Centrum voor Wiskunde en Informatica
REPORT*RAPPORT*

On the Representation of Functions and Finite Difference Operators on Adaptive Sparse Grids

P.W. Hemker, F. Sprengel

Modelling, Analysis and Simulation (MAS)

MAS-R9933 November 30, 1999

Report MAS-R9933
ISSN 1386-3703

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

On the Representation of Functions and Finite Difference Operators on Adaptive Sparse Grids

P.W. Hemker

CWI

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

and

F. Sprengel

SCAI Institute for Algorithms and Scientific Computing

GMD German National Research Center for Information Technology

Schloss Birlinghoven, D–53754 Sankt Augustin, Germany

ABSTRACT

In this paper we describe methods to approximate functions and differential operators on adaptive sparse grids. We distinguish between several representations of a function on the sparse grid, and we describe how finite difference (FD) operators can be applied to these representations.

For general variable coefficient equations on sparse grids, FD operators allow a more efficient operator evaluation than finite element operators. However, the structure of the FD operators is more complex. In order to examine the possibility to construct efficient solution methods, we analyze the discrete FD (Laplace) operator and compare its hierarchical representation on sparse and on full grids. The analysis gives a motivation for a MG solution algorithm.

1991 Mathematics Subject Classification: 65D15, 65D25, 65N06, 65N55

Keywords and Phrases: adaptiv sparse grids, function representation, operator approximation, finite differences

Note: Work carried out under project MAS2.1, as part of the CWI - GMD cooperation

1. Introduction

When applied to d dimensions, $d \geq 2$, all numerical methods using regular rectangular grids have one problem in common: *the curse of dimension*. This means, that if one refines the grid — for instance by repeatedly halving the mesh size — the number of grid points grows exponentially with the dimension, i.e., like N^d in $\mathbb{R}^d$ where $N = \mathcal{O}(2^n)$ denotes the number of points per direction. One way out is to use *sparse grids*, where the number of points only grows like $N\,(\log N)^{d-1}$. Nevertheless, under certain conditions on the mixed derivatives of the function, the approximation accuracy degrades only by a logarithmic factor compared with the accuracy achieved on the regular grid with the same N, cf. Section 2.3 or [13, 23, 9].

Sparse grids and related methods already have a long tradition in numerical quadrature and approximation theory (see e.g. [14, 19, 2, 3, 22]). During the last decade, since the introduction of sparse grid methods into the numerical treatment of elliptic boundary problems by Zenger [23], several authors (e.g. [1, 5, 9, 10, 15, 16]) have contributed to this field. Most of these papers were concerned with finite element methods.

Because the efficient evaluation of general, variable coefficient 3-dimensional finite element operators appears to be an unsolved problem, recently also finite difference methods for sparse grids have been developed [6, 8, 17, 18]. Note, that consistent finite difference operators on sparse grids *cannot* be chosen as simple stencils involving the nearest neighbors of a point, cf. [17, Appendix]. In [6, 17], the whole machinery for finite differences on sparse grids (which needs linear combination, multiplication

and approximate differentiation) is provided. In Section 3, we present an alternative way of motivating and describing the finite difference discretization of first and second order derivatives on sparse grids.

In Section 4, we restrict ourselves to the special case of Poisson's equation and regular sparse grids in order to analyze the finite difference operator in detail. The resulting matrix is ill-conditioned. So, it takes (even in the preconditioned version [17]) many iterations for an iterative solver (like BiCGStab) to obtain a solution. Our purpose is to construct Galerkin relations which may motivate multilevel algorithms, and to propose such algorithms suited for solving the finite difference system of linear equations in a better way. A few results for these algorithms are shown. More results, with different possible solution procedures, will be published elsewhere.

2. Representation

2.1 Basic notation

To be able to describe adaptive sparse grid (ASG) function representation, we first summarize some necessary notation. For background see, e.g., [13].

- Domain $\Omega \subset \mathbb{R}^d$, with coordinates x_j, $j = 1, \ldots, d$.
- Multi-integer: $\mathbf{m} = (m_1, m_2, \ldots .m_d) \in \mathbb{Z}^d$,
  - $\mathbf{0} = (0, 0, \ldots, 0)$, — $\mathbf{e} = (1, 1, \ldots, 1)$,
  - $\mathbf{e}_j = (\ldots, 0, 1, 0, \ldots)$, the j-th unit vector,
  - $|\mathbf{m}| = \sum_{j=1}^{d} m_j$, — $\mathbf{m} < \mathbf{n} \quad \Leftrightarrow \quad m_j < n_j \ \forall j = 1, 2, \ldots, d$,
  - $\lfloor \mathbf{m} \rfloor = \min_{j=1,\ldots,d} m_j$, — $\lceil \mathbf{m} \rceil = \max_{j=1,\ldots,d} m_j$,
  - $\|\mathbf{m}\| = \sqrt{\sum_{j=1}^{d} m_j^2}$, — $\|\!|\mathbf{m}|\!\| = \prod_{j=1}^{d} m_j$.
- Dyadic mesh $\Omega_{\mathbf{k}}$, $\mathbf{k} \geq \mathbf{0}$, mesh with dyadic mesh-width $\mathbf{h}_{\mathbf{k}}$,
  - Mesh-width: $\mathbf{h} \in \mathbb{R}^d$; $\quad \mathbf{h} = (h_1, h_2, \ldots, h_d)$,
  - Mesh-size: $\|\mathbf{h}\| = \sqrt{\sum_{i=1}^{d} h_i^2}$,
  - Mesh-volume: $\|\!|\mathbf{h}|\!\| = \prod_{i=1}^{d} h_i$,
  - Dyadic mesh-width: $\mathbf{h}_{\mathbf{k}} = (h_{k_1}, h_{k_2}, \ldots, h_{k_d})$ with $h_{k_i} = 2^{-k_i}$.
  - Dyadic grid: $\Omega_{\mathbf{k}}^{+} = \{\mathbf{x}_{\mathbf{k},\mathbf{j}} \mid \mathbf{x}_{\mathbf{k},\mathbf{j}} = \mathbf{j} \cdot \mathbf{h}_{\mathbf{k}} = (j_1 h_{k_1}, j_2 h_{k_2}, \ldots, j_d h_{k_d})\} \cap \overline{\Omega}$,
  - Sparse grid: $\Omega_{\ell}^{+} = \bigcup_{|\mathbf{k}|=\ell} \Omega_{\mathbf{k}}^{+}$.
- Derivatives: $D^{\mathbf{m}} = \prod_{j=1}^{d} \frac{\partial^{m_j}}{\partial x_j}$.
- Univariate hat function: $\varphi(x) = \max(0, 1 - |x|)$.
- Univariate Haar function: $\eta(x) = 1$ for $0 < x < 1$, and $\eta(x) = 0$ for $x < 0$ or $x > 1$. In this way, the Haar function is defined only almost everywhere. Therefore, we define the left- and right-continuous Haar function by respectively $\eta^L(x) = \lim_{\xi \nearrow x} \eta(x)$ and $\eta^R(x) = \lim_{\xi \searrow x} \eta(x)$
- Basis hat functions: $\varphi_{\mathbf{k},\mathbf{j}}(\mathbf{x}) = \prod_{i=1}^{d} \varphi(x_i / h_{k_i} - j_i)$.
- Basis Haar functions: $\eta_{\mathbf{k},\mathbf{j}}(\mathbf{x}) = \prod_{i=1}^{d} \eta(x_i / h_{k_i} - j_i)$.
- Space of piecewise d-linear functions on $\Omega_{\mathbf{k}}$: $V_{\mathbf{k}} = \operatorname{span}\{\varphi_{\mathbf{k},\mathbf{j}} \mid \mathbf{j} \in \mathbb{Z}^d, \mathbf{x}_{\mathbf{k},\mathbf{j}} \in \Omega_{\mathbf{k}}^{+}\}$.
- Space of hierarchical surpluses on $\Omega_{\mathbf{k}}$:

$$W_{\mathbf{k}} = \operatorname{span}\{\varphi_{\mathbf{k},\mathbf{j}} \mid \|\!|\mathbf{j}|\!\| \text{ odd}, \mathbf{j} \in \mathbb{Z}^d, \mathbf{x}_{\mathbf{k},\mathbf{j}} \in \Omega_{\mathbf{k}}^{+}\}.$$

Without loss of generality, we assume here that $\mathbf{k} = \mathbf{0}$ yields "the coarsest grid".

With "$|||\mathbf{j}|||$ odd" we mean: for all $i = 1, 2, \ldots, d$, either j_i is an odd integer, or $k_i = 0$, (i.e., $\mathbf{x}_{\mathbf{k},\mathbf{j}}$ lives on the coarsest grid in the i-direction). If j_i is even and $k_i > 0$, there is a coarser grid on which the same point can be found. This implies that the phrase "$(\mathbf{k},\mathbf{j})$ such that $|||\mathbf{j}|||$ odd" means that $(\mathbf{k},\mathbf{j})$ is the index of the point $\mathbf{x}_{\mathbf{k},\mathbf{j}}$ on the coarsest grid in which this particular dyadic point appears.

- Norms for the function spaces used in Section 2.3 ($p \geq 1$, $0 \leq \ell \leq d$):
  - $\|u\|_\infty = \|u\|_{L_\infty} = \text{supess}_{\mathbf{x}\in\Omega}|u(\mathbf{x})|$,
  - $\|u\|_p = \|u\|_{L_p} = \left(\int_\Omega |u(\mathbf{x})|^p \, d\Omega\right)^{1/p}$,
  - $\|u\|_{W_p^{\mathbf{n},\ell}} = \left(\sum_{\substack{|\mathbf{m}|=\ell \\ \mathbf{0}\leq\mathbf{m}\leq\mathbf{e}}} \|D^{\mathbf{n}+\mathbf{m}}u\|_p^p \, d\Omega\right)^{1/p}$.
  - $\|u\|_{W_p^1} = \|u\|_{W_p^{\mathbf{0},1}}$, the usual first order Sobolev norm.

2.2 Representation of ASG functions

The H-condition Given a continuous function $u \in C(\Omega)$, we can approximate it by a function $u_\mathbf{n} \in V_\mathbf{n} = \text{span}\{\varphi_{\mathbf{n},\mathbf{j}}\}$ by means of interpolation on the grid $\Omega_\mathbf{n}^+$, i.e.,

$$u_\mathbf{n}(\mathbf{x}_{\mathbf{n},\mathbf{j}}) = u(\mathbf{x}_{\mathbf{n},\mathbf{j}}) \quad \forall \mathbf{x}_{\mathbf{n},\mathbf{j}} \in \Omega_\mathbf{n}^+ .$$

Obviously, the function $u_\mathbf{n}$ on $\Omega_\mathbf{n}$ is given by

$$u_\mathbf{n} = \sum_\mathbf{j} a_{\mathbf{n},\mathbf{j}} \varphi_{\mathbf{n},\mathbf{j}} , \tag{2.1}$$

where $a_{\mathbf{n},\mathbf{j}} = u(x_{\mathbf{n},\mathbf{j}})$. The error of approximation is well-known (see e.g. Section 2.3). However, in contrast to classical approximation we are not interested in approximation for a fixed $\mathbf{n}$, but in approximation on (the union of) a number of grids $\Omega_\mathbf{n}^+$.

We can make the approximation (2.1) for all grids $\Omega_\mathbf{n}^+$ with $\mathbf{n} \geq \mathbf{0}$. For large enough $\mathbf{n}$ the approximation can be arbitrarily accurate, but the number of degrees of freedom increases geometrically with $|\mathbf{n}|$. Therefore, in practice we select a 'smallest' $\mathbf{n}$ such that an accuracy criterion is satisfied. Notice that keeping the representations in *all* coarser $V_\mathbf{k}$ (all $V_\mathbf{k}$, $\mathbf{0} \leq \mathbf{k} \leq \mathbf{n}$) does not take essentially more coefficients than the representation on the finest grid (i.e., in $V_\mathbf{n}$) alone.

In order to obtain an efficient approximation, it may be useful to distinguish different subregions of the domain Ω, in each of which we make the finest approximation of u in different $V_\mathbf{n}$. We make full — and efficient — use of the system $\{V_\mathbf{n} \mid \mathbf{n} \in \mathbb{N}_0^d\}$, by *in principle* approximating a given function $u \in C(\Omega)$ in *all* $\{V_\mathbf{n} \mid \mathbf{n} \in \mathbb{N}_0^d\}$, but using *in practice* only those coefficients that contribute to a sufficiently accurate representation. This implies that possibly, in practice, the function u is represented in a particular $V_\mathbf{n}$ only on part of the domain Ω. To introduce a (minimal) structure in the family of approximating basis functions $\{\varphi_{\mathbf{n},\mathbf{j}}\}$, we introduce the following condition **H**.

Condition H: If a basis function $\varphi_{\mathbf{n},\mathbf{j}}(\mathbf{x})$ is used in the representation (2.1), all corresponding coarser basis functions (i.e., functions $\varphi_{\mathbf{k},\mathbf{i}}$ for which $\text{supp}(\varphi_{\mathbf{k},\mathbf{i}}) \supset \text{supp}(\varphi_{\mathbf{n},\mathbf{j}})$) are also used for the approximation.

E-, C-, D- and H-representation We call the representation of the approximation of a function $u \in C(\Omega)$ by a collection of such (partial) approximations in the family of spaces $\{V_\mathbf{n}\}$, the *nodal representation*, or the *E-representation* of the approximation. This E-representation requires the coefficients $a_{\mathbf{n},\mathbf{j}} = u(\mathbf{x}_{\mathbf{n},\mathbf{j}})$ corresponding with grid-points $\mathbf{x}_{\mathbf{n},\mathbf{j}}$, to be equal on the different grids $\Omega_\mathbf{n}^+$ at coinciding grid-points $\mathbf{x}_{\mathbf{n},\mathbf{j}}$. Thus, because points from coarser grids coincide with those from

finer ones, a certain consistency is required (and an redundancy exists) in the E-representation of an approximation.

During the computation in an approximation process, the representations of the approximations on all different grids $\Omega_{\mathbf{k}}^+$ do not necessarily always satisfy the consistency condition required for the E-representation. In that case an approximation exists, that is of the form (2.1) on each separate grid $\Omega_{\mathbf{n}}^+$, and the approximation on the whole system is not uniquely determined. Such a representation we call the *C-representation* of a (non-unique) approximation. There, for different $\mathbf{n}$, the approximations $u_{\mathbf{n}}(\mathbf{x})$ do not necessarily coincide at corresponding grid-points $\mathbf{x}_{\mathbf{n},\mathbf{j}}$.

Another way of representing approximations on the family of grids $\{\Omega_{\mathbf{n}}^+\}$ is by partitioning the approximation over the different grids. Then, instead of (2.1) the approximation reads

$$u_h = \sum_{\mathbf{n}} \sum_{\mathbf{j}} a_{\mathbf{n},\mathbf{j}} \varphi_{\mathbf{n},\mathbf{j}} \,. \tag{2.2}$$

In this case, of course, the set of coefficients $\{a_{\mathbf{n},\mathbf{j}}\}$ always determines a unique function u_h. An approximation in this form we call a *D-representation*. However, for a given function u_h, now the coefficients $\{a_{\mathbf{n},\mathbf{j}}\}$ are not uniquely determined because the $\{\varphi_{\mathbf{n},\mathbf{j}}\}$ are linearly dependent.

One way to select a special unique D-representation is by choosing the coefficients $a_{\mathbf{n},\mathbf{j}}$ such that $a_{\mathbf{n},\mathbf{j}} \neq 0$ only for those $(\mathbf{n},\mathbf{j})$ for which $\|\mathbf{j}\|$ is odd[1]. This implies that $a_{\mathbf{n},\mathbf{j}} = 0$ except for a pair $(\mathbf{n},\mathbf{j})$ for which $\Omega_{\mathbf{n}}^+$ is the coarsest grid which contains the nodal point $\mathbf{x}_{\mathbf{n},\mathbf{j}}$. This representation

$$u_h = \sum_{(\mathbf{n},\mathbf{j}),\|\mathbf{j}\| \text{ odd}} a_{\mathbf{n},\mathbf{j}} \varphi_{\mathbf{n},\mathbf{j}} \tag{2.3}$$

we call the *H-representation* because it represents the approximation in the *hierarchical basis*

$$\left\{\varphi_{\mathbf{n},\mathbf{j}} \,|\; \mathbf{n} \in \mathbb{N}_0^d, \mathbf{j} \in \mathbb{Z}^d,\ \|\mathbf{j}\| \text{ odd}, \mathbf{x}_{\mathbf{n},\mathbf{j}} \in \Omega_{\mathbf{n}}^+\right\} \,, \tag{2.4}$$

and the part of u_h in

$$W_{\mathbf{n}} = \operatorname{span}\{\varphi_{\mathbf{n},\mathbf{j}} \mid \mathbf{j} \in \mathbb{Z}^d,\ \|\mathbf{j}\| \text{ odd}, \mathbf{x}_{\mathbf{n},\mathbf{j}} \in \Omega_{\mathbf{n}}^+\}$$

is the *hierarchical contribution* from the grid $\Omega_{\mathbf{n}}^+$ to the approximation. We notice that

$$V_{\mathbf{n}} = W_{\mathbf{n}} + \sum_{j=1}^{d} V_{\mathbf{n}-\mathbf{e}_j} = \sum_{\mathbf{0} \leq \mathbf{m} \leq \mathbf{n}} V_{\mathbf{m}} \,,$$

and the sparse grid space is defined by

$$V_L = \sum_{0 \leq |\mathbf{m}| \leq L} V_{\mathbf{m}} \,.$$

Interpolating the function u at the nodal points $\mathbf{x}_{\mathbf{n},\mathbf{j}}$, the hierarchical coefficients $a_{\mathbf{n},\mathbf{j}}$ in

$$u(\mathbf{x}_{\mathbf{n},\mathbf{j}}) = \sum_{(\mathbf{n},\mathbf{j}),\|\mathbf{j}\| \text{ odd}} a_{\mathbf{n},\mathbf{j}}\ \varphi_{\mathbf{n},\mathbf{j}}(\mathbf{x}_{\mathbf{n},\mathbf{j}}) \tag{2.5}$$

are determined by (cf. [13])

$$a_{\mathbf{n},\mathbf{j}} = \prod_{i=1}^{d} \left[-\frac{1}{2}, 1, -\frac{1}{2}\right]_{h_{n_i}\mathbf{e}_i} u(\mathbf{j}\mathbf{h}_{\mathbf{n}}) \,, \tag{2.6}$$

[1] With "$\|\mathbf{j}\|$ is odd" we mean: for all $i = 1, 2, \ldots, d$, either j_i is an odd integer, or $k_i = 0$ (i.e., j_i lives on the coarsest grid in the i-direction), see Section 2.1.

where $\left[-\frac{1}{2}, 1, -\frac{1}{2}\right]_{h_{n_i}\mathbf{e}_i}$ denotes the difference stencil for the mesh-size h_{n_i} in the i-th coordinate direction. Notice that this expression is well-defined for each odd $\mathbf{j}$ because Condition **H** requires that all h_i-neighbors are nodal points in the approximation. Another expression for the coefficient $a_{\mathbf{n},\mathbf{j}}$ is found in the following lemma, see [13].

Lemma 2.1 Let $u \in C^{\mathbf{e}+\mathbf{m}}$, for a given $\mathbf{m}$ with $\mathbf{0} \leq \mathbf{m} \leq \mathbf{e}$. Then, for each $\varphi_{\mathbf{n},\mathbf{j}} \in W_{\mathbf{n}}$, we have

$$\begin{aligned} a_{\mathbf{n},\mathbf{j}} &= \prod_{i=1}^{d} \left[-\frac{1}{2}, 1, -\frac{1}{2}\right]_{h_{n_i}\mathbf{e}_i} u(\mathbf{j}\mathbf{h}_{\mathbf{n}}) \\ &= (-1)^{|\mathbf{e}+\mathbf{m}|}\, 2^{-d-|\mathbf{n}|} \int_{\Omega} D^{\mathbf{e}+\mathbf{m}} u(\mathbf{x})\, D^{\mathbf{e}-\mathbf{m}} \varphi(2^{\mathbf{n}}\mathbf{x} - \mathbf{j})\, d\Omega\,. \end{aligned} \tag{2.7}$$

Transformation of representations From the above, it is clear that each H-representation is a D-representation and each E-representation a C-representation. For piecewise d-linear functions, it is often described [4, 5, 6] how a pyramid algorithm can be used to convert an E-representation to a H-representation, and vice versa. Such a conversion can be executed in $\mathcal{O}(N)$ operations, where N is the total number of coefficients (degrees of freedom). The transformation from a D-representation to an H-representation is equally straightforward.

The E-, H-, D-, and C-representations can also be used for piecewise constant functions, and – because of the tensor product structure – discrete function representations can be combined in the different coordinate directions. E.g., a discrete function can be piecewise constant in one and piecewise linear in the other coordinate directions. Also for the piecewise constant functions, efficient pyramid conversion algorithms exist between the different (H-, D-, E-) representation styles. In this case, it is often useful to decide on left- or right-continuity at the discontinuities in the representation.

The data structure The data structure used to implement all the above possibilities of an adaptive (sparse) grid representation can be efficient and relatively simple. For the d-dimensional case ($d = 1, 2, 3$), we use the data structure BASIS3 [12] that takes the 'patch' $P_{\mathbf{n},\mathbf{j}}$ as an elementary entity. This $P_{\mathbf{n},\mathbf{j}}$ takes all information related to a right-open left-closed cell

$$\prod_{k=1}^{3} \left[j_k 2^{-n_k}, (j_k+1)2^{-n_k}\right)\,.$$

This implies that there exist as many patches in the data structure as there are points used in the description of the approximation. The patches are related to each other by means of pointers in an intertwined tree structure, where each patch has at most 15 pointers to related patches (3 fathers, 6 neighbors and 6 kids). The data structure is symmetric with respect to any of the coordinate directions.

2.3 Approximation by ASG functions

The representation of functions as discussed in Section 2.2 includes as special cases the representation, i.e. the d-linear interpolation of continuous functions both on full grids $\Omega^+_{\mathbf{n}}$, especially on regular grids $\Omega^+_{n\mathbf{e}}$, and on sparse grids Ω^+_{ℓ}. In this section, we collect some approximation results from [13] in order to recall the motivation for using sparse grids in the numerical treatment of elliptic boundary value problems.

Errors on full grids Let $u_{\mathbf{n}} \in V_{\mathbf{n}}$ be the piecewise d-linear interpolant at $\bigcup_{\mathbf{k}\leq\mathbf{n}} \Omega^+_{\mathbf{k}} = \Omega^+_{\mathbf{n}}$ of a continuous function u with finite norm $\|D^{\mathbf{e}+\mathbf{m}}u\|_p$, with $p = 2, \infty$, and $\mathbf{0} \leq \mathbf{m} \leq \mathbf{e}$. Error bounds for the approximation on full grids are well-known (e.g., [13, Thm.3.3]):

- $\|u - u_{\mathbf{n}}\|_2 \leq 2^{-d} 3^{-3|\mathbf{m}|/2} \quad \sum_{i=1}^{d} h_{n_i}^{(1+m_i)} \quad \|D^{\mathbf{e}+\mathbf{m}}u\|_2;$

- $\|u - u_{\mathbf{n}}\|_\infty \leq 6^{-|\mathbf{m}|} \quad \sum_{i=1}^d h_{n_i}^{(1+m_i)} \quad \|D^{\mathbf{e}+\mathbf{m}} u\|_\infty$.

It follows that for square grids, for which $\|\mathbf{h}_{n\cdot\mathbf{e}}\|^2 = \sum_{i=1}^d \left(2^{-n}\right)^2 = d2^{-2n}$, we have

- $\|u - u_{n\cdot\mathbf{e}}\|_2 \leq 3^{-2d} \ \|\mathbf{h}_{n\cdot\mathbf{e}}\|^2 \quad \|D^{2\cdot\mathbf{e}} u\|_2$;
- $\|u - u_{n\cdot\mathbf{e}}\|_\infty \leq 6^{-d} \ \|\mathbf{h}_{n\cdot\mathbf{e}}\|^2 \quad \|D^{2\cdot\mathbf{e}} u\|_\infty$.

By simple counting, it is seen that for this approximation the number of degrees of freedom is $\mathcal{O}(2^{|\mathbf{n}|})$, and therefore for a square grid $\mathcal{O}(2^{nd})$.

Errors on sparse grids Let $\hat{u}_n \in V_n$ be the piecewise d-linear interpolant at the sparse grid $\bigcup_{|\mathbf{k}|\leq n} \Omega_{\mathbf{k}}^+ = \Omega_n^+$ of a continuous function u with finite norm $\|D^{\mathbf{e}+\mathbf{m}} u\|_p$, with $p = 2, \infty$, and $\mathbf{0} \leq \mathbf{m} \leq \mathbf{e}$. Error bounds for approximations on these sparse grids [13, Thm.3.6] are given by

- $\|u - \hat{u}_n\|_p \leq C \ \|\mathbf{h}\|^2 \ (\log \|\mathbf{h}\|^{-1})^{d-1} \ \|D^{2\cdot\mathbf{e}} u\|_p$;
- $\|u - \hat{u}_n\|_p \leq C \ \|\mathbf{h}\| \ (\log \|\mathbf{h}\|^{-1})^{d-1-|\mathbf{m}|} \ \|D^{\mathbf{e}+\mathbf{m}} u\|_p$;
- $\|u - \hat{u}_n\|_p \leq C \ \|\mathbf{h}\|^{1+\ell/d} \ (\log \|\mathbf{h}\|^{-1})^{d-1} \ \|u\|_{W_p^{\mathbf{e},\ell}}$.

where ℓ is any integer, $0 \leq \ell \leq d$. Bounds for the errors on sparse grids in energy norm are [13, Thm.3.7]:

- $\|u - \hat{u}_n\|_{W_p^1} \leq C \ \|\mathbf{h}\| \ (\log \|\mathbf{h}\|^{-1})^{d-1} \ \|u\|_{W_p^{\mathbf{e},1}}$;
- $\|u - \hat{u}_n\|_{W_p^1} \leq C \ \|\mathbf{h}\| \ \|D^{2\cdot\mathbf{e}} u\|_p$.

Notice that here $\|\mathbf{h}\| = 2^{-|\mathbf{k}|}$ is the volume of the finest cells in the sparse grid.

For this approximation, by counting, we see that the number of degrees of freedom for the sparse grid approximation is $\mathcal{O}(2^n n^{d-1})$. This number is significantly less than the $\mathcal{O}(2^{nd})$ for the full grid.

3. Evaluation of difference operators to ASG functions

Although finite element discretization of a PDE on a sparse grid is feasible for a constant coefficient problem in two dimensions, finite elements for more-dimensional problems and variable coefficients give problems. The difficulty arises because — with the hierarchical basis (2.4) for test and trial space — the computational complexity of the evaluation of the discrete operator becomes too large. This is caused by the fact that the intersection of the supports of an arbitrary trial and test function is much smaller than the supports of these functions themselves. This has as a consequence that the advantage of sparse grids is lost if the FEM discrete operator is evaluated.

The alternative, as it was already suggested in [6, 17], is the use of a finite difference discretization. Therefore, in order to solve PDEs on sparse grids, we should be able to apply (approximate) differentiation to discrete representations of approximations as described in Section 2.2. The application of linear difference operators approximating the linear differential operator

$$\sum_{i,j} \frac{\partial}{\partial x_i} \left(A_{ij}(\mathbf{x}) \frac{\partial}{\partial x_j} \right) + \sum_i B_i(\mathbf{x}) \frac{\partial}{\partial x_i} + C(\mathbf{x}) \tag{3.1}$$

comes down to the construction of linear combinations of ASG functions, and pointwise multiplication and approximate differentiation of such functions. In any of the representations (E-, H-, D- or C-), the construction of a linear combination over the real numbers is directly computed by application of the linear combination to its coefficients. Pointwise multiplication is only possible in the E-representation, in which the function values at grid-points are directly available (see [6, 17] for details). Below we describe differentiation, which requires some more attention, distinguishing between the evaluation of first and second order derivatives.

3.1 First order derivatives

For a piecewise d-linear ASG function, the derivative $\frac{\partial}{\partial x_i} u_h(\mathbf{x})$ is well-defined almost everywhere, and, written in D-representation (2.2) the derivative is simply described by

$$\begin{aligned}
\frac{\partial}{\partial x_i} u_h(\mathbf{x}) &\equiv D_i u_h(\mathbf{x}) = \sum_{\mathbf{n},\mathbf{j}} a_{\mathbf{n},\mathbf{j}} D_i \varphi_{\mathbf{n},\mathbf{j}}(\mathbf{x}) \\
&= \sum_{\mathbf{n},\mathbf{j}} a_{\mathbf{n},\mathbf{j}}\, D_i \prod_{k=1}^{d} \varphi(2^{n_k} x_k - j_k) \\
&= \sum_{\mathbf{n},\mathbf{j}} a_{\mathbf{n},\mathbf{j}}\, D_i \left(\varphi(2^{n_i} x_i - j_i)\right) \prod_{k=1,k\neq i}^{d} \varphi(2^{n_k} x_k - j_k) \\
&= \sum_{\mathbf{n},\mathbf{j}} a_{\mathbf{n},\mathbf{j}}\, \left(\eta(2^{n_i} x_i - j_i + 1) - \eta(2^{n_i} x_i - j_i)\right) \prod_{k=1,k\neq i}^{d} \varphi(2^{n_k} x_k - j_k)
\end{aligned}$$

This, again, is a function in D-representation, piecewise constant in the i-direction and piecewise linear in the other directions. It can be described by coefficients associated with nodal points if we decide to choose either a left- or a right-continuous representation.

The functions are piecewise linear in all coordinate directions except for the i-th direction, where the derivative is piecewise constant. Because of the discontinuities of $\frac{\partial}{\partial x_j} \varphi_{\mathbf{n},\mathbf{m}}(\mathbf{x})$ along segments parallel to the j-th axis, the function is not in representation (2.1) because the coefficients do not correspond to function values. In order to identify such derivative by values at nodal points, we have to agree on either left-continuity (i.e., $\varphi(x) = \lim_{\xi \searrow x} \varphi(\xi)$) or on right-continuity (i.e., $\varphi(x) = \lim_{\xi \nearrow x} \varphi(\xi)$). Like forward and backward differences, both possibilities have their advantages and disadvantages, and in the implementation we allow left- as well as right-continuity, which are identified by the labels $\alpha = -1$ or $\alpha = 1$, respectively.

3.2 Second order derivatives

The computation of second order derivatives of piecewise d-linear ASG functions, $D_i^2 u_h(x)$, seems to be less obvious because second derivatives of the piecewise d-linear functions vanish almost everywhere on Ω. Nevertheless the approximation of the second order derivatives is useful and can be easily derived from the representation of u_h.

In order to approximate $D_i^2 u_h$, we first construct the representation that is the H-representation in the i-th coordinate direction and the E-representation in the other coordinate directions, i.e., we apply the pyramid algorithm only in the i-th direction. This implies, cf. Equation (2.7),

$$\begin{aligned}
a_{\mathbf{n},\mathbf{j}} &= \left[-\frac{1}{2}, 1, -\frac{1}{2}\right]_{h_{n_i}\mathbf{e}_i} u(\mathbf{j}\mathbf{h}_\mathbf{n}) \qquad (3.2) \\
&= -2^{-1-n_i} \int D_i^2 u(\mathbf{x})\, \varphi(2^{n_i} x_i - j_i)\, dx_i \,.
\end{aligned}$$

where $x_k = 2^{-n_k} j_k$ for $k \neq i$.

For the H-representation of the second derivative, we see that the coefficients in the expression (2.2) are given by (2.7). It follows that the hierarchical coefficient $a_{\mathbf{n},\mathbf{j}}$ with respect to the i-th coordinate direction corresponds with a measure of $D_{x_i}^2 u_h(\mathbf{x})$ in the $\mathbf{h}_\mathbf{n}$-neighborhood of $\mathbf{x}_{\mathbf{n},\mathbf{j}}$.

Considering the i-th coordinate in (3.2), writing $h = 2^{-n}$, and omitting higher order terms in h, we

see

$$\begin{aligned} a_{n,j} &= -\frac{h}{2}\int D_i^2 u(x)\varphi(x/h-j)\,dx \\ &= -\frac{h}{2}\int \left(D_i^2 u(jh) + (x-jh)D_i^3 u(jh) + \frac{(x-jh)^2}{2}D_i^4 u(z)\right)\varphi(x/h-j)\,dx \\ &= -\frac{h}{2}\int \left(D_i^2 u(jh) + (\xi h)D_i^3 u(jh) + \frac{(\xi h)^2}{2}D_i^4 u(z)\right)\varphi(\xi)\,d\xi h \\ &= -\frac{h}{2}\left(D_i^2 u(jh)\int \varphi(\xi)\,d\xi h + D_i^3 u(jh)\int (\xi h)\varphi(\xi)\,d\xi h + R\right) \\ &= -2^{-1-2n}D_i^2 u(j2^{-n}) + R \end{aligned} \tag{3.3}$$

with

$$|R| \leq \frac{h}{2}\|D_i^4 u(z)\|\int \frac{(\xi h)^2}{2}\varphi(\xi)\,d\xi h = \frac{4}{3}\left(\frac{h}{2}\right)^4 \|D_i^4 u(z)\|\,.$$

Here $\|\cdot\|$ denotes the maximum norm in the $\mathbf{h_n}$-neighborhood of $\mathbf{x_{n,j}}$. Such coefficients, and hence such approximate second derivatives, are directly available at the hierarchical points (points with odd j). At points with even j (not at the boundary), we can use

$$\begin{aligned} a_{n,2j} &= -\frac{h}{2}\int D_i^2 u(x)\ \varphi(\frac{x}{h}-2j)\,dx \\ &= -\frac{h}{2}\int D_i^2 u(x)\left(\varphi\left(\frac{x}{2h}-2j\right) - \frac{\varphi(\frac{x}{h}-(2j-1)) + \varphi(\frac{x}{h}-(2j+1))}{2}\right)dx \\ &= \frac{1}{2}(a_{n-1,j} - a_{n,j-1} - a_{n,j+1})\,. \end{aligned} \tag{3.4}$$

The values $a_{n-1,j}$ are available (by recursion) from coarser grids, where we have

$$a_{n-1,j} \approx -2^{1-2n}D_i^2 u(j2^{1-n})\,.$$

This procedure can be used in any of the coordinate directions i. As a result, we find the nodal representation of the (approximate) second derivative in the i-th direction. If we start with the H-representation (in every direction) of the function, this procedure would deliver the second derivative in the i-th direction in E-representation in the i-th direction and in H-representation in all other directions.

4. Properties of the finite difference discretization of the Laplacian

In the remaining part of this paper, we analyze the discretized operator (3.1) as described in Section 3. For simplicity, we restrict ourselves to the model problem of Poisson's equation with homogeneous Dirichlet boundary conditions,

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ u|_{\delta\Omega} &= 0, \end{aligned} \tag{4.1}$$

on the cube $\Omega = (0,1)^d$ and a regular sparse grid. For this special case, we give explicit formulae for the entries of the sparse grid finite difference matrix. This will help us to establish relations between the finite difference discretization on different sparse grid levels and the difference between sparse and full grids. From this, we will finally propose multilevel-type algorithms to solve the linear system of equation resulting from the finite difference discretization (on the sparse grid).

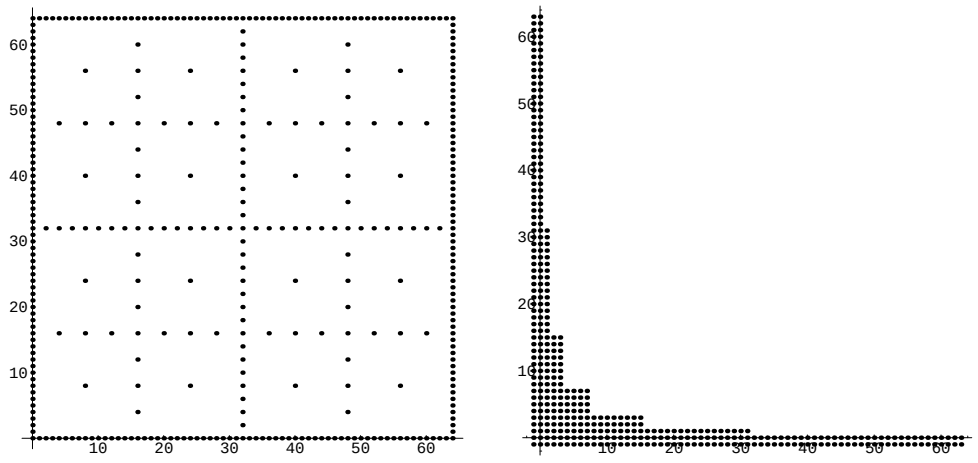

Figure 1: The index sets $\mathcal{J}_6$ and $\mathcal{H}_6$

4.1 Characterization of sparse grid and hierarchical points

To start with, it is useful to investigate the index sets associated with sparse grid points and hierarchical points, respectively. Consider a sparse grid Ω_L^+ of level L on $\Omega = [0,1]^d$ with a finest cell volume $h = 2^{-L}$. Then, every sparse grid point is also a point of a regular grid $\Omega_{L\mathbf{e}}^+$ (with step size h in each direction). Now, we want to characterize the points of the regular grid which belong to the sparse grid, i.e., we wish to characterize the index set $\mathcal{J}_L$ with

$$\mathcal{J}_L = \{\boldsymbol{\ell} \mid x_{L\mathbf{e},\boldsymbol{\ell}} \in \Omega_L^+\}\,.$$

For this, we use the notation of *bit reversing*. Let the integer k, satisfying $0 \le k < 2^\ell$, have the binary representation

$$k = \sum_{s=0}^{\ell-1} k_s\, 2^s, \qquad k_s \in \{0,1\}\,, \tag{4.2}$$

then, the binary representation *reflected at position* $\ell - 1$ (or *in bit reversed order*) is given by[2]

$$M_\ell(k) = \sum_{s=0}^{\ell-1} k_s\, 2^{\ell-1-s}\,.$$

We define the bit reversing of a multi-integer $\mathbf{k}$ with respect to a multi-integer $\boldsymbol{\ell}$ as $M_{\boldsymbol{\ell}}(\mathbf{k}) = (M_{\ell_1}(k_1), \ldots, M_{\ell_d}(k_d))$ and with a scalar m by $M_m(\mathbf{k}) = M_{m\mathbf{e}}(\mathbf{k})$.

For $\mathbf{r} \in \mathbb{N}_0^d$, we also introduce the set of multi-integers (see Figure 1) within a dyadic range:

$$\mathcal{K}_{\mathbf{r}} = \{\mathbf{k} \mid k_i = 2^{r_i-1}, \ldots, 2^{r_i} - 1,\ i = 1, \ldots, d\}$$

in order to define[3] the family of multi-integers with dyadic range up to level L:

$$\mathcal{H}_L = \bigcup_{|\mathbf{r}| \le L} \mathcal{K}_{\mathbf{r}}\,.$$

[2] For $0 \le k \le 2^\ell$, we additionally define $M_\ell(2^\ell) = -1$ and $M_\ell(-1) = 2^\ell$.

[3] To include multi-integers $\mathbf{k}$ with $\lceil \mathbf{k} \rceil = 2^\ell$, we denote $N_\ell = 2^\ell$ for $\ell \in \mathbb{N}_0$ and $N_{-1} = -1$. and $\mathcal{K}_{\mathbf{r}} = \{\mathbf{k},\ k_i = N_{r_i-1}, \ldots, N_{r_i} - 1, i = 1, \ldots, d\}$.

Then one can easily prove (generalization of [20, 21]) that the mapping of bit reversing between the two index sets

$$M_L : \mathcal{H}_L \rightarrow \mathcal{J}_L$$

is a bijection. Furthermore, the set $\bigcup_{|\mathbf{r}|=L} \mathcal{K}_\mathbf{r}$ exactly characterizes the grid points that are added to the sparse grid of level $L-1$ to obtain the sparse grid of level L.

One single set $\mathcal{K}_\mathbf{r}$ represents (i.e., are bit reversed indices of indices of) the hierarchical points in the subset $\Omega^+_\mathbf{r} \subset \Omega^+_L$. Thus, the hierarchical points in $\Omega_\mathbf{r}$ are characterized by

$$\{x_{\mathbf{r},M_\mathbf{r}(\mathbf{k})} = x_{L\mathbf{e},M_L(\mathbf{k})} \mid \mathbf{k} \in \mathcal{K}_\mathbf{r}\}.$$

Remark. The same technique can be used starting on a coarsest grid with smaller mesh width $1/2^{t_0}$ instead of 1. Then M_ℓ has to be defined as a bit reversing w.r.t. $\ell + t_0$ instead of ℓ. If we do not start with powers of 2 the situation is more complex but in principle the same approach can be followed (see [20]).

4.2 Hierarchical representation for finite differences in the univariate case

In Section 3.2, we gave a recursive expression for the nodal representation of the second derivative. In this section, we find an explicit expression for the hierarchical coefficients of the second derivatives found in Section 3.2.

For notational convenience, we first develop the hierarchical representation for second order finite differences in the univariate case. Then, our model problem reads as

$$-u''(x) = f \quad \text{for } x \in (0,1), \qquad u(0) = u(1) = 0.$$

We interpolate the function u on the equidistant grid with step size h_n by a linear spline u_n. This spline function has the hierarchical representation

$$u_n = \sum_{m=1}^{n} \sum_{k \text{ odd}} a_{m,k}\varphi_{m,k}.$$

Note that, because of the homogeneous boundary conditions, the hierarchical coefficients vanish on the coarsest level. We approximate the second derivative of a function u by the second difference

$$\Delta_{2,h}u(x) = \frac{1}{h^2}(u(x+h) - 2u(x) + u(x-h)).$$

In hierarchical points on the finest grid, the hierarchical coefficients represent the second difference $a_{n,k} = -2^{-1-n}\Delta_{2,h_n}u(x_{n,k})$ (cf. Section 3.2). In non-hierarchical points, we can use the recurrence relation (3.4). Using bit reversing, we write this now in closed form. Let $x_{m,k} \in \Omega^+_n$ be a hierarchical point (k odd) on the subgrid Ω^+_m with step size h_m, then $M_m(k) \in \mathcal{K}_m$ and writing out the recursion (3.4), we obtain

$$-\Delta_{2,h_n}u(x_{m,k}) = 2^{1+n}\left(2^m a_{m,k} - \sum_{\ell=m+1}^{n} 2^{\ell-1}(a_{\ell,2^{\ell-m}k-1} + a_{\ell,2^{\ell-m}k+1})\right). \tag{4.3}$$

Now, we interpolate the nodal approximation on level n of the second derivative of the function u by a linear spline[4] and obtain

$$-\tilde{u}_n = \sum_{k=1}^{2^n-1} -\Delta_{2,h_n}u(x_{n,k})\,\varphi_{n,k} = \sum_{m=1}^{n} \sum_{k \text{ odd}} b_{m,k}\varphi_{m,k}\,. \tag{4.4}$$

[4] For notational convenience, we set in the end points $\Delta_{2,h_n}u(0) = \Delta_{2,h_n}u(1) = 0$.

This defines the coefficients $b_{m,k}$ which clearly satisfy

$$b_{m,k} = \frac{1}{2}(\Delta_{2,h_n} u(x_{m,k-1}) - 2\Delta_{2,h_n} u(x_{m,k}) + \Delta_{2,h_n} u(x_{m,k+1}))$$

for hierarchical points $x_{m,k}$. The non-hierarchical points $x_{m,k-1}$ and $x_{m,k+1}$ appear in this formula, too. If we want to apply formula (4.3), we need to characterize these points (non-hierarchical on level m) as hierarchical points on some coarser level. For this, we need some further notation.

Let k have the binary representation (4.2). Now we need to characterize the indices m_- and m_+ for which $M_m(k-1) \in \mathcal{K}_{m_-}$ and $M_m(k+1) \in \mathcal{K}_{m_+}$, where $m_\pm$ denotes the level on which $x_{m,k\pm1}$ is hierarchical. We see that

$$m_- = \begin{cases} 0 & \text{for } k = 1, \\ \max\limits_{s=1,\dots,m-1} \{m-1-s, k_s = 1\} & \text{otherwise,} \end{cases}$$

and

$$m_+ = \begin{cases} 0 & \text{for } k = 2^m - 1, \\ \max\limits_{s=1,\dots,m-1} \{m-1-s, k_s = 0\} & \text{otherwise.} \end{cases}$$

Thus, if the points $x_{m,k}$ are not next to the boundary, i.e., for[5] $m_\pm \neq 0$,

$$\begin{aligned} &-\Delta_{2,h_n} u(x_{m,k\pm1}) \\ &\quad = \quad 2^{1+n}\left(2^{m_\pm} a_{m_\pm, 2^{m_\pm - m}(k\pm1)} - \sum_{\ell=m_\pm+1}^{n} 2^{\ell-1}(a_{\ell,2^{\ell-m}(k\pm1)-1} + a_{\ell,2^{\ell-m}(k\pm1)+1})\right). \end{aligned}$$

Collecting all terms, we get the hierarchical coefficient

$$\begin{aligned} b_{m,k} \quad = \quad & 2^{1+n}\Bigg(2^m a_{m,k} - \sum_{\ell=m+1}^{n} 2^{\ell-1}(a_{\ell,2^{\ell-m}k-1} + a_{\ell,2^{\ell-m}k+1}) \\ & - 2^{m_- - 1} a_{m_-, 2^{m_- - m}(k-1)} + \sum_{\ell=m_-+1}^{n} 2^{\ell-2}(a_{\ell,2^{\ell-m}(k-1)-1} + a_{\ell,2^{\ell-m}(k-1)+1}) \\ & - 2^{m_+ - 1} a_{m_+, 2^{m_+ - m}(k+1)} + \sum_{\ell=m_++1}^{n} 2^{\ell-2}(a_{\ell,2^{\ell-m}(k+1)-1} + a_{\ell,2^{\ell-m}(k+1)+1})\Bigg) \end{aligned} \tag{4.5}$$

with modifications for $m_\pm = 0$. Note that this expression depends on m, k and n, i.e., it depends not only on the point $x_{m,k}$ but also on the finest grid chosen.

We can write the coefficients $a_{m,k}$ and $b_{m,k}$ in vector notation and describe the transformation process by a matrix. For this, we combine hierarchical coefficients in a vector as

$$\boldsymbol{v}_n = \big(v_{m,k}\big)_{1 \le m \le n, M_m(k) \in \mathcal{K}_m} \tag{4.6}$$

in order to define the vectors $\boldsymbol{a}_n$ and $\boldsymbol{b}_n$, and nodal coefficients simply as

$$\boldsymbol{u}_n = \big(u(x_{n,1}), \dots, u(x_{n,2^n-1})\big)^T.$$

In this way, we define the matrices $\boldsymbol{A}_n$ and $\boldsymbol{H}_n$ by

$$\boldsymbol{A}_n\, \boldsymbol{a}_n = \boldsymbol{b}_n \qquad \text{and} \qquad \boldsymbol{H}_n\, \boldsymbol{u}_n = \boldsymbol{a}_n.$$

[5] For points next to the boundary, i.e., for $m_\pm = 0$, we use the notation from the last three footnotes and set $\Delta_{2,h_n} u(x_{m,k\pm1}) = 0$.

By construction, $\boldsymbol{H}_n^{-1}\boldsymbol{A}_n\boldsymbol{H}_n$ is the usual matrix of second differences (applied to nodal values) and hence symmetric positive definite. Thus, the matrix $\boldsymbol{A}_n$ of finite differences in terms of hierarchical coefficients has the same, i.e., only real positive eigenvalues. However, it is not symmetric and can not be symmetric positive definite.

Solving the discretized system in its hierarchical form now comes down to solving

$$\boldsymbol{A}_n\,\boldsymbol{a}_n = \boldsymbol{H}_n\left(f(x_{n,1}),\ldots,f(x_{n,2^n-1})\right)^T.$$

4.3 Hierarchical representation for finite differences on the sparse grid
To gain some more insight in the structure of the sparse grid FD operator, we want to obtain an explicit expression for the hierarchical coefficients of the discrete Laplacian. I.e., we look for an expression for the elements of the FD matrix $\boldsymbol{A}_L$. To this end, we investigate the discrete second derivative in the x_1-direction first.

Finite differences on one fixed level Let L be the (highest) level of the sparse grid. Let u_L be given in H-representation

$$u_L = \sum_{|\mathbf{m}|\le L}\ \sum_{\|\boldsymbol{\ell}\|\ \mathrm{odd}} a_{\mathbf{m},\boldsymbol{\ell}}\,\varphi_{\mathbf{m},\boldsymbol{\ell}}.$$

Because of the boundary conditions, here and in the sequel the notation $|\mathbf{m}| \le L$ means: $\mathbf{m} > \mathbf{0}$, $|\mathbf{m}| \le L$. (The coefficients responsible for the boundary are zero). Let $\tilde{u}_L$ be the approximation of the resulting function Δu on the sparse grid Ω_L^+ with

$$-\tilde{u}_L = \sum_{|\mathbf{m}|\le L}\ \sum_{\|\boldsymbol{\ell}\|\ \mathrm{odd}} b^{\Delta}_{\mathbf{m},\boldsymbol{\ell}}\,\varphi_{\mathbf{m},\boldsymbol{\ell}}\,.$$

We denote the approximation of the second derivative in x_ν-direction on the sparse grid Ω_L^+ by

$$-\tilde{u}_L^{(\nu)} = \sum_{|\mathbf{m}|\le L}\ \sum_{\|\boldsymbol{\ell}\|\ \mathrm{odd}} b^{(\nu)}_{\mathbf{m},\boldsymbol{\ell}}\,\varphi_{\mathbf{m},\boldsymbol{\ell}}.$$

Then, the hierarchical coefficients obviously fulfill

$$b^{\Delta}_{\mathbf{m},\boldsymbol{\ell}} = b^{(1)}_{\mathbf{m},\boldsymbol{\ell}} + b^{(2)}_{\mathbf{m},\boldsymbol{\ell}} + \cdots + b^{(d)}_{\mathbf{m},\boldsymbol{\ell}}\,. \tag{4.7}$$

Let $\mathbf{x}_{\mathbf{m},\mathbf{k}} \in \Omega_L^+$ be a hierarchical point on grid $\Omega^+_{\mathbf{m}}$. The full grid $\Omega^+_{\mathbf{j}} \subset \Omega_L$, finest in x_1-direction, such that $\mathbf{x}_{\mathbf{m},\mathbf{k}} \in \Omega^+_{\mathbf{j}}$ is characterized by the multi-index $\mathbf{j} = (m_1 + L - |\mathbf{m}|, m_2, \ldots, m_d)$. Then the hierarchical coefficient of the second difference in the x_1-direction at point $\mathbf{x}_{\mathbf{m},\mathbf{k}}$ is

$$b^{(1)}_{\mathbf{m},\mathbf{k}} = -\alpha_{\mathbf{m},\mathbf{k}} + \frac{1}{2}\left(\alpha_{\mathbf{m},\mathbf{k}-\mathbf{e}_1} + \alpha_{\mathbf{m},\mathbf{k}+\mathbf{e}_1}\right), \tag{4.8}$$

where $\alpha_{\mathbf{m},\boldsymbol{\ell}}$ denotes the coefficient of the approximation of the second derivative in x_1-direction (E-representation in x_1-direction, H-representation in all other directions) at the point $\mathbf{x}_{\mathbf{m},\boldsymbol{\ell}}$.

By construction $M_{\mathbf{m}}(\mathbf{k}) \in \mathcal{K}_{\mathbf{m}}$ and using the results of the previous section, we obtain from the finite difference operator in the x_1-direction, see (4.3),

$$\begin{aligned}\alpha_{\mathbf{m},\mathbf{k}} \;=\; -2^{1+2m_1+L-|\mathbf{m}|}\Bigg(a_{\mathbf{m},\mathbf{k}} - \sum_{\ell=m_1+1}^{m_1+L-|\mathbf{m}|} 2^{\ell-m_1-1}\big(&a_{(\ell,m_2,\ldots,m_d),(2^{\ell-m_1}k_1-1,k_2,\ldots,k_d)}\\ &+ a_{(\ell,m_2,\ldots,m_d),(2^{\ell-m_1}k_1+1,k_2,\ldots,k_d)}\big)\Bigg).\end{aligned}$$

In the same way as in the previous section, we denote by m_{1-} and m_{1+} the indices for which $M_{m_1}(k_1-1) \in \mathcal{K}_{m_{1-}}$ and $M_{m_1}(k_1+1) \in \mathcal{K}_{m_{1+}}$ (i.e., m_{1-} and m_{1-} are the x_1-levels on which $\mathbf{x}_{\mathbf{m},\mathbf{k}}$ are hierarchical points in the x_1-direction). Now we can characterize the remaining terms in (4.8) and obtain

$$\begin{aligned} b^{(1)}_{\mathbf{m},\mathbf{k}} \;=\; & 2^{1+2m_1+L-|\mathbf{m}|} \cdot \Bigg(a_{\mathbf{m},\mathbf{k}} \\ & - \sum_{\ell=m_1+1}^{m_1+L-|\mathbf{m}|} 2^{\ell-m_1-1} \big(a_{(\ell,m_2,\dots,m_d),(2^{\ell-m_1}k_1-1,k_2,\dots,k_d)} \\ & \qquad + a_{(\ell,m_2,\dots,m_d),(2^{\ell-m_1}k_1+1,k_2,\dots,k_d)} \big) \\ & - 2^{m_{1+}-m_1-1} a_{(m_{1+},m_2,\dots,m_d),(2^{m_{1+}-m_1}(k_1+1),k_2,\dots,k_d)} \\ & + \sum_{\ell=m_{1+}+1}^{m_1+L-|\mathbf{m}|} 2^{\ell-m_1-2} \big(a_{(\ell,m_2,\dots,m_d),(2^{\ell-m_1}(k_1+1)-1,k_2,\dots,k_d)} \\ & \qquad + a_{(\ell,m_2,\dots,m_d),(2^{\ell-m_1}(k_1+1)+1,k_2,\dots,k_d)} \big) \\ & - 2^{m_{1-}-m_1-1} a_{(m_{1-},m_2,\dots,m_d),(2^{m_{1-}-m_1}(k_1-1),k_2,\dots,k_d)} \\ & + \sum_{\ell=m_{1-}+1}^{m_1+L-|\mathbf{m}|} 2^{\ell-m_1-2} \big(a_{(\ell,m_2,\dots,m_d),(2^{\ell-m_1}(k_1-1)-1,k_2,\dots,k_d)} \\ & \qquad + a_{(\ell,m_2,\dots,m_d),(2^{\ell-m_1}(k_1-1)+1,k_2,\dots,k_d)} \big) \Bigg) \end{aligned} \tag{4.9}$$

(with obvious modifications for points next to the boundary, where $m_{1-} = 0$ or $m_{1+} = 0$). Here, the finest grid in x_1-direction depends on the other coordinates of the evaluation point. This makes $b^{(1)}_{\mathbf{m},\mathbf{k}}$ depending on $\mathbf{m}, \mathbf{k}$ and the highest level L.

Now, we write the hierarchical coefficients $b^{(\nu)}_{\mathbf{m},\mathbf{k}}$, $b^{\Delta}_{\mathbf{m},\mathbf{k}}$, and $a_{\mathbf{m},\mathbf{k}}$ in vector form as $\boldsymbol{b}^{(\nu)}_L$, $\boldsymbol{b}^{\Delta}_L$, and $\boldsymbol{a}_L$, respectively, using the d-dimensional version of (4.6):

$$\boldsymbol{v}_L = \big(v_{\mathbf{m},\mathbf{k}}\big)_{|\mathbf{m}|\le L, M_{\mathbf{m}}(\mathbf{k})\in\mathcal{K}_{\mathbf{m}}} .$$

We define the matrices $\boldsymbol{A}_L$ and $\boldsymbol{A}^{(\nu)}_L$ by

$$\boldsymbol{A}_L \boldsymbol{a}_L = \boldsymbol{b}^{\Delta}_L \quad \text{and} \quad \boldsymbol{A}^{(\nu)}_L \boldsymbol{a}_L = \boldsymbol{b}^{(\nu)}_L , \tag{4.10}$$

respectively. Then, obviously $\boldsymbol{A}_L = \boldsymbol{A}^{(1)}_L + \boldsymbol{A}^{(2)}_L + \cdots + \boldsymbol{A}^{(d)}_L$.

Finite differences on sparse grids of different levels For the hierarchical coefficient vectors, we introduce matrices that represent the piecewise linear prolongation $\boldsymbol{P}_{L,L-1} : V_{L-1} \to V_L$, defined by $\boldsymbol{w}_L = \boldsymbol{P}_{L,L-1}\boldsymbol{v}_{L-1}$ with

$$w_{\mathbf{m},\mathbf{k}} = \begin{cases} v_{\mathbf{m},\mathbf{k}} & \text{for } |\mathbf{m}| \le L-1, \\ 0 & \text{for } |\mathbf{m}| = L. \end{cases}$$

and similarly a restriction matrix corresponding with $\boldsymbol{R}_{L-1,L} : V_L \to V_{L-1}$, defined by $\boldsymbol{w}_{L-1} = \boldsymbol{R}_{L-1,L}\boldsymbol{v}_L$ with

$$w_{\mathbf{m},\mathbf{k}} = \frac{1}{2}\, v_{\mathbf{m},\mathbf{k}} \quad \text{for } |\mathbf{m}| \le L-1.$$

With these definitions, as can be seen from (4.9), the following Galerkin relation holds

$$\boldsymbol{A}_{L-1} = \boldsymbol{R}_{L-1,L}\boldsymbol{A}_L\boldsymbol{P}_{L,L-1}.$$

This relation is the motivation for further research in [11].

4.4 Hierarchical representation for finite differences on full grids

The finite difference discretization of (4.1) on a full grid is well-known (the usual 7-point stencil in 3D). Here we rewrite its matrix in the hierarchical representation in order to see the relation between the finite difference operators on full and sparse grids, with the aim to find proper multigrid algorithms. Let a discrete function on $\Omega_{\mathbf{n}}^+$, say $u_{\mathbf{n}}$, be given in its hierarchical representation

$$u_{\mathbf{n}} = \sum_{\mathbf{0}<\mathbf{m}\le\mathbf{n}} \sum_{\|\mathbf{k}\| \text{ odd}} a_{\mathbf{m},\mathbf{k}}\, \varphi_{\mathbf{m},\mathbf{k}}\,.$$

Let $\tilde{u}_{\mathbf{n}}^{\Delta}$ denote the approximation of Δu, similar to (4.4) in the one-dimensional case in Section 4.2,

$$-\tilde{u}_{\mathbf{n}}^{\Delta} = \sum_{\mathbf{0}<\mathbf{m}\le\mathbf{n}} \sum_{\|\mathbf{k}\| \text{ odd}} \tilde{b}_{\mathbf{m},\mathbf{k}}^{\Delta}\, \varphi_{\mathbf{m},\mathbf{k}}.$$

Assume the approximation of the second derivative in x_ν-direction $\tilde{u}_{\mathbf{n}}^{(\nu)}$ to have hierarchical coefficients $\tilde{b}_{\mathbf{m},\mathbf{k}}^{(\nu)}$. The coefficients corresponding to the subgrid $\Omega_{\mathbf{n}}^+ \subset \Omega_L^+$ are written in vector-form as

$$\boldsymbol{v}_{\mathbf{n}} = \big(v_{\mathbf{m},\mathbf{k}}\big)_{\mathbf{m}\le\mathbf{n},\, M_{\mathbf{m}}(\mathbf{k})\in\mathcal{K}_{\mathbf{m}}}.$$

Similar to (4.7) in the previous section, we have $\tilde{\boldsymbol{b}}_{\mathbf{n}}^{\Delta} = \tilde{\boldsymbol{b}}_{\mathbf{n}}^{(1)} + \tilde{\boldsymbol{b}}_{\mathbf{n}}^{(2)} + \cdots + \tilde{\boldsymbol{b}}_{\mathbf{n}}^{(d)}$, and we write

$$\boldsymbol{A}_{\mathbf{n}}^{(\nu)} \boldsymbol{a}_{\mathbf{n}} = \tilde{\boldsymbol{b}}_{\mathbf{n}}^{(\nu)} \qquad \text{and} \qquad \boldsymbol{A}_{\mathbf{n}} \boldsymbol{a}_{\mathbf{n}} = \tilde{\boldsymbol{b}}_{\mathbf{n}}^{\Delta}.$$

Obviously,

$$\boldsymbol{A}_{\mathbf{n}} = \boldsymbol{A}_{\mathbf{n}}^{(1)} + \boldsymbol{A}_{\mathbf{n}}^{(2)} + \ldots + \boldsymbol{A}_{\mathbf{n}}^{(d)}\,. \tag{4.11}$$

We introduce a one-dimensional index set by $\mathcal{I}_n = \bigcup_{m=1}^{n} \big\{(m,k) \mid\ M_m(k) \in \mathcal{K}_m\big\}$ and the corresponding identity matrix by

$$\boldsymbol{I}_n = \big(\delta_{\mu,\nu}\big)_{\mu,\nu=1}^{\#(\mathcal{I}_n)}.$$

Then, for the more-dimensional case $\big\{(\mathbf{m},\mathbf{k}) \mid\ M_{\mathbf{m}}(\mathbf{k}) \in \mathcal{K}_{\mathbf{m}}\big\} = \mathcal{I}_{n_1} \times \mathcal{I}_{n_2} \times \cdots \times \mathcal{I}_{n_d}$ and we find the matrix of second differences in the ν-th direction as the Kronecker product

$$\boldsymbol{A}_{\mathbf{n}}^{(\nu)} = \boldsymbol{I}_{n_1} \otimes \cdots \otimes \boldsymbol{I}_{n_{\nu-1}} \otimes \boldsymbol{A}_{n_\nu} \otimes \boldsymbol{I}_{n_{\nu+1}} \otimes \cdots \otimes \boldsymbol{I}_{n_d}.$$

with $\boldsymbol{A}_{n_\nu}$ the same as in Section 4.2. As a Kronecker sum of the matrices $\boldsymbol{A}_{n_\nu}$, the matrix $\boldsymbol{A}_{\mathbf{n}}$ also has only positive eigenvalues. This is not surprising because $\boldsymbol{A}_{\mathbf{n}}$ has the same eigenvalues as the finite difference matrix for the nodal representation on the full grid which is symmetric positive definite.

4.5 Relations between finite differences on full and sparse grids

In this section, we establish relations between the discrete Laplacian on full and on sparse grids. If one considers the Galerkin or FEM approach for discretization on sparse grids, using the hierarchical basis or the generating system of nodal bases functions as test and trial functions, one obtains simple relations between the stiffness matrices for the sparse and the full grids. The basis functions for a full grid $\Omega_{\mathbf{n}}^+ \subset \Omega_L^+$ simply form a subset of the hierarchical basis/generating system for the sparse grid. So, by construction it can be written as a Galerkin product using the matrix for the sparse grid. As a result, one can immediately write down multiplicative subspace correction algorithms like in [7] and interpret them as block iteration methods.

In the finite difference case things are different. In contrast with the sparse grid, on a full grid, for a constant coefficient differential equation the evaluation of the finite difference at a certain point does

not depend on the location of the point. For this reason, the finite difference matrix of a full grid $\Omega^+_{\mathbf{n}}$ can not be written as a Galerkin product including the finite difference matrix of the sparse grid Ω^+_L of level $L \geq |\mathbf{n}|$. This will become immediately clear from equation (4.13), below.

In what follows, we discuss the kind of relation between the finite difference operators for the Laplacian discretized on a full and on a sparse grid, and we propose a solution algorithm resulting from this.

Full grids as subgrids of the sparse grid Assume $|\mathbf{n}| \leq L$. We define a prolongation $\boldsymbol{P}_{L,\mathbf{n}} : V_{\mathbf{n}} \to V_L$ by $\boldsymbol{w}_L = \boldsymbol{P}_{L,\mathbf{n}} \boldsymbol{v}_{\mathbf{n}}$ with

$$w_{\mathbf{m},\mathbf{k}} = \begin{cases} v_{\mathbf{m},\mathbf{k}} & \text{for } \mathbf{m} \leq \mathbf{n}, \\ 0 & \text{otherwise,} \end{cases}$$

where $v_{\mathbf{m},\mathbf{k}}$ and $w_{\mathbf{m},\mathbf{k}}$ are hierarchical coefficients, and a direction dependent restriction $\boldsymbol{R}^{(\nu)}_{\mathbf{n},L} : V_L \to V_{\mathbf{n}}$ by $\boldsymbol{w}_{\mathbf{n}} = \boldsymbol{R}^{(\nu)}_{\mathbf{n},L} \boldsymbol{v}_L$ with

$$w_{\mathbf{m},\mathbf{k}} = 2^{n_\nu - m_\nu - L + |\mathbf{m}|} v_{\mathbf{m},\mathbf{k}} \quad \text{for } \mathbf{m} \leq \mathbf{n}\,. \tag{4.12}$$

The scaling factor in the above definition arises from the difference in scaling between (4.5) and (4.9). Comparing these two formulae and keeping (4.10) and (4.11) in mind, we see that, with these $\boldsymbol{P}_{L,\mathbf{n}}$ and $\boldsymbol{R}^{(\nu)}_{\mathbf{n},L}$, we can write the matrix of finite differences for the Laplacian on the full grid as a sum of Galerkin products using the direction (and point) dependent weighted finite differences for the second derivatives in each direction

$$\boldsymbol{A}_{\mathbf{n}} = \big(\boldsymbol{R}^{(1)}_{\mathbf{n},L} \boldsymbol{A}^{(1)}_L + \boldsymbol{R}^{(2)}_{\mathbf{n},L} \boldsymbol{A}^{(2)}_L + \cdots + \boldsymbol{R}^{(d)}_{\mathbf{n},L} \boldsymbol{A}^{(d)}_L\big) \boldsymbol{P}_{L,\mathbf{n}}. \tag{4.13}$$

Full grids as subgrids of the sparse grid — An alternative approach In the previous paragraph, we scaled the different parts, $A^{(\nu)}_L$, of the discrete Laplacian for the sparse grid for the different directions. On the other hand, we can scale $A^{(\nu)}_{\mathbf{n}}$, i.e., the directional parts of the matrix responsible for the full grid. With the scaling matrices

$$\begin{aligned} \boldsymbol{M}^{(\nu)}_{\mathbf{n}} &= \operatorname{diag}\big(2^{n_\nu - m_\nu}\big)_{\mathbf{m} \leq \mathbf{n}, M_{\mathbf{m}}(\mathbf{k}) \in \mathcal{K}_{\mathbf{m}}} \\ &= \boldsymbol{I}_{n_1} \otimes \cdots \otimes \boldsymbol{I}_{n_{\nu-1}} \otimes \big(\operatorname{diag}\big(2^{m_\nu - n_\nu}\big)_{(m_\nu, k_\nu) \in \mathcal{I}_{n_\nu}}\big) \otimes \boldsymbol{I}_{n_{\nu+1}} \otimes \cdots \otimes \boldsymbol{I}_{n_d} \end{aligned}$$

and (4.11), we define

$$\tilde{\boldsymbol{A}}_{\mathbf{n}} = \boldsymbol{M}^{(1)}_{\mathbf{n}} \boldsymbol{A}^{(1)}_{\mathbf{n}} + \boldsymbol{M}^{(2)}_{\mathbf{n}} \boldsymbol{A}^{(2)}_{\mathbf{n}} + \cdots + \boldsymbol{M}^{(d)}_{\mathbf{n}} \boldsymbol{A}^{(d)}_{\mathbf{n}}. \tag{4.14}$$

This matrix is again a Kronecker sum of matrices with positive eigenvalues and so it has only positive eigenvalues. Introducing another, differently scaled, restriction that is now independent of the direction ν, $\tilde{\boldsymbol{R}}_{\mathbf{n},L} : V_L \to V_{\mathbf{n}}$, defined by $\boldsymbol{w}_{\mathbf{n}} = \tilde{\boldsymbol{R}}_{\mathbf{n},L} \boldsymbol{v}_L$ with

$$w_{\mathbf{m},\mathbf{k}} = 2^{-L + |\mathbf{m}|} v_{\mathbf{m},\mathbf{k}} \quad \text{for } \mathbf{m} \leq \mathbf{n}\,,$$

the new matrix can be written as the Galerkin product

$$\tilde{\boldsymbol{A}}_{\mathbf{n}} = \tilde{\boldsymbol{R}}_{\mathbf{n},L} \boldsymbol{A}_L \boldsymbol{P}_{L,\mathbf{n}}.$$

This relation leads to the following solution algorithm which uses the matrices $\tilde{\boldsymbol{A}}_{\mathbf{n}}$. Denote the vector of hierarchical coefficients of the right-hand side by $\boldsymbol{f}_L$, and let an approximation $\boldsymbol{a}_L$ of the hierarchical coefficients of u_L be given. Then the next iteration step is

for all $|\mathbf{n}| = k$
do $\boldsymbol{a}_L := \boldsymbol{a}_L + \boldsymbol{P}_{L,\mathbf{n}} \tilde{\boldsymbol{A}}_{\mathbf{n}}^{-1} \tilde{\boldsymbol{R}}_{\mathbf{n},L} (\boldsymbol{f}_L - \boldsymbol{A}_L \boldsymbol{a}_L)$ (4.15)
enddo

This defect correction algorithm of multiplicative type corresponds to a block iterative solver for the discrete Laplacian on the sparse grid with overlapping blocks. To some extent this is similar to the FEM case (see [7]), but in case of finite differences the blocks (containing point dependent scaling) cannot really be seen as a finite difference discretization on a certain full grid. Nevertheless, in case of the Laplacian it is still possible to give a relation (4.14) between the blocks and the finite difference discretizations.

To improve the convergence rate of the above algorithm we might think of using more than one level and/or more than one iteration per level which leads to

for $k = \ell$ **to** L
do for $i = 1$ **to** ν
  do for all $|\mathbf{n}| = k$
    do $\boldsymbol{a}_L := \boldsymbol{a}_L + \boldsymbol{P}_{L,\mathbf{n}} \tilde{\boldsymbol{A}}_{\mathbf{n}}^{-1} \tilde{\boldsymbol{R}}_{\mathbf{n},L} (\boldsymbol{f}_L - \boldsymbol{A}_L \boldsymbol{a}_L)$ (4.16)
    enddo
  enddo
enddo

with a fixed lower level $\ell \geq d$ and a number ν of iterations per level.

Example: We apply the Algorithms (4.15) and (4.16) to the following 3D-problem. Solve (4.1) with the right-hand side

$$f(\mathbf{x}) = -3\pi^2 \Big(\prod_{i=1}^{3} \sin \pi x_i + 8 \prod_{i=1}^{3} \sin 8\pi x_i \Big)$$

and starting from the zero function $u_L^{(0)} \equiv 0$. We obtain the convergence behavior shown in Figure 2. We see that we get better convergence if we include also lower levels (right). In both cases, the speed of convergence slows down with growing level. Approximately, the reduction factor gets worse with L^2, the square of the highest level.

The sparse grid as a subgrid of a full grid Thus far, in Sections 4.5 and 4.5, we considered full grids which are contained in a sparse grid. But the sparse grid itself is contained in the full grid with step size 2^{-L} in each direction. So for completeness, we establish the relation between the finite difference matrices of these grids. Define the prolongation $\boldsymbol{P}_{L\mathbf{e},L} : V_L \to V_{L\mathbf{e}}$ by $\boldsymbol{w}_{L\mathbf{e}} = \boldsymbol{P}_{L\mathbf{e},L} \boldsymbol{v}_L$ with

$$w_{\mathbf{m},\mathbf{k}} = \begin{cases} v_{\mathbf{m},\mathbf{k}} & \text{for } |\mathbf{m}| \leq L, \\ 0 & \text{otherwise,} \end{cases}$$

for the hierarchical coefficients $v_{\mathbf{j},\mathbf{k}}$ and $w_{\mathbf{m},\mathbf{k}}$, and the restriction $\boldsymbol{R}_{L,L\mathbf{e}} : V_{L\mathbf{e}} \to V_L$ by $\boldsymbol{w}_L = \boldsymbol{R}_{L,L\mathbf{e}} \boldsymbol{v}_{L\mathbf{e}}$, where

$$w_{\mathbf{m},\mathbf{k}} = 2^{-|\mathbf{m}|} v_{\mathbf{m},\mathbf{k}} \quad \text{for } |\mathbf{m}| \leq L.$$

Further, we need scaling matrices

$$\begin{aligned} \tilde{\boldsymbol{M}}_{L\mathbf{e}}^{(\nu)} &= \operatorname{diag}\big(2^{m_\nu}\big)_{\mathbf{m} \leq L\mathbf{e}, M_{\mathbf{m}}(\mathbf{k}) \in \mathcal{K}_{\mathbf{m}}} \\ &= \boldsymbol{I}_L \otimes \cdots \otimes \boldsymbol{I}_L \otimes \big(\operatorname{diag}\big(2^{m_\nu}\big)_{(m_\nu,k_\nu) \in \mathcal{I}_L} \big) \otimes \boldsymbol{I}_L \otimes \cdots \otimes \boldsymbol{I}_L, \end{aligned}$$

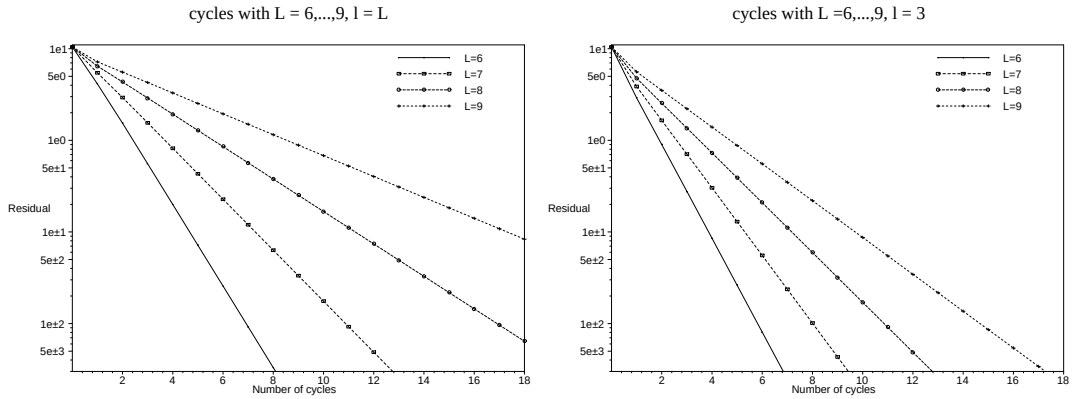

Figure 2: Left: Convergence of Algorithm (4.15) for the levels $L = 6, \ldots, 9$. Right: Convergence of Algorithm (4.16) for the levels $L = 6, \ldots, 9$, $\ell = 3$ with $\nu = 1$.

then, reversing the procedure followed above, we obtain

$$\boldsymbol{A}_L = \boldsymbol{R}_{L,L\mathbf{e}}\big(\tilde{\boldsymbol{M}}^{(1)}_{L\mathbf{e}}\boldsymbol{A}^{(1)}_{L\mathbf{e}} + \tilde{\boldsymbol{M}}^{(2)}_{L\mathbf{e}}\boldsymbol{A}^{(2)}_{L\mathbf{e}} + \cdots + \tilde{\boldsymbol{M}}^{(d)}_{L\mathbf{e}}\boldsymbol{A}^{(d)}_{L\mathbf{e}}\big)\boldsymbol{P}_{L\mathbf{e},L}.$$

The matrix between the parentheses is again a Kronecker sum of matrices with positive eigenvalues and so it has positive eigenvalues. The whole matrix $\boldsymbol{A}_L$ is a scaled submatrix of a matrix with positive eigenvalues.

In numerical experiments by Schiekofer [17] in 2D, it turned out that $\boldsymbol{A}_L$ itself has only positive eigenvalues. (For s.p.d. matrices this is obvious for a submatrix, for non-symmetric ones this need not to be the case.) Of course, because of the size of the full grid matrices involved, the above Galerkin relation generally has no practical computational value.

Conclusion

Because the evaluation of finite element stiffness matrices for variable coefficient equations on sparse grids (in more dimensions, $d > 2$) still yields difficulties, one might be tempted to use finite differences (FD) instead. In this paper, the relation between FD operators on sparse and full grids are studied, and Galerkin relations are established.

In an obvious way, such relations lead to iterative (defect correction) solution algorithms that can also be applied in a multilevel setting. However, no spectral equivalence could be established, and the convergence of the iterative schemes appears to depend on the maximum discretization level used, so that the convergence rate is lower on finer grids.

References

1. H.-J. Bungartz and T. Dornseifer. Sparse grids: Recent developments for elliptic partial differential equations. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, Lecture Notes in Computational Science and Engineering 3. Springer, Berlin, 1998.
2. F.-J. Delvos and H. Posdorf. N–th order blending. In W. Schempp and K. Zeller, editors, *Constructive Theory of Functions of Several Variables*, pages 53–64. Springer, Berlin, 1977.
3. F.-J. Delvos and W. Schempp. *Boolean Methods in Interpolation and Approximation.* Pitman Research Notes in Mathematics Series 230. Longman Scientific & Technical, Harlow, 1989.
4. M. Griebel. A parallelizable and vectorizable multi-level algorithm on sparse grids. In *Parallel algorithms for partial differential equations*, Notes Numer. Fluid Mech. 31, pages 94–100. Vieweg, Wiesbaden, 1991.
5. M. Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen.* Teubner Skripten zur Numerik. Teubner, Stuttgart, 1994.
6. M. Griebel. Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences. *Computing*, 1999. to appear.
7. M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70:163–180, 1995.
8. M. Griebel and T. Schiekofer. An adaptive sparse grid Navier–Stokes solver in 3D based on finite differences. In *Proceedings ENUMATH97*, Heidelberg, 1997.
9. M. Griebel, M. Schneider, and C. Zenger. A combination technique for the solution of sparse grid problems. In R. Beauwens and P. de Groen, editors, *Iterative Methods in Linear Algebra*, pages 263–281. Elsevier, IMACS, Amsterdam, 1992.
10. P. W. Hemker. Sparse–grid finite–volume multigrid for $3D$–problems. *Adv. Comput. Math.*, 4:83–110, 1995.
11. P. W. Hemker. Some remarks on the solution of a finite difference discretization on sparse grids. in preparation, 1999.
12. P. W. Hemker and P. de Zeeuw. BASIS3, a data structure for 3-dimensional sparse grids. In H. Deconinck, editor, *Euler and Navier-Stokes Solvers Using Multi-dimensional Upwinds Schemes and Multigrid Acceleration*, Notes Numer. Fluid Mech. 57, pages 443–484. Vieweg, Wiesbaden, 1997.

13. P. W. Hemker and C. Pflaum. Approximation on partially ordered sets of regular grids. *Appl. Numer. Math.*, 25:55–87, 1997.

14. N. M. Korobov. Approximate calculation of multiple integrals with the aid of methods in the theory of numbers. *Dokl. Akad. Nauk SSSR*, 115:1062–1065, 1957. Russian.

15. J. Noordmans and P. W. Hemker. Convergence results for 3D sparse grid approaches. *Numer. Linear Algebra Appl.*, 5:363–376, 1999.

16. P. Oswald. *Multilevel Finite Element Approximation. Theory and Applications.* Teubner Skripten zur Numerik. Teubner, Stuttgart, 1994.

17. T. Schiekofer. *Die Methode der Finiten Differenzen auf dünnen Gittern zur Lösung elliptischer und parabolischer partieller Differentialgleichungen.* PhD thesis, Universität Bonn, 1998.

18. T. Schiekofer and G. Zumbusch. Software concepts of a sparse grid finite difference code. In W. Hackbusch and G. Wittum, editors, *Concepts of Numerical Software*, Notes in Numerical Fluid Mechanics. Vieweg, Braunschweig, 1998.

19. S. A. Smolyak. Quadrature and interpolation formulas the classes W_s^a and E_s^a. *Dokl. Akad. Nauk SSSR*, 131:1028–1031, 1960. Russian, Engl. Transl.: Soviet Math. Dokl. 1:384–387, 1963.

20. F. Sprengel. *Interpolation and Wavelet Decomposition of Multivariate Periodic Functions.* PhD thesis, University of Rostock, 1997.

21. F. Sprengel. Interpolation and wavelets on sparse Gauß–Chebyshev grids. In W. Haußmann, K. Jetter, and M. Reimer, editors, *Multivariate Approximation, Recent Trends and Results*, Mathematical Research, Vo. 101, pages 269–286. Akademie–Verlag, Berlin, 1997.

22. V. N. Temlyakov. *Approximation of Periodic Functions.* Nova Science, New York, 1993.

23. C. Zenger. Sparse grids. In W. Hackbusch, editor, *Parallel Algorithms for Partial Differential Equations*, Notes on Numerical Fluid Mechanics 31, pages 297–301. Vieweg, Braunschweig, 1991.