|\vec{l}|_1}$$

$\rightarrow$ In the following estimation, leave out $\vec{l}$-independent factor C
$\rightsquigarrow$ can be appended to the estimate in the end

Quality of Approximation of Full Grid V_n (2)

- We can estimate

$$\sum_{\vec{l}\in L_n} s(\vec{l}) = \sum_{\vec{l}\in L_n} 2^{-2|\vec{l}|_1} = \sum_{l_1=1}^{n} \cdots \sum_{l_d=1}^{n} 2^{-2(l_1+\cdots+l_n)} = \left(\sum_{k=1}^{n} 2^{-2k}\right)^d$$

$$= \left(\frac{1}{4}\cdot\frac{1-\frac{1}{4}^n}{1-\frac{1}{4}}\right)^d = \left(\frac{1}{3}\right)^d\left(1-2^{-2n}\right)^d \geq \left(\frac{1}{3}\right)^d\left(1-d\cdot 2^{-2n}\right)$$

using $(1-\epsilon)^d \geq 1-d\epsilon$ for $0\leq\epsilon\leq 1$ and $d\in\mathbb{N}$

$\Rightarrow$ For $n\to\infty$ we obtain

$$\sum_{\vec{l}\in\mathbb{N}^d} s(\vec{l}) = \left(\frac{1}{3}\right)^d \qquad \text{and thus} \qquad \sum_{\vec{l}\notin L_n} s(\vec{l}) = \left(\frac{1}{3}\right)^d\cdot d\cdot 2^{-2n}$$

- Leads to bounds for the approximation error in L^2- and maximum norm

$$\|u-u_{L_n}\| \leq C\cdot\sum_{\vec{l}\notin L_n} s(\vec{l}) \leq \frac{C\cdot d}{3^d}2^{-2n} \in O(h_n^2)$$

with constant C (independent of n)

Part III

Sparse Grids

Sparse Grids

Final steps to high-dimensional numerics

- Consider sum of benefits/contributions (for L^2 and maximum norm)

$$\sum_{\vec{l}\in L_n} 2^{-2|\vec{l}|_1}$$

$\Rightarrow$ Equal benefit of hierarchical increments $W_{\vec{l}}$ for constant $|\vec{l}|_1$

- Same for cost $c(\vec{l}) = 2^{|l|_1 - d}$ (number of grid points of $W_{\vec{l}}$)

$\Rightarrow$ **Constant *cost-benefit ratio* $c(\vec{l})/s(\vec{l})$ for constant $|\vec{l}|_1$**

Full grids?

- Quadratic extract of subspaces is not economical:
  We take large subgrids with low contribution
- We could have taken others with much higher contribution

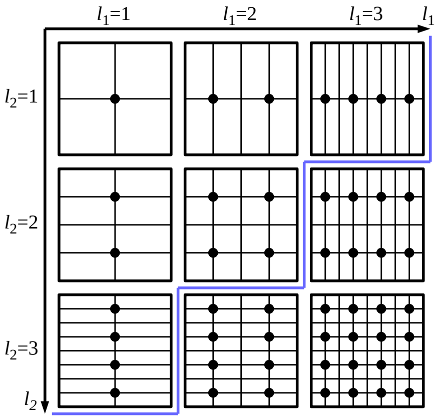
Sparse Grids!

- cost-benefit analysis:
  equal contribution of hierarchical increments $W_{\vec{l}}$ for constant $|\vec{l}|_1$
- Best choice: Cut diagonally in tableau of subspaces:

$$L_n^1 := \{\vec{l} : |\vec{l}|_1 \leq n + d - 1\}$$

$\Rightarrow$ Resulting *sparse grid space*

$$V_n^1 := \bigoplus_{|\vec{l}|_1 \leq n+d-1} W_{\vec{l}}$$



Sparse Grids!

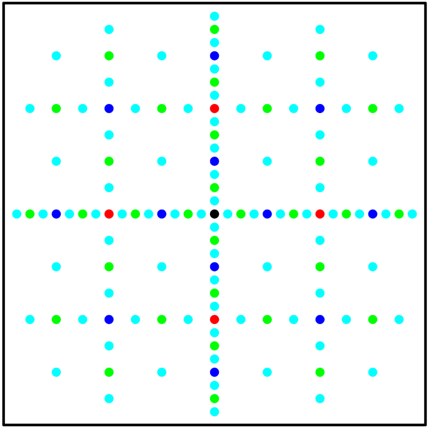
- Diagonall cut in tableau of subspaces:

$$L_n^1 := \{\vec{l} : |\vec{l}|_1 \leq n + d - 1\}$$

$\Rightarrow$ Resulting *sparse grid space*

$$V_n^1 := \bigoplus_{|\vec{l}|_1 \leq n+d-1} W_{\vec{l}}$$

- Sparse grid for $d = 2$ and overall level $n = 5$
- Grid points $x_{\vec{l},\vec{i}}$ of same cost/benefit ratio in same color



Sparse Grids – Cost

Number of grid points?

- For $d = 2$:

$$\dim V_n^1 = \sum_{|\vec{l}|_1 \le n+1} \dim W_{\vec{l}} = \sum_{|\vec{l}|_1 \le n+1} 2^{|\vec{l}|_1 - 2} = \sum_{k=1}^{n} k \cdot 2^{k-1} = 2^n(n-1) + 1,$$

- For $d = 3$:

$$\dim V_n^1 = \sum_{k=1}^{n} \frac{k(k+1)}{2} \cdot 2^{k-1} = 2^n \left(\frac{n^2}{2} - \frac{n}{2} + 1 \right) - 1,$$

$\Rightarrow$ **Both in** $\mathcal{O}(2^n \cdot n^{d-1})$

- Holds for general d as well (proof with some combinatorics)
- Expressed in terms of $N = 2^n$ (max. points per dimension):
  $\Rightarrow \mathcal{O}\left(N (\log N)^{d-1}\right)$

Sparse Grids – Cost (2)

In numbers…

Compare cost for full grid V_n and sparse grid V_n^1:

$d = 2$:

n	1	2	3	4	5	…	10
$\dim V_n = (2^n - 1)^2$	1	9	49	225	961	…	1,046,529
$\dim V_n^1 = 2^n(n-1)+1$	1	5	17	49	129	…	9,217

Even more distinct for $d = 3$:

n	1	2	3	4	…	10
$\dim V_n = (2^n - 1)^3$	1	27	343	3,375	…	1,070,590,167
$\dim V_n^1 = 2^n\left(\frac{n^2}{2} - \frac{n}{2} + 1\right) - 1$	1	7	31	111	…	47,103

Sparse Grids – Cost (3)

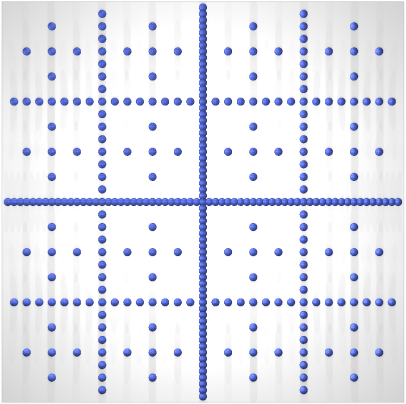
…and for overall level $n = 5$ in different dimensions

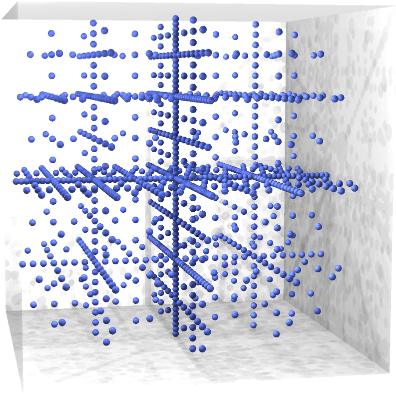
d	V_5	V_5^1
1	31	31
2	961	129
3	29,791	351
4	923,521	769
5	28,629,151	1,471
6	887,503,681	2,561
7	27,512,614,111	4,159
8	852,891,037,441	6,401
9	26,439,622,160,671	9,439
10	819,628,286,980,801	13,441

- The higher the dimension, the higher the benefit of sparse grids!

Sparse Grids – Examples

Sparse Grids of overall level $n = 6$ in $d = 2$ and $d = 3$





Sparse Grids – Accuracy

Much fewer grid points $\Rightarrow$ much lower accuracy?

- Would force us to choose larger n to obtain similar accuracy (and spoil everything)
- Error in L^2 and maximum norm:
  Compute sum ($|\vec{l}|_1 = k+1$):

$$\sum_{\vec{l} \notin L_n^1} s(\vec{l}) = \sum_{k=n+1}^{\infty} k \cdot 2^{-2(k+1)} = \left(\frac{n}{12} + \frac{1}{9}\right) 2^{-2n}$$

- And for $d = 3$ (with $|\vec{l}|_1 = k+2$):

$$\sum_{\vec{l} \notin L_n^1} s(\vec{l}) = \sum_{k=n+3}^{\infty} \frac{k(k+1)}{2} \cdot 2^{-2(k+2)} = \left(\frac{n^2}{96} + \frac{11n}{288} + \frac{1}{27}\right) 2^{-2n}$$

Sparse Grids – Accuracy (2)

In general, it can be shown

- Error of interpolation in L^2 and maximum norm is $\mathcal{O}(2^{-2n} n^{d-1})$
  $\Rightarrow$ or, expressed in mesh size $h := 2^{-n}$: $\mathcal{O}\left(h^2 \left(\log \frac{1}{h}\right)^{d-1}\right)$
- Only polynomial (in n) factor worse than full grid with $\mathcal{O}(2^{-2n})$
  $\Rightarrow$ or, expressed in mesh size $h := 2^{-n}$: $\mathcal{O}(h^2)$

Outlook on Energy norm: ($\rightsquigarrow$ Algorithms for Scientific Computing II)

- Analysis is more complicated (lines through subspaces with similar $s(\vec{l})$, and thus $c(\vec{l})/s(\vec{l})$, are more complicated)
- Overall result even better:
  obtain accuracy of $\mathcal{O}(2^{-n})$ with only $\mathcal{O}(2^n)$ grid points
  $\rightarrow$ no polynomial terms (of type n^d) left!