Algorithms for Scientific Computing

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

Michael Bader
Technical University of Munich
Summer 2016
Intermezzo/“Big Picture”: Archimedes’ Quadrature

- Start with 2d example (compare tutorials):
  \[
  f := 16x_1(x_1 - 1)x_2(x_2 - 1), \quad \Omega = [0, 1]^2 \quad \Rightarrow f|_{\partial \Omega} = 0
  \]
- Consider hierarchical surplus at grid points with \( n = 3, h_3 = 2^{-3} \)

<table>
<thead>
<tr>
<th></th>
<th>( \frac{1}{256} )</th>
<th>( \frac{1}{64} )</th>
<th>( \frac{1}{256} )</th>
<th>( \frac{1}{16} )</th>
<th>( \frac{1}{256} )</th>
<th>( \frac{1}{64} )</th>
<th>( \frac{1}{256} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{64} )</td>
</tr>
<tr>
<td></td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{256} )</td>
</tr>
<tr>
<td></td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{16} )</td>
<td>( 1 )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{16} )</td>
</tr>
<tr>
<td></td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{256} )</td>
</tr>
<tr>
<td></td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{64} )</td>
</tr>
<tr>
<td></td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{16} )</td>
<td>( \frac{1}{256} )</td>
<td>( \frac{1}{64} )</td>
<td>( \frac{1}{256} )</td>
</tr>
</tbody>
</table>
“Big Picture”: Archimedes’ Quadrature (2)

\[ \int_{\Omega} f \, d\vec{x} = \frac{4}{9} = 0.4 \quad \sum = \frac{441}{1024} = 0.4306640625 \]

- Consider volume of subvolumes (pagodas) for quadrature
“Big Picture”: Archimedes’ Quadrature (3)

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

- 49 grid points (full grid) ⇒ 17 grid points (sparse grid)

- Approximation of volume:

\[
\frac{441}{1024} = 0.4306640625 \quad \Rightarrow \quad \frac{27}{64} = 0.421875
\]
Part I

Hierarchical Decomposition, $d$-Dimensional
Hierarchical Decomposition – Step by Step

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

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

- Functions in multiple variables $\vec{x} = (x_1, \ldots, x_d)$
- 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
  \[ \vec{l} = (l_1, \ldots, l_d) \in \mathbb{N}^d \]
- Grids have different mesh-widths in different dimensions:
  \[ \vec{h_l} := (h_1, \ldots, h_d) := (2^{-l_1}, \ldots, 2^{-l_d}) =: 2^{-\vec{l}} \]
Hierarchical Decomposition, $d > 1$

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

- 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 $|\cdot|_{\infty}$*

- Comparisons of multi-indices component-wise:
  \[ \vec{l} \leq \vec{i} \iff l_k \leq i_k, \quad k = 1, \ldots, d \]

- Grid points (for function evaluations):
  \[ \vec{x}_{\vec{l},\vec{i}} = (i_1 \cdot h_{l_1}, \ldots, i_d \cdot h_{l_d}) \]
Practicing Identifiers $\vec{l}, \vec{h}, \vec{x}_{l,i}$

<table>
<thead>
<tr>
<th>$l_1$</th>
<th>$l_2=1$</th>
<th>$l_2=2$</th>
<th>$l_2=3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l_1=1$</td>
<td><img src="image" alt="Grid" /></td>
<td><img src="image" alt="Grid" /></td>
<td><img src="image" alt="Grid" /></td>
</tr>
<tr>
<td>$l_1=2$</td>
<td><img src="image" alt="Grid" /></td>
<td><img src="image" alt="Grid" /></td>
<td><img src="image" alt="Grid" /></td>
</tr>
<tr>
<td>$l_1=3$</td>
<td><img src="image" alt="Grid" /></td>
<td><img src="image" alt="Grid" /></td>
<td><img src="image" alt="Grid" /></td>
</tr>
</tbody>
</table>
Piecewise $d$-linear Functions

Suitable generalization of piecewise linear functions

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

Alternative point of view:

- Define suitable basis $\Phi_{\vec{i}}$
- Regard $V_{\vec{i}}$ as span of $\Phi_{\vec{i}}$
- $d$-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 \textit{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
  \[ \Phi_{\vec{l}} := \{ \phi_{\vec{l}, i}, 1 \leq i < 2^{|\vec{l}|} \} \]

- Function space
  \[ V_{\vec{l}} := \text{span}\{ \Phi_{\vec{l}} \} \]
  with
  \[ \dim V_{\vec{l}} = (2^{l_1} - 1) \cdot \ldots \cdot (2^{l_d} - 1) \in O(2^{|\vec{l}|}) \]

- Special case $l_1 = \ldots = l_d \Rightarrow$ 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

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

⇒ 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 $\mathcal{W}_i$ vs. Nodal Basis

$W_{(1,1)^T}$  $W_{(2,1)^T}$  $W_{(3,1)^T}$

$W_{(1,2)^T}$  $W_{(2,2)^T}$  $W_{(3,2)^T}$

$W_{(1,3)^T}$  $W_{(2,3)^T}$  $W_{(3,3)^T}$
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 $w_{\vec{l}} \in W_{\vec{l}}$

$\Rightarrow$ Representation in the *hierarchical basis*

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

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

We now compute the hierarchical surpluses $v_{l,i}$ for some $u \in V_n$:

$$u = \sum_{\phi_{l,i} \in \Phi(n,...,n)} u(x_{l,i}) \cdot \phi_{l,i}$$

First step

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

$$u = \sum_{l_1=1}^{n} \sum_{i_1 \in \mathcal{I}_{l_1}} v_{l_1,i_1}(x_2, \ldots, x_d) \phi_{l_1,i_1}(x_1)$$

with $1d$ surplus

$$v_{l_1,i_1}(x_2, \ldots, x_d) = \frac{u(x_{l_1,i_1}, x_2, \ldots, x_d) - u(x_{l_1,i_1-1}, x_2, \ldots, x_d) + u(x_{l_1,i_1+1}, x_2, \ldots, x_d)}{2}$$
Determining the Hierarchical Surpluses (2)

A bit more intuitive:
We mark the grid points of the corresponding ansatz functions we use (before and after)
Determining the Hierarchical Surpluses (3)

Second step

- Hierarchize every \( v_{l_1,i_1} : \mathbb{R}^{d-1} \rightarrow \mathbb{R} \) (separately) in its first argument:

\[
U = \sum_{l_1=1}^{n} \sum_{i_1 \in I_{l_1}} \sum_{l_2=1}^{n} \sum_{i_2 \in I_{l_2}} v(l_1,l_2),(i_1,i_2)(x_3, \ldots, x_d) \phi_{l_1,i_1}(x_1) \phi_{l_2,i_2}(x_2)
\]
Determining the Hierarchical Surpluses (3)

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)
- Representation

$$u = \sum_{\vec{l}} \sum_{\vec{i} \in I_{\vec{l}}} v_{\vec{l},\vec{i}} \phi_{l_1,i_1}(x_1) \phi_{l_2,i_2}(x_2) \cdots \phi_{l_d,i_d}(x_d)$$

$$= \sum_{\vec{l}} \sum_{\vec{i} \in I_{\vec{l}}} v_{\vec{l},\vec{i}} \phi_{\vec{l},\vec{i}}(\vec{x})$$

$$= \sum_{\vec{l}} w_{\vec{l}}.$$
Part II

Hierarchical Decomposition – Outlook on Cost and Accuracy
Analysis of Hierarchical Decomposition

• Contribution of summands in hierarchical decomposition
  → in 1D:

\[ u = \sum_{l=1}^{n} w_l = \sum_{l=1}^{n} \sum_{i \in I_l} v_{l,i} \phi_{l,i} \]

→ in dD:

\[ u = \sum_{\vec{l}} w_{\vec{l}} = \sum_{\vec{l}} \sum_{\vec{i} \in I_{\vec{l}}} v_{\vec{l},\vec{i}} \phi_{\vec{l},\vec{i}}(\vec{x}) \]

• start analysis in univariate setting
• and port to multivariate setting
  → 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{align*}
\| \phi_{l,i} \|_\infty &= 1 \\
\| \phi_{l,i} \|_2 &= \sqrt{\frac{2h_l}{3}} \\
\| \phi_{l,i} \|_E &= \sqrt{\frac{2}{h_l}}
\end{align*}
Estimation of Surpluses

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

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

• $u$ two times differentiable

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

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

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

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

$\rightarrow$ 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

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

- 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
  \[ \|w_l\|_\infty \leq \frac{h_l^2}{2} \cdot \mu_\infty(u) \in O(h_l^2) \]
- $L^2$ norm
  \[ \|w_l\|_2^2 = \sum_{i \in \mathcal{I}_l} |v_{l,i}|^2 \cdot \|\phi_{l,i}\|_2^2 \leq \frac{h_l^3}{6} \cdot \frac{2h_l}{3} \cdot \sum_{i \in \mathcal{I}_l} \mu_2(u|_{T_i})^2 = \frac{h_l^4}{9} \mu_2(u)^2 \]
  \[ \Rightarrow \|w_l\|_2 \in O(h_l^2) \]
Estimation of $w_l$ (2)

- Energy norm

$$\|w_l\|_E^2 = \sum_{i \in I_l} |v_{l,i}|^2 \cdot \|\phi_{l,i}\|_E^2 = \sum_{i \in I_l} |v_{l,i}|^2 \frac{2}{h_l}$$

$$\leq \frac{2}{h_l} \cdot \frac{h_l^4}{4} \cdot \frac{1}{2h_l} \mu_\infty(u)^2 = \frac{h_l^2}{4} \mu_\infty(u)^2$$

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

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

\[
\|\phi_{\vec{l},i}\|_E := \sqrt{\int_{[0,1]^d} \nabla \phi_{\vec{l},i}(\vec{x}) \cdot \nabla \phi_{\vec{l},i}(\vec{x}) \, d\vec{x}} = \ldots =
\]

\[
= \sqrt{2 \left(\frac{2}{3}\right)^{d-1} \sum_{j=1}^{d} \frac{h_1 \cdot \ldots \cdot h_d}{h_j^2}}
\]

\[
= \sqrt{2 \left(\frac{2}{3}\right)^{d-1} 2^{-|\vec{l}|_1} \sum_{j=1}^{d} 2^{2l_j}}
\]

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

\[
\|\phi_{\vec{l},i}\|_E = \sqrt{\frac{4}{3} \left(\frac{h_1}{h_2} + \frac{h_2}{h_1}\right)}
\]
Estimation of Surpluses

- Hierarchical surpluses now depend on mixed 2nd derivatives

\[ \partial^{2d} u := \frac{\partial^{2d} u}{\partial x_1^2 \cdots \partial x_d^2} \]

- 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 obtain 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 $1^d$)
  \[ \mu_2(u) := \| \partial^{2d} u \|_2 \quad \text{and} \quad \mu_\infty(u) := \| \partial^{2d} u \|_\infty \]

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

and

\[
|v_{\vec{l}, \vec{i}}| \leq \left( \prod_{j=1}^{d} \frac{h_j}{2} \right) \| \phi_{\vec{l}, \vec{i}} \|_2 \cdot \mu_2(u \mid T_{\vec{i}}) = \sqrt{\frac{h_1^3 \cdot \ldots \cdot h_d^3}{6^d}} \cdot \mu_2(u \mid T_{\vec{i}})
\]

\[
= \left( \frac{1}{6} \right)^{d/2} \cdot 2^{-3|\vec{l}|/2} \cdot \mu_2(u \mid T_{\vec{i}})
\]
Estimation of $w_\vec{l}$

- Obtain estimates for $w_\vec{l}$ in subspace $W_\vec{l}$ analogously as in 1d:
  - 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
  $$\|w_\vec{l}\|_E \leq \sqrt{\frac{1}{4} \left(\frac{1}{12}\right)^{d-1} \sum_{j=1}^{d} \frac{h_1^4 \cdots h_d^4}{h_j^2}} \cdot \mu_\infty(u) = \sqrt{\frac{1}{4} \left(\frac{1}{12}\right)^{d-1} 2^{-4|\vec{l}|_1} \sum_{j=1}^{d} 2^{2l_j} \cdot \mu_\infty(u)}$$
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”)

\[
c(\vec{l}) = |I_{\vec{l}}| = 2^{||\vec{l}||_1 - d}
\]

Benefit
- How to measure benefit? \( \leadsto \) 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 $\mu$.

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

- 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 $2^{-n}$ in each direction (full grid) for function space $V_n$
- 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 $C$
$\sim\Rightarrow$ can be appended to the estimate in the end
Quality of Approximation of Full Grid $V_n$ (2)

- We can estimate

$$
\sum_{\vec{l} \in L_n} s(\vec{l}) = \sum_{\vec{l} \in L_n} 2^{-2|\vec{l}|_1} = \left( \sum_{k=1}^{n} 2^{-2k} \right)^d = \left( \frac{1}{4} \cdot \frac{1 - \frac{1}{4}^n}{1 - \frac{1}{4}} \right)^d
$$

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

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

⇒ For $n \to \infty$ we obtain

$$
\sum_{\vec{l} \in \mathbb{N}^d} s(\vec{l}) = \left( \frac{1}{3} \right)^d
$$

- 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 $n$)
Sparse Grids

Final steps to high-dimensional numerics

- Consider sum of benefits/contributions (for $L^2$ and maximum norm)
  \[ \sum_{\vec{l} \in L_n} 2^{-2|\vec{l}|_1} \]

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

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

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

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
Part III

Sparse Grids
Sparse Grids!

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

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

⇒ Resulting sparse grid space

$$V_n^1 := \bigoplus_{|\overrightarrow{i}|_1 \leq n + d - 1} W_i$$
Sparse Grids!

- Diagonally cut in tableau of subspaces:
  \[ L_n^1 := \{ \vec{i} : |\vec{i}|_1 \leq n + d - 1 \} \]
  \[ \Rightarrow \] Resulting sparse grid space
  \[ V_n^1 := \bigoplus_{|\vec{i}|_1 \leq n + d - 1} W_{\vec{i}} \]

- Sparse grid for \( d = 2 \) and overall level \( n = 5 \)
- Grid points \( x_{\vec{i},i} \) of same importance in same color
Sparse Grids – Cost

**Number of grid points?**

- For $d = 2$:
  \[
  \dim V_n^1 = \sum_{|\vec{l}|_1 \leq n+1} \dim W_{\vec{l}} = \sum_{|\vec{l}|_1 \leq n+1} 2^{|\vec{l}|_1 - 2} = \sum_{k=1}^{n} k \cdot 2^{k-1} = 2^n(n - 1) + 1,
  \]

- For $d = 3$:
  \[
  \dim V_n^1 = \sum_{k=1}^{n} \frac{k(k + 1)}{2} \cdot 2^{k-1} = 2^n \left( \frac{n^2}{2} - \frac{n}{2} + 1 \right) - 1,
  \]

⇒ **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):
  \[
  \Rightarrow \mathcal{O}(N(\log N)^{d-1})
  \]
Sparse Grids – Cost (2)

In numbers...

Compare cost for full grid $V_n$ and sparse grid $V^1_n$:

$d = 2$:

<table>
<thead>
<tr>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>…</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\dim V_n = (2^n - 1)^2$</td>
<td>1</td>
<td>9</td>
<td>49</td>
<td>225</td>
<td>961</td>
<td>…</td>
<td>1,046,529</td>
</tr>
<tr>
<td>$\dim V^1_n = 2^n(n - 1) + 1$</td>
<td>1</td>
<td>5</td>
<td>17</td>
<td>49</td>
<td>129</td>
<td>…</td>
<td>9,217</td>
</tr>
</tbody>
</table>

Even more distinct for $d = 3$:

<table>
<thead>
<tr>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>…</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\dim V_n = (2^n - 1)^3$</td>
<td>1</td>
<td>27</td>
<td>343</td>
<td>3,375</td>
<td>…</td>
<td>1,070,590,167</td>
</tr>
<tr>
<td>$\dim V^1_n = 2^n\left(\frac{n^2}{2} - \frac{n}{2} + 1\right) - 1$</td>
<td>1</td>
<td>7</td>
<td>31</td>
<td>111</td>
<td>…</td>
<td>47,103</td>
</tr>
</tbody>
</table>
Sparse Grids – Cost (3)

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

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

- 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 ⇒ 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 ($|I|_1 = k + 1$):

$$
\sum_{\vec{l} \not\in L_n^1} s(\vec{l}) = \sum_{k=n+1}^{\infty} k \cdot 2^{-2(k+1)} = \left( \frac{n}{12} + \frac{1}{9} \right) 2^{-2n}
$$

• And for $d = 3$ (with $|I|_1 = k + 2$):

$$
\sum_{\vec{l} \not\in L_n^1} s(\vec{l}) = \sum_{k=n+3}^{\infty} \frac{k(k+1)}{2} \cdot 2^{-2(k+2)} = \left( \frac{n^2}{96} + \frac{11n}{288} + \frac{1}{27} \right) 2^{-2n}
$$
Sparse Grids – Accuracy (2)

In general, it can be shown

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

For Energy norm:

- 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 $O(2^{-n})$ with only $O(2^n)$ grid points
  - no polynomial terms (of type $n^d$) left!