Approximation on partially ordered sets of regular grids

P.W. Hemker [a,*], C. Pflaum [b,1]

[a] *Centrum voor Wiskunde en Informatica, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands*
[b] *Institut für Angewandte Mathematik und Statistik, Am Hubland, 97074 Würzburg, Germany*

Abstract

In this paper we analyze the approximation of functions on partially ordered sequences of regular grids. We start with the formulation of minimal requirements for useful grid transfer operators. We introduce the notions of nested and of commutative transfer operators. We define mutual coherence for representations on grids that are not related by coarsening or refining. We show necessary and sufficient conditions for mutual coherence and we show how a hierarchical decomposition is generated by a set of commutative transfer operators. The usual piecewise constant and piecewise d-linear approximations are identified as special instances of tensor product type.

In the second part of the paper we derive error estimates for approximation in these spaces, in different norms on general d-dimensional dyadic sequences of regular and sparse grids. Some of these results have been published before, e.g., in doctoral theses by Bungartz and Pflaum. Here, the results are presented in a unified framework and the proofs are much simplified. We pay special attention to a convenient notation. © 1997 Published by Elsevier Science B.V.

1. Introduction and notation

1.1. Introduction

Recently, in the research on multigrid methods for problems in three dimensions more and more attention is paid to semi-coarsening [5,12,14,17] and sparse grid approaches [2,6–8,13,16]. This can be understood if we notice that the classical multigrid approach, where a linear sequence of nested grids is used for the approximation on different grids, requires very strong relaxation techniques. The selection of a suitable relaxation is difficult because of the large number of possible choices, each with their particular advantages and disadvantages.

---

* Corresponding author. E-mail: p.w.hemker@cwi.nl.
[1] E-mail: pflaum@mathematik.uni-wuerzburg.de.

In semi-coarsening, where more coarser grids are introduced, each coarsened in a single direction, the role of the smoothing procedure is reduced, and simpler relaxation procedures can be applied [10]. This makes it attractive to study partially ordered sets of grids, rather than sequentially ordered ones.

Another difficulty, that particularly arises when regular grids are used for the approximation of functions, is the *curse of three dimensions*: the number of cells increases cubically with each refinement in all directions. This results in enormous amounts of degrees of freedom in the approximation, and in very large systems of algebraic equations to solve. This difficulty can be removed to a large extent by adaptive refinement, i.e., by adding only those degrees of freedom that contribute significantly to the improvement of the accuracy. Of course, what each additional degree of freedom adds to the higher accuracy depends on the choice of basis-functions that span the approximating function space. If, on a regular rectangular grid, a hierarchical basis is chosen, for a sufficiently smooth function the degrees of freedom associated with a "sparse grid" are the optimal choice. The sparse grid can be seen as a combination of regular grids, each with a different cell aspect ratio. In this way semi-coarsening multigrid and sparse grid approximations are much related and make an interesting match.

It appears that the relations between the approximations on the different grids in the partially ordered set are not always clear [4], and that the requirements for the prolongations and restrictions between the approximations on the different grids are often chosen in an *ad hoc* way. In the present paper we study the approximation of functions on partially ordered sets of regular grids (on a *grids of grids*). In particular we are interested in the minimal requirements that are needed to introduce the necessary grid transfer operators. Analyzing these requirements results naturally in the introduction of a hierarchical decomposition of the approximation on the grid of grids, and we are able to show how the usual approximations by piecewise constant and piecewise linear basis functions appear as a special case of tensor product form. In the next sections we concentrate on the piecewise constant and piecewise linear approximations. We define their construction in a systematic way and we derive error estimates for the approximations in different norms.

So, the purpose of this paper is twofold: (1) we show what are the essential requirements for prolongations and restrictions to be able to prove the useful properties that are used in computational schemes, such as mutual coherence [4], and (2) we show the approximation properties of piecewise constant and multilinear approximations on anisotropic and sparse grids. In the first part (Section 2) we see that the requirement of *nested transfer operators* is useful to define *restrictions between different grids* and to define *mutual coherence*. Furthermore, we see that we need the stronger requirement of *commutativity* in order to construct a *pre-wavelet decomposition* (28). Of course, tensor product spaces give the most interesting and useful examples. In the second part of this paper (Section 3) we derive general error estimates for the simplest cases: piecewise constant and piecewise multilinear approximations.

Studying the approximation on a grid of grids, it appeared that a simple and convenient notation was lacking and that the data structures that are used in practice to realize the related algorithms, are rather complicated. Therefore, in the treatment much attention is given to a convenient notation that can be used in general for the description and analysis of algorithms on a grid of grids [11].

Sparse grids yield a way for obtaining approximations with a high accuracy relative to the number of degrees of freedom (unknowns) used. This was first observed by Smoljak [15] for numerical integration and interpolation with trigonometric functions. A different approach of constructing sparse grids is presented in [16]. This approach uses hierarchical basis functions for interpolation with piecewise multilinear functions. Error estimates in different norms and with different assumptions are found

in [2,13,16]. For obtaining optimal estimates, it is necessary to assume that suitable derivatives of the functions are bounded. In case of singularities these assumptions may not hold. Then, optimal estimates can be obtained on adaptive sparse grids, which can be constructed in a natural way with hierarchical basis functions [2].

1.2. Notation

Let $\boldsymbol{k} \in \mathbb{Z}^d$ be a *multi-integer* in d dimensions, then $\boldsymbol{k} = (k_1, k_2, \ldots, k_d)$, with $k_i \in \mathbb{Z}$ for $i = 1, 2, \ldots, d$. We define *relational operators* between multi-integers by

$$\boldsymbol{k} < \boldsymbol{m} \Leftrightarrow (k_1 < m_1 \text{ and } k_2 < m_2 \text{ and } \ldots \text{ and } k_d < m_d),$$

and analogously we define $\boldsymbol{k} \leqslant \boldsymbol{m}$, $\boldsymbol{k} > \boldsymbol{m}$, $\boldsymbol{k} \geqslant \boldsymbol{m}$ and $\boldsymbol{k} = \boldsymbol{m}$. Further we define

$$\max(\boldsymbol{m}, \boldsymbol{n}) = \big(\max(m_1, n_1), \max(m_2, n_2), \ldots, \max(m_d, n_d)\big),$$

and $\min(\boldsymbol{m}, \boldsymbol{n})$ similarly. In a few instances we will use these operators with the same meaning for real vectors $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$.

With $\boldsymbol{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ we denote $|\boldsymbol{n}| = n_1 + \cdots + n_d$. We also use the notation $\boldsymbol{0} = (0, \ldots, 0) \in \mathbb{N}^d$; $2^{\boldsymbol{n}} = (2^{n_1}, \ldots, 2^{n_d})$; $2^{\boldsymbol{n}}\boldsymbol{x} = (2^{n_1}x_1, \ldots, 2^{n_d}x_d)$; $\boldsymbol{n} \circ \boldsymbol{m} = \sum_{i=1,\ldots,d} n_i m_i$, and $\|\boldsymbol{n}\| = n_1 \cdots n_d$. Further we introduce in $\mathbb{Z}^d$ the unit vectors $\boldsymbol{e}_k$, $k = 1, \ldots, d$, as follows: $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$; $\boldsymbol{e}_2 = (0, 1, 0, \ldots, 0)$; $\boldsymbol{e}_d = (0, \ldots, 0, 1)$, and we use $\boldsymbol{e} = (1, \ldots, 1)$. Finally we define $E = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}$.

Let either $\Omega = \mathbb{R}^d$ be the d-dimensional Euclidean space, or let $\Omega = (0, 1)^d \subset \mathbb{R}^d$ be the d-dimensional open unit cube. With any multi-integer $\boldsymbol{n} \in \mathbb{Z}^d$ we associate a function space $V_{\boldsymbol{n}}$, e.g., the space of piecewise constant or piecewise linear (bi-linear, tri-linear, d-linear) functions on a uniform grid with mesh size $\boldsymbol{h} = (h_1, \ldots, h_d) = (2^{-n_1}, \ldots, 2^{-n_d})$. These grids are uniformly spaced in each of the d coordinate directions, but possibly with a different mesh size in the different directions. The volume of these cells is denoted by $\|\boldsymbol{h}\| = 2^{-|\boldsymbol{n}|}$. The functions in $V_{\boldsymbol{n}}$ all are constant or d-linear on each *dyadic block* or *cell*

$$\Omega_{\boldsymbol{n},\boldsymbol{k}} = \big[k_1 2^{-n_1}, (k_1 + 1)2^{-n_1}\big] \times \cdots \times \big[k_d 2^{-n_d}, (k_d + 1)2^{-n_d}\big],$$

and this family of cells forms the *grid*

$$\Omega_{\boldsymbol{n}} = \big\{\Omega_{\boldsymbol{n},\boldsymbol{k}} \mid \Omega_{\boldsymbol{n},\boldsymbol{k}} \subset \overline{\Omega}, \ \boldsymbol{k} \in \mathbb{Z}^d\big\}.$$

The family of *cell centers* or *cell nodes* is denoted by

$$\Omega_{\boldsymbol{n}}^* = \big\{z_{\boldsymbol{n},\boldsymbol{k}} \mid z_{\boldsymbol{n},\boldsymbol{k}} = (\boldsymbol{k} + \boldsymbol{e}/2)2^{-\boldsymbol{n}}; \ \boldsymbol{k} \in \mathbb{Z}^d; \ z_{\boldsymbol{n},\boldsymbol{k}} \in \Omega\big\}.$$

Other grids are obtained by considering the *cell vertices* or *vertex nodes* of the cells in $\Omega_{\boldsymbol{n}}$ as a grid of points. We denote these grids by $\Omega_{\boldsymbol{n}}^+$.

Apparently, all grids are identified by a multi-integer $\boldsymbol{n}$; the number $|\boldsymbol{n}|$ is called the *level* of the grid $\boldsymbol{n}$. Notice that—different from classical multigrid theory—we make a clear distinction between the grid-identification $\boldsymbol{n}$ and the level number $|\boldsymbol{n}|$.

We also use the following notation for partial derivatives, with $\boldsymbol{n} \in \mathbb{N}_0^d$,

$$D^{\boldsymbol{n}} = D^{n_1, \ldots, n_d} = \left(\frac{\partial}{\partial x_1}\right)^{n_1} \cdots \left(\frac{\partial}{\partial x_d}\right)^{n_d}.$$

For the Banach spaces of continuously differentiable functions we use, with $\boldsymbol{n} \in \mathbb{N}_0^d$, the notation $C^{\boldsymbol{n}}(\Omega)$ for the space of functions with finite norm

$$\|u\|_{C^{\boldsymbol{n}}} = \max_{\boldsymbol{0} \leqslant \boldsymbol{m} \leqslant \boldsymbol{n}} \max_{\boldsymbol{x} \in \Omega} |D^{\boldsymbol{m}} u(\boldsymbol{x})|.$$

For $l \in \mathbb{N}_0$, we introduce the notation $C^{\boldsymbol{n},l}(\Omega) = \bigcap_{|\boldsymbol{m}|=l} C^{\boldsymbol{n}+\boldsymbol{m}}(\Omega)$. This is a generalization which combines $C^{\boldsymbol{n}}(\Omega) = C^{\boldsymbol{n},0}(\Omega)$ with the usual space of l times continuously differentiable functions $C^l(\Omega) = C^{\boldsymbol{0},l}(\Omega)$. With a $C_0^{\boldsymbol{n}}(\Omega)$ and $C_0^{\boldsymbol{n},l}(\Omega)$ we denote the corresponding subspaces with homogeneous boundary conditions.

For the Banach spaces of integrable functions, $1 \leqslant p \leqslant \infty$, we, similarly, use the notation $W_p^{\boldsymbol{n}}(\Omega)$ for the space of functions with finite norm

$$\|u\|_{W_p^{\boldsymbol{n}}} = \left(\sum_{\boldsymbol{0} \leqslant \boldsymbol{m} \leqslant \boldsymbol{n}} |u|_{W_p^{\boldsymbol{m}}}^p \right)^{1/p}.$$

Further we use the semi-norm $|u|_{W_p^{\boldsymbol{n}}}$ or, with $0 \leqslant k \leqslant d$, the norm $\|u\|_{W_p^{\boldsymbol{n},k}}$, defined by

$$|u|_{W_p^{\boldsymbol{n}}} = \left(\int_{\boldsymbol{x} \in \Omega} |D^{\boldsymbol{n}} u(x)|^p \right)^{1/p} \quad \text{and} \quad \|u\|_{W_p^{\boldsymbol{n},k}} = \left(\sum_{\boldsymbol{0} \leqslant \boldsymbol{m} \leqslant \boldsymbol{e}, |\boldsymbol{m}|=k} \|D^{\boldsymbol{n}+\boldsymbol{m}} u\|_p^p \right)^{1/p}.$$

For $l \in \mathbb{N}_0$, we write $\mathcal{W}_p^{\boldsymbol{n},l}(\Omega) = \bigcap_{|\boldsymbol{m}|=l} W_p^{\boldsymbol{n}+\boldsymbol{m}}(\Omega)$ and we obtain the Sobolev space $W_p^l(\Omega) = \mathcal{W}_p^{\boldsymbol{0},l}(\Omega)$. Again, with a $W_{p,0}^l(\Omega)$ and $\mathcal{W}_{p,0}^{\boldsymbol{n},l}(\Omega)$ we denote the corresponding subspaces with homogeneous boundary conditions. For $p = \infty$ we use the standard modifications, and for W_2 we also write H. Thus, for the Hilbert spaces of square integrable functions we use the notation $H^{\boldsymbol{n}}(\Omega) = W_2^{\boldsymbol{n}}(\Omega)$, and for the semi-norm and norm $|u|_{H^{\boldsymbol{n}}} = |u|_{W_2^{\boldsymbol{n}}}$ and $\|u\|_{H^{\boldsymbol{n}}} = \|u\|_{W_2^{\boldsymbol{n}}}$. For $l \in \mathbb{N}_0$, we write $\mathcal{H}^{\boldsymbol{n},l}(\Omega) = \bigcap_{|\boldsymbol{m}|=l} H^{\boldsymbol{n}+\boldsymbol{m}}(\Omega)$ and we obtain the usual Sobolev space $H^l(\Omega) = \mathcal{H}^{\boldsymbol{0},l}(\Omega)$. Again, with a $H_0^l(\Omega)$ and $\mathcal{H}_0^{\boldsymbol{n},l}(\Omega)$ we denote the corresponding subspaces with homogeneous boundary conditions.

2. Space decomposition

In this section we introduce *nested* and *commutative* sets of *restrictions* and *prolongations*. We see that the nested property is required to define restrictions between coarser and finer spaces. The property of *commutativity* is needed to check coherence of approximations on different grids, and, moreover, it allows characterization of approximating spaces by a hierarchical decomposition.

2.1. Nested restrictions and prolongations

Let X be a Banach space; e.g., $X = C^0(\Omega)$, $X = L_p(\Omega)$ or $X = L_p^{\text{loc}}(\Omega)$, where $\Omega \subset \mathbb{R}^d$. Let $\boldsymbol{k} \in \mathbb{Z}^d$ and let

$$R_{\boldsymbol{k}} : X \to V_{\boldsymbol{k}} \tag{1}$$

be a *restriction*, i.e., a linear surjection. Possibly $V_k \subset X$, but this is not necessary. We notice that for any such R_k, because of the surjection, there exists the right-inverse or reconstruction

$$P_k : V_k \to X, \tag{2}$$

such that

$$R_k P_k = I_k \tag{3}$$

is the identity operator on V_k. We notice that P_k is an injection (and hence a prolongation) and $\mathrm{Ran}(P_k) \subset X$, but P_k is *not* uniquely determined by a given R_k. In this section we study properties of such sets of *transfer operators* $\{R_k\}_{k \in \mathbb{Z}^d}$ and $\{P_k\}_{k \in \mathbb{Z}^d}$.

It is a consequence of (3) that

$$\Pi_k = P_k R_k$$

is a projection

$$\Pi_k : X \to \mathrm{Ran}(\Pi_k) = \mathrm{Ran}(P_k) \subset X,$$

as is

$$I - \Pi_k : X \to \mathrm{Ker}(\Pi_k) = \mathrm{Ker}(R_k) \subset X,$$

and we observe that X can be written as a direct sum $X = \mathrm{Ran}(P_k) \oplus \mathrm{Ker}(R_k)$.

Definition 2.1. A set $\{R_k\}_{k \in \mathbb{Z}^d}$ is called a *nested set of restrictions* (or *NSR*) iff

$$k \geqslant m \Rightarrow \mathrm{Ker}(R_k) \subset \mathrm{Ker}(R_m). \tag{4}$$

A set $\{P_k\}_{k \in \mathbb{Z}^d}$ is called a *nested set of prolongations* (or *NSP*) iff

$$k \geqslant m \Rightarrow \mathrm{Ran}(P_k) \supset \mathrm{Ran}(P_m). \tag{5}$$

It is obvious that for an NSR $\{R_k\}$ a set of corresponding reconstructions is not necessarily an NSP. On the other hand, given an NSP, the corresponding set of restrictions is not necessarily an NSR. However, in some cases both the restrictions and their reconstructions may form nested sets. Then we say that the *transfer operators are nested* and $\{V_k\}$ forms a nested set of representations of functions in X.

Theorem 2.2. *Let* $\{R_k\}_{k \in \mathbb{Z}^d}$ *be a nested set of restrictions, and* $\{P_k\}_{k \in \mathbb{Z}^d}$ *a set of corresponding reconstructions, then*

$$\forall n \geqslant m \; \exists! \; R_{mn} : V_n \to V_m,$$

with the properties:

(1) R_{mn} *is a restriction*;

(2) $R_{mn} R_n = R_m$; $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (6)

(3) $R_{mn} = R_m P_n$ *(independent of the choice of* P_n*!)*.

Proof. (i) Define $R^1_{mn} = R_m P^1_n$ and $R^2_{mn} = R_m P^2_n$. Then we know that $R_n P^1_n = I_n = R_n P^2_n$ and hence $\forall v_n \in V_n \; R_n(P^1_n - P^2_n)v_n = 0$. Because $\{R_k\}$ is an NSR and $m \leqslant n$ it follows that

$R_m(P_n^1 - P_n^2)v_n = 0$ and hence $R_{mn}^1 = R_{mn}^2$. So that there exists a unique R_{mn}. This means that we can write $R_{mn} = R_m P_n$, and R_{mn} is independent of the choice of P_n.

(ii) $R_{mn}R_n = R_m P_n R_n = R_m \Pi_n = R_m$ on $\mathrm{Ran}(P_n)$. Now, because $X = \mathrm{Ran}(P_n) \oplus \mathrm{Ker}(R_n)$ we may write $\forall v \in X$: $v = v_p + v_n$ so that

$$R_{mn}R_n v = R_{mn}R_n v_p + R_{mn}R_n v_n = R_m v_p + 0 = R_m v_p.$$

Further, because of $\mathrm{Ker}(R_n) \subset \mathrm{Ker}(R_m)$ we see $R_m v = R_m v_p + R_m v_n = R_m v_p + 0 = R_m v_p$, so that $R_{mn}R_n v = R_m v_p = R_m v \ \forall v \in X$ and hence $R_{mn}R_n = R_m$.

(iii) Because R_m is a surjection, and by (6), R_{mn} is necessarily a surjection. Of course, R_{mn} is linear (trivial). Therefore R_{mn} is a restriction. $\square$

Given an NSR $\{R_m\}_{m \in \mathbb{Z}^d}$, and a set of corresponding reconstructions $\{P_m\}_{m \in \mathbb{Z}^d}$, we introduce, for $m \leqslant n$,

$$P_{nm} = R_n P_m : V_m \to V_n. \tag{7}$$

Notice that there are many possible choices of P_m for a give R_m. Of course, some actual properties of P_{nm} may depend on this choice!

Lemma 2.3. P_{nm} *is a right inverse of* R_{nm}:

$$R_{mn}P_{nm} = I_m. \tag{8}$$

Proof. $R_{mn}P_{nm} = R_m P_n R_n P_m = R_m P_m - R_m(I - P_n R_n)P_m = I_m + 0$, because $\mathrm{Ker}(R_n) \subset \mathrm{Ker}(R_m)$. $\square$

Corollary 2.4. *With* $n \geqslant m$:
(1) P_{nm} *is a prolongation (i.e., a linear injection);*
(2) $P_{nm}R_{mn}$ *is a projection* $V_n \to \mathrm{Ran}(P_{nm}) \subset V_n$;
(3) $I_n - P_{nm}R_{mn}$ *is a projection* $V_n \to \mathrm{Ker}(R_{mn}) \subset V_n$;
(4) $V_n = \mathrm{Ran}(P_{nm}) \oplus \mathrm{Ker}(R_{mn})$;
(5) $P_{nm} : V_m \to \mathrm{Ran}(P_{nm}) \subset V_n$ *is a bijection.*

Lemma 2.5. *Let* $\{R_k\}_{k \in \mathbb{Z}^d}$ *be an NSR and let* $k \geqslant n \geqslant l$, *then with a given set of corresponding reconstructions* $\{P_k\}_{k \in \mathbb{Z}^d}$ *we have*

$$P_{kl} = P_{kn}P_{nl}.$$

Proof.

$$P_{kn}P_{nl} = R_k P_n R_n P_l = R_k P_l - R_k(I - P_n R_n)P_l = R_k P_l - R_k(I - \Pi_n)P_l$$
$$= R_k P_l - 0 = R_k P_l,$$

because $\mathrm{Ker}(R_k) \subset \mathrm{Ker}(\Pi_n)$. Hence

$$P_{kl} = R_k P_l = P_{kn}P_{nl}.   \square$$

Lemma 2.6. *Let* $\{R_k\}_{k \in \mathbb{Z}^d}$ *be an NSR with the corresponding* $\{P_k\}_{k \in \mathbb{Z}^d}$ *an NSP, then*

$$m \geqslant l \Rightarrow P_m P_{ml} = P_l.$$

Proof. $P_m P_{ml} = P_m R_m P_l = \Pi_m P_l = P_l$. The last equality holding because $\Pi_m u = u$ for all $u \in \mathrm{Ran}(\Pi_m) = \mathrm{Ran}(P_m) \supset \mathrm{Ran}(P_l)$. $\quad\square$

Lemma 2.7. *If $\{R_k\}_{k \in \mathbb{Z}^d}$ is an NSR, then*

$$m \geqslant l \Rightarrow \Pi_l \Pi_m = \Pi_l.$$

If, in addition, $\{P_k\}_{k \in \mathbb{Z}^d}$ is an NSP, then also

$$m \geqslant l \Rightarrow \Pi_m \Pi_l = \Pi_l.$$

Proof. The first equality follows by

$$\Pi_l \Pi_m = P_l R_l P_m R_m = P_l R_{lm} R_m = P_l R_l = \Pi_l,$$

and the second equality by

$$\Pi_m \Pi_l = P_m R_m P_l R_l = P_m R_m P_m P_{ml} R_{lm} R_m = P_m P_{ml} R_{lm} R_m = P_l R_l = \Pi_l. \quad\square$$

If $\{R_k\}_{k \in \mathbb{Z}^d}$ is an NSR, and if $m \leqslant n$, then a bijection exists between $\mathrm{Ran}(P_m)$ and a subset of $\mathrm{Ran}(P_n)$. We denote this relation by $\mathrm{Ran}(P_m) \preceq \mathrm{Ran}(P_n)$. I.e., a function that can be found in $\mathrm{Ran}(P_m)$, can uniquely be associated with a function in $\mathrm{Ran}(P_n)$. This follows because a bijection exists between $\mathrm{Ran}(P_n)$ and V_n, and between $\mathrm{Ran}(P_m)$ and V_m; and also a bijection exists between V_m and $\mathrm{Ran}(P_{nm}) \subset V_n$.

Hence, given a Banach space X with a nested set of restrictions $\{R_k\}_{k \in \mathbb{Z}^d}$, a family of subspaces $\mathrm{Ran}(P_n)$ exists, with a partial ordering corresponding with the partial ordering of $\{n\}$. This means that, although not necessarily $\mathrm{Ran}(P_m) \subset \mathrm{Ran}(P_n)$, a partial ordering exists such that

$$m \leqslant n \Leftrightarrow \mathrm{Ran}(P_m) \preceq \mathrm{Ran}(P_n) \Leftrightarrow P_{nm} R_m \mathrm{Ran}(P_m) = R_n \mathrm{Ran}(P_m) \subset R_n \mathrm{Ran}(P_n).$$

If $\{P_k\}_{k \in \mathbb{Z}^d}$ is an NSP, this partial ordering simply reduces to $\mathrm{Ran}(P_m) \subset \mathrm{Ran}(P_n)$. Then, in the case that $V_n \subset X$ and we take P_n to be the natural injection, this means that $\mathrm{Ran}(P_n)$ can be identified with V_n and $m \leqslant n \Leftrightarrow V_m \subset V_n$ and, thus, we find $\{V_n\}$ to be a partially ordered family of subsets of X.

Definition 2.8. Functions $f_m \in V_m$ and $f_n \in V_n$ are *mutually coherent*[2] iff

$$\exists f_k \in V_k \quad \text{with } k \geqslant m, \ k \geqslant n,$$
$$\text{such that} \quad f_m = R_{mk} f_k \quad \text{and} \quad f_n = R_{nk} f_k.$$

Theorem 2.9. *Let $\{R_k\}_{k \in \mathbb{Z}^d}$ be an NSR and $\{P_k\}_{k \in \mathbb{Z}^d}$ an NSP. Then, if f_m and f_n are mutually coherent, we have*

$$\forall l \in \mathbb{Z}^d \text{ with } l \leqslant m, \ l \leqslant n, \text{ we have } R_{lm} f_m = R_{ln} f_n. \tag{9}$$

Moreover, under the additional condition that $\Pi_{\min(m,n)} = \Pi_m \Pi_n$, also the reverse holds: it follows from (9) that f_m and f_n are mutually coherent.

---

[2] For the practical use of mutual coherence see, e.g., [4].

Proof. ($\Rightarrow$) First we assume that $f_m \in V_m$ and $f_n \in V_n$ are mutually coherent. Then $\exists f_k \in V_k$ with $k \geqslant m$, $k \geqslant n$ such that $f_m = R_{mk} f_k$, $f_n = R_{nk} f_k$, and hence

$$R_{lm} f_m = R_{lm} R_{mk} f_k = R_{lk} f_k = R_{ln} R_{nk} f_k = R_{ln} f_n,$$

which proves (9).

($\Leftarrow$) Now we assume (9). Let $k = \max(m, n)$ and take $l = \min(m, n)$. We introduce f_k by

$$f_k = P_{km} f_m + P_{kn} f_n - P_{kl} R_{lm} f_m = P_{km} f_m + P_{kn} f_n - P_{kl} R_{ln} f_n.$$

Then

$$\begin{aligned} R_{mk} f_k &= R_{mk} P_{km} f_m + R_{mk} P_{kn} f_n - R_{mk} P_{kl} R_{ln} f_n \\ &= R_{mk} P_{km} f_m + R_{mk} P_{kn} (I_n - P_{nl} R_{ln}) f_n. \end{aligned}$$

Now, because $k \geqslant m$ and $k \geqslant n \geqslant l$, we know $R_{mk} P_{km} = I_m$ (Lemma 2.3).

The additional condition and Lemma 2.7 show $\Pi_m \Pi_n = \Pi_l = \Pi_m \Pi_l \Pi_n$, so that $\Pi_m (I - \Pi_l) \Pi_n = 0$. Hence,

$$\begin{aligned} R_{mk} P_{kn} (I_n - P_{nl} R_{ln}) &= R_m P_k R_k P_n (I_n - R_n P_l R_l P_n) \\ &= R_m P_k R_k (I - P_n R_n P_l R_l) P_n \\ &= R_m \Pi_m \Pi_k (I - \Pi_n \Pi_l) \Pi_n P_n \\ &= R_m \Pi_m (I - \Pi_l) \Pi_n P_n = 0. \end{aligned}$$

Thus, we find $R_{mk} f_k = f_m + 0 = f_m$. Analogously we prove $R_{nk} f_k = f_n$. $\quad\square$

2.2. Commutative restrictions and projections

Definition 2.10. A set $\{R_k\}_{k \in \mathbb{Z}^d}$ is called a *commutative set of restrictions* (or a CSR) iff for all $m, n \in \mathbb{Z}^d$

$$\begin{aligned} &\mathrm{Ker}(R_n) \cap \mathrm{Ker}(R_m) \\ &\qquad\qquad \oplus \\ \mathrm{Ker}(R_{\min(m,n)}) = \ &\mathrm{Ker}(R_n) \cap \mathrm{Ran}(P_m) \ . \\ &\qquad\qquad \oplus \\ &\mathrm{Ran}(P_n) \cap \mathrm{Ker}(R_m) \end{aligned} \qquad (10)$$

A set $\{P_k\}_{k \in \mathbb{Z}^d}$ is called a *commutative set of prolongations* (or a CSP) iff for all $m, n \in \mathbb{Z}^d$

$$\mathrm{Ran}(P_{\min(m,n)}) = \mathrm{Ran}(P_n) \cap \mathrm{Ran}(P_m). \qquad (11)$$

If $\{R_k\}_{k \in \mathbb{Z}^d}$ is a CSR and $\{P_k\}_{k \in \mathbb{Z}^d}$ is a CSP, then we say that we have *commutative transfer operators*.

It is immediate that each CSR is an NSR and each CSP is an NSP: which simply follows from the equivalence $n \leqslant m \Leftrightarrow n = \min(n, m)$. On the other hand not necessarily every NSP is a CSP nor every NSR a CSR.

In the next theorem we show how the above definition of a CSR and an NSP lead to commutative projection operators indeed. To prove the theorem, we derive the following three lemmas.

Lemma 2.11. *If* (12) *holds, then* $\{R_k\}_{k \in \mathbb{Z}^d}$ *is a CSR.*

Proof. Let $m, n \in \mathbb{Z}^d$ be arbitrary, and let $l = \min(m, n)$, then we know by assumption $\Pi_m \Pi_n = \Pi_l = \Pi_n \Pi_m$. We define

$$M = \mathrm{Span}\big((\mathrm{Ker}(R_n) \cap \mathrm{Ker}(R_m)), (\mathrm{Ker}(R_n) \cap \mathrm{Ran}(P_m)), (\mathrm{Ran}(P_n) \cap \mathrm{Ker}(R_m)) \big).$$

To prove the lemma we show:
  (i) $M \subset \mathrm{Ker}(\Pi_l)$,
  (ii) $\mathrm{Ker}(\Pi_l) \subset M$,
  (iii) $M = (\mathrm{Ker}(R_n) \cap \mathrm{Ker}(R_m)) \oplus (\mathrm{Ker}(R_n) \cap \mathrm{Ran}(P_m)) \oplus (\mathrm{Ran}(P_n) \cap \mathrm{Ker}(R_m))$.

To prove (i), let $x \in M$, then $x = x_m + x_n + x_l$ with $x_n \in \mathrm{Ker}(\Pi_n) \cap \mathrm{Ran}(\Pi_m)$, $x_m \in \mathrm{Ker}(\Pi_m) \cap \mathrm{Ran}(\Pi_n)$, $x_l \in \mathrm{Ker}(\Pi_n) \cap \mathrm{Ker}(\Pi_m)$. Then

$$\Pi_l x = \Pi_n \Pi_m x_m + \Pi_m \Pi_n x_n + \Pi_m \Pi_n x_l = \Pi_n 0 + \Pi_m 0 + \Pi_m 0 = 0.$$

So that $x \in \mathrm{Ker}(\Pi_l)$.

To prove (ii), let $x \in \mathrm{Ker}(\Pi_l)$ be arbitrary and define $x_m := \Pi_m(I - \Pi_n)x = \Pi_m x$ and $x_n := \Pi_n(I - \Pi_m)x = \Pi_n x$ and

$$x_0 := (I - \Pi_m - \Pi_n + \Pi_n \Pi_m)x = (I - \Pi_m - \Pi_n)x.$$

Then a simple calculation shows: $x_0 \in \mathrm{Ker}(R_n) \cap \mathrm{Ker}(R_m)$ and $x_m \in \mathrm{Ker}(R_n) \cap \mathrm{Ran}(P_m)$ and $x_n \in \mathrm{Ran}(P_n) \cap \mathrm{Ker}(R_m)$ and $x_0 + x_m + x_n = x$.

To prove (iii), we show that M is a direct sum of the three spanning spaces. For this we have to prove: if $x_0 \in \mathrm{Ker}(R_n) \cap \mathrm{Ker}(R_m)$ and $x_m \in \mathrm{Ker}(R_n) \cap \mathrm{Ran}(P_m)$ and $x_n \in \mathrm{Ran}(P_n) \cap \mathrm{Ker}(R_m)$ and $x_0 + x_m + x_n = 0$, then $x_0 = x_m = x_n = 0$. This is seen by $0 = \Pi_m(I - \Pi_n)(x_0 + x_m + x_n) = x_m$ and $0 = \Pi_n(I - \Pi_m)(x_0 + x_m + x_n) = x_n$. This implies $x_0 = 0$. $\square$

Lemma 2.12. *If* (12) *holds, then* $\{P_k\}_{k \in \mathbb{Z}^d}$ *is an NSP.*

Proof. Let $n < m$ then (12) implies $\Pi_m \Pi_n = \Pi_n$, and hence, for all $x \in X$ we have $\Pi_m \Pi_n x = \Pi_n x$. It follows that for all $x \in X$ holds $\Pi_n x \in \mathrm{Ran}(\Pi_m)$. Therefore $\mathrm{Ran}(\Pi_n) \subset \mathrm{Ran}(\Pi_m)$ for all $m, n \in \mathbb{Z}^d$ with $n < m$. Hence $\{P_k\}_{k \in \mathbb{Z}^d}$ is a NSP. $\square$

Lemma 2.13. *If* $\{R_k\}_{k \in \mathbb{Z}^d}$ *is a CSR and* $\{P_k\}_{k \in \mathbb{Z}^d}$ *is an NSP then* (12) *holds.*

Proof. Let $l = \min(m, n)$ and let $x \in X$ be arbitrary. We know that

$$\begin{aligned}
X &= \mathrm{Ran}(P_l) \oplus \mathrm{Ker}(R_l) \\
&= \mathrm{Ran}(P_l) \oplus \big(\mathrm{Ker}(R_n) \cap \mathrm{Ker}(R_m) \big) \oplus \big(\mathrm{Ran}(P_n) \cap \mathrm{Ker}(R_m) \big) \oplus \big(\mathrm{Ker}(R_n) \cap \mathrm{Ran}(P_m) \big)
\end{aligned}$$

because $\{R_k\}$ is a CSR. Hence, we may split $x = x_l + x_o + x_m + x_n$ accordingly. Now we know $\Pi_l x = x_l$ and

$$\Pi_m \Pi_n x = \Pi_m \Pi_n (x_l + x_o + x_m + x_n)$$
$$= \Pi_m \Pi_n x_l + \Pi_m \Pi_n x_o + \Pi_m \Pi_n x_m + \Pi_m \Pi_n x_n$$
$$= \Pi_m \Pi_n x_l + \Pi_m 0 + \Pi_m x_m + \Pi_m 0$$
$$= \Pi_m \Pi_n x_l.$$

Because $\{P_k\}_{k \in \mathbb{Z}^d}$ is an NSP, we know that $x_l \in \mathrm{Ran}(P_l) \subset \mathrm{Ran}(P_n)$ and $x_l \in \mathrm{Ran}(P_l) \subset \mathrm{Ran}(P_m)$; hence $\Pi_m \Pi_n x_l = x_l$. We conclude that, for arbitrary $x \in X$ holds $\Pi_m \Pi_n x = x_l = \Pi_l x$; which proves the lemma. □

Theorem 2.14. *The two following statements are equivalent:*
 (1) $\{R_k\}_{k \in \mathbb{Z}^d}$ *is a CSR, and* $\{P_k\}_{k \in \mathbb{Z}^d}$ *is an NSP; and*
 (2)

$$\Pi_m \Pi_n = \Pi_{\min(m,n)} \quad \forall m, n \in \mathbb{Z}^d. \tag{12}$$

Proof. The theorem is a direct combination of the three lemmas above. □

It is an immediate consequence of the theorem that operators Π_n associated with commutative transfer operators $\{R_k\}_{k \in \mathbb{Z}^d}$ and $\{P_k\}_{k \in \mathbb{Z}^d}$ do commute:

$$\Pi_m \Pi_n = \Pi_{\min(m,n)} = \Pi_n \Pi_m.$$

Further, combination of Theorem 2.14 with Theorem 2.9 gives the following corollary, which is a direct generalization of [4, Proposition 2.5]. By the present framework we recognize immediately the essential conditions that lead to this result.

Corollary 2.15. *Let* $\{R_k\}_{k \in \mathbb{Z}^d}$ *be a CSR and* $\{P_k\}_{k \in \mathbb{Z}^d}$ *an NSP, then* f_m *and* f_n *are mutually coherent if and only if*

$$\forall l \in \mathbb{Z}^d \text{ with } l \leqslant m, \ l \leqslant n, \text{ we have } R_{lm} f_m = R_{ln} f_n. \tag{13}$$

Having seen what the essential structure of commutative transfer operators is, further in this section we assume all transfer operators to be commutative.

2.3. The merging operator

Now we have seen how information about a function $u \in X$ can be represented on V_n, and how the representations $R_n u$ are related for different $n \in \mathbb{Z}^d$. An important question is how these $R_n u$, given for a limited number of $n \in \mathbb{Z}^d$, can be used to restore the picture of the original function u as complete as possible.
 We start with the situation where information is available from two representations, viz. in V_n, V_m. Therefore we introduce the *merging* operator Π_{mn}, which selects for an $x \in X$ the information that can be represented by the combined representations in V_n and V_m.

Definition 2.16. The *merging operator* $\Pi_{mn} : X \to X$ is defined by

$$\Pi_{mn} = \Pi_m + \Pi_n - \Pi_{\min(m,n)}. \tag{14}$$

Lemma 2.17. *Let* $\{R_n\}$ *and* $\{P_n\}$ *be a set of commutative transfer operators, then:*

(1) $\Pi_{mn} = \Pi_{nm},$ (15)

(2) *if* $m \leqslant n$ *then* $\Pi_{mn} = \Pi_n,$ (16)

(3) Π_{mn} *is a projection,*   (17)

$$\text{Ran}(P_n) \cap \text{Ran}(P_m)$$
$$\oplus$$
(4) $\text{Ran}(\Pi_{mn}) = \text{Ran}(P_n) \cap \text{Ker}(R_m)$ (18)
$$\oplus$$
$$\text{Ker}(R_n) \cap \text{Ran}(P_m)$$

(5) $\text{Ker}(\Pi_{mn}) = \text{Ker}(R_n) \cap \text{Ker}(R_m).$ (19)

Proof. The first two statements are trivial by the definition of Π_{mn}. Now set $l = \min(m,n)$. Because $\Pi_m \Pi_n = \Pi_n \Pi_m = \Pi_{\min(m,n)}$ we have

$$\Pi_{mn}\Pi_{mn} = (\Pi_m + \Pi_n - \Pi_{\min(m,n)})(\Pi_m + \Pi_n - \Pi_{\min(m,n)})$$
$$= \Pi_m + \Pi_l - \Pi_l + \Pi_l + \Pi_n - \Pi_l - \Pi_l - \Pi_l + \Pi_l$$
$$= \Pi_m + \Pi_n - \Pi_l = \Pi_{mn}.$$

Hence, Π_{mn} is a projection.

$x \in \text{Ran}(\Pi_{mn})$ implies $\exists z$:

$$x = \Pi_{mn}z = \Pi_m z + \Pi_n z - \Pi_l z = (\Pi_m - \Pi_l)z + (\Pi_n - \Pi_l)z + \Pi_l z = z_m + z_n + z_l.$$

It is clear that this is a decomposition according to the direct sum in (18): $z_l \in \text{Ran}(\Pi_l) \subset \text{Ran}(\Pi_m) \cap \text{Ran}(\Pi_n)$; $z_m \in \text{Ran}(\Pi_m)$ and $z_m \in \text{Ker}(\Pi_n)$ because

$$\Pi_n z_m = \Pi_n(\Pi_m - \Pi_l) = \Pi_{\min(n,m)} - \Pi_l z = 0;$$

and similarly $z_n \in \text{Ran}(\Pi_n) \cap \text{Ker}(\Pi_m)$.

On the other hand, if $z = z_l + z_n + z_m$ is a splitting according to the direct sum, then

$$\Pi_l z = \Pi_l(z_l + z_m + z_n)$$
$$= \Pi_l z_l + \Pi_n \Pi_m z_m + \Pi_m \Pi_n z_n$$
$$= z_l + \Pi_n z_m + \Pi_m z_n = z_l,$$

$$(\Pi_m - \Pi_l)z = (\Pi_m - \Pi_l)(z_l + z_m + z_n)$$
$$= (\Pi_m - \Pi_l)z_l + (\Pi_m - \Pi_l)z_m + (\Pi_m - \Pi_l)z_n$$
$$= z_l - z_l + \Pi_m(I - \Pi_n)z_m + \Pi_m(I - \Pi_n)z_n$$
$$= \Pi_m z_m = z_m.$$

Analogously $(\Pi_n - \Pi_l)z = z_n$. Thus

$$z = z_l + z_m + z_n = \Pi_l z + (\Pi_m - \Pi_l)z + (\Pi_n - \Pi_l)z = (\Pi_m + \Pi_n - \Pi_l)z = \Pi_{mn}z.$$

Hence $z \in \text{Ran}(\Pi_{mn})$.

Assume that $0 = z = z_l + z_n + z_m$ is a splitting as above, then it follows that $0 = z_l = z_n = z_m$ because

$$0 = \Pi_l(z) = (\Pi_m \Pi_n)(z_l + z_n + z_m) = z_l,$$
$$0 = (\Pi_m - \Pi_l)(z) = (I - \Pi_n)\Pi_m(z_l + z_n + z_m) = z_m$$

and

$$0 = (\Pi_n - \Pi_l)(z) = (I - \Pi_m)\Pi_n(z_l + z_n + z_m) = z_n.$$

$\Pi_{mn} = P_m R_m + P_n R_n - P_l R_l$, hence $\text{Ker}(\Pi_{mn}) \supset \text{Ker}(R_n) \cap \text{Ker}(R_m)$ is trivial. $\text{Ker}(\Pi_{mn}) \subset \text{Ker}(R_n) \cap \text{Ker}(R_m)$ is shown as follows. Let $x \in \text{Ker}(\Pi_{nm})$, then

$$0 = \Pi_{nm} x = (\Pi_n - \Pi_l)x + (\Pi_m - \Pi_l)x + \Pi_l x =: x_n + x_m + x_l.$$

This implies $x_l \in \text{Ran}(\Pi_l)$ and $x_n, x_m \in \text{Ker}(\Pi_l)$, so that $x_l = 0$ and $x_n + x_m = 0$. Further, from

$$x_n = -x_m \in \text{Ran}(\Pi_m) \cap \text{Ran}(\Pi_n) = \text{Ran}(\Pi_{\min(n,m)}) = \text{Ran}(P_l)$$

it follows that $x_m = x_n = 0$, and hence $\text{Ker}(\Pi_{mn}) = \text{Ker}(R_n) \cap \text{Ker}(R_m)$.   □

Now we can introduce the hierarchical surplus, $H_{mn} u$, of a function u. This hierarchical surplus represents the amount of information in an approximation $u \in V_{\max(m,n)}$ that cannot be represented on the $\text{Span}(V_n, V_m)$.

Definition 2.18. Let $\{R_n\}$ and $\{P_n\}$ be a commutative set of transfer operators, then we define $H_{mn}: X \to \text{Ran}(P_{\max(m,n)}) \subset X$, the *hierarchical surplus, relative to* the grids m and n, by

$$H_{mn} = \Pi_{\max(m,n)} - \Pi_{mn}. \tag{20}$$

Clearly, the hierarchical surplus is a projection operator, and we can write

$$H_{mn} = (\Pi_{\max(n,m)} - \Pi_n)(\Pi_{\max(n,m)} - \Pi_m).$$

This general idea of a hierarchical surplus leads naturally to a partitioning of the spaces V_n in more elementary subspaces (pre-wavelet spaces) in the following section.

2.4. The hierarchical decomposition

In this section, again, we assume $\{R_n\}$ and $\{P_n\}$ to be commutative sets of transfer operators.

Definition 2.19. For a fixed $n_0 \in \{-\infty, \mathbb{Z}\}^d$, which indicates a coarsest grid, we define for arbitrary $n \geqslant n_0$, $n \in \mathbb{Z}^d$, the operator $Q_n: X \to \text{Ran}(P_n) \subset X$, by

$$Q_n = \prod_{j=1,\ n_j \neq n_{0j}}^{d} (\Pi_n - \Pi_{n-e_j}). \tag{21}$$

We use the convention that $Q_{n_0} = \Pi_{n_0}$. The operator Q_n is called the *(direct) hierarchical surplus* at grid n.

If $n_0 \in \mathbb{Z}^d$ we call $n_0 \in \mathbb{Z}^d$ the *coarsest grid*, and, without loss of generality we may assume $n_0 = 0$. If a $n_{0j} = -\infty$ then no coarsest grid exists in the jth direction.

Lemma 2.20. Q_n *is a projection and* $Q_m Q_n = 0$ *for all* $m \neq n$.

Proof. First we show that Q_n is a projection. For simplicity of notation, we set $c = n_0$.

$$
\begin{aligned}
Q_n Q_n &= \prod_{j=1, n_j \neq c_j}^{d} (\Pi_n - \Pi_{n-e_j})(\Pi_n - \Pi_{n-e_j}) \\
&= \prod_{j=1, n_j \neq c_j}^{d} (\Pi_n - \Pi_{n-e_j} - \Pi_{n-e_j} + \Pi_{n-e_j}) = \prod_{j=1, n_j \neq c_j}^{d} (\Pi_n - \Pi_{n-e_j}) = Q_n. \quad (22)
\end{aligned}
$$

To show that $Q_m Q_n = 0$ for all $m \neq n$, let $m \neq n$. Without loss of generality we may assume $n_i < m_i$ for some $i \in \{1, \ldots, d\}$. We consider the case $n_i \neq c_i$; the other cases are similar.

$$
\begin{aligned}
Q_m Q_n &= \prod_{j} (\Pi_m - \Pi_{m-e_j})(\Pi_n - \Pi_{n-e_j}) \\
&= (\Pi_m - \Pi_{m-e_i})(\Pi_n - \Pi_{n-e_i}) \prod_{j \neq i} \cdots \\
&= (\Pi_{(\ldots, n_l, \ldots)} - \Pi_{(\ldots, n_l, \ldots)} - \Pi_{(\ldots, n_l-1, \ldots)} + \Pi_{(\ldots, n_l-1, \ldots)}) \prod_{j \neq i} \cdots = 0. \quad (23)
\end{aligned}
$$

The indices indicated by dots correspond with those of $\min(n, m)$. □

Notice that, for $n > n_0$, the two-dimensional case the relation (21) reads

$$
Q_n u = \Pi_n u - \Pi_{n-e_1} u - \Pi_{n-e_2} u + \Pi_{n-e} u, \quad (24)
$$

where $e = (1, 1)$, and in the one-dimensional case we have

$$
Q_n u = \Pi_n u - \Pi_{n-e} u. \quad (25)
$$

Corollary 2.21. *From Definition 2.19 it is immediately obvious that the projection* Π_n *can be decomposed as*

$$
\Pi_n = \sum_{n_0 \leqslant m \leqslant n} Q_m, \quad (26)
$$

and, hence, $\mathrm{Ran}(\Pi_n) = \mathrm{Span}(\mathrm{Ran}(Q_m))_{n_0 \leqslant m \leqslant n}$. *Because of Lemma 2.20 we can write*

$$
\mathrm{Ran}(\Pi_n) = \bigoplus_{n_0 \leqslant m \leqslant n} \mathrm{Ran}(Q_m). \quad (27)
$$

If $V_n \subset X$ and P_n is the natural injection, then we see that $\mathrm{Ran}(P_n) = \mathrm{Ran}(\Pi_n) = V_n$ and, defining the *pre-wavelet space* $W_n = \mathrm{Ran}(Q_m)$, we find

$$
V_n = \bigoplus_{n_0 \leqslant m \leqslant n} W_m. \quad (28)
$$

Notice that, given the set $\{W_m\}_{n_0 \leqslant m}$, the expansion (28) essentially describes the complete partially ordered set of approximating spaces. In fact, if we can write $X = \bigoplus_{n_0 \leqslant m} W_m$, we induce automatically a structure for prolongations and restrictions, which—now by nature—form a commutative and hence nested sequence.

2.5. The tensor product case

In this section we introduce the usual and by far the most interesting examples of commutative sets of operators, constructed by means of tensor products. Of course it is also possible to construct—by trivial means—cases where (28) is no tensor product space, e.g., by taking arbitrary linearly independent spaces W_m. Although such non-tensor product spaces maybe useful in particular cases, tensor product spaces give more regular and interesting examples.

For any $i = 1, \ldots, d$, let $\Omega_i \subset \mathbb{R}$ and let $\Omega = \bigotimes_{i=1}^{d} \Omega_i \subset \mathbb{R}^d$ be their Cartesian product. Let $X_i(\Omega_i)$ be a function space on Ω_i with functions $u_{i,a_i}(x_i)$ so that $X_i(\Omega_i) = \text{Span}(\{u_{i,a_i}\}_{a_i \in A_i})$, for some index set A_i. Let $A = \bigotimes_{i=1}^{d} A_i$ be the Cartesian product of index sets. Then the tensor product space $X(\Omega)$ is defined by

$$X(\Omega) = \bigotimes_{i=1}^{d} X_i(\Omega_i) = \underset{(a_1, a_2, \ldots, a_d) \in A}{\text{Span}} \prod_{i=1}^{d} u_{i,a_i}. \tag{29}$$

It is well known, e.g., that for $X_i(\Omega_i) = C_0^\infty(\mathbb{R})$, the tensor product space $X(\Omega)$ is densely embedded both in $C^l(\mathbb{R}^d)$ and in $H^l(\mathbb{R}^d)$.

For each $i = 1, \ldots, d$, let $\{R_{i,n}\}_{n \in \mathbb{Z}}$ be a sequence of restrictions (for functions in one dimension) defined on $X_i(\Omega_i)$, with $R_{i,n}: X_i(\Omega_i) \to V_{i,n}(\Omega_i) \subset X_i(\Omega_i)$, then, similar to $X(\Omega)$, for each $\boldsymbol{n} \in \mathbb{Z}^d$ we may define a tensor product space $V_{\boldsymbol{n}}(\Omega) = \bigotimes_{i=1}^{d} V_{i,n_i}(\Omega_i)$.

If the elements of $V_{i,n_i}(\Omega_i)$ are all determined by values associated with Ω_{n_i} ($\Omega_{n_i}^+$ or $\Omega_{n_i}^*$), then we find a bijection $V_{\boldsymbol{n}} = V_{\boldsymbol{n}}(\Omega) \cong V_{\boldsymbol{n}}(\Omega_{\boldsymbol{n}})$ (or $\cong V_{\boldsymbol{n}}(\Omega_{\boldsymbol{n}}^+)$ or $\cong V_{\boldsymbol{n}}(\Omega_{\boldsymbol{n}}^*)$). This notation indicates that its elements are determined by their values on the Cartesian product space $\Omega_{\boldsymbol{n}} = \bigotimes_{i=1}^{d} \Omega_{n_i}$ (or $\Omega_{\boldsymbol{n}}^+$ or $\Omega_{\boldsymbol{n}}^*$).

Definition 2.22. We define the *tensor product restriction* $R_{\boldsymbol{n}}: X(\Omega) \to V_{\boldsymbol{n}}(\Omega)$ by its action on a typical basis function

$$u_{\boldsymbol{a}}(\boldsymbol{x}) = \prod_{i=1}^{d} u_{i,a_i}(x_i) \mapsto R_{\boldsymbol{n}} u_{\boldsymbol{a}}(\boldsymbol{x}) = \prod_{i=1}^{d} R_{i,n_i} u_{i,a_i}(x_i). \tag{30}$$

We also write $R_{\boldsymbol{n}} = \bigotimes_i R_{i,n_i}$. Since $V_{\boldsymbol{n}}(\Omega) \subset X(\Omega)$, we can take the natural injection $P_{\boldsymbol{n}}$ as the corresponding reconstruction for $R_{\boldsymbol{n}}$. This $P_{\boldsymbol{n}}$ we call the *tensor product prolongation*.

Theorem 2.23. *For each $i \in \{1, \ldots, d\}$ let $\Omega_i \subset \mathbb{R}$, and let $\{V_{i,n}(\Omega_i)\}_{n \in \mathbb{Z}}$ form a nested sequence of subspaces of the function space $X(\Omega_i)$, with*

$$V_{i,p}(\Omega_i) \subset V_{i,q}(\Omega_i) \subset X(\Omega_i) \quad \text{for } p \leqslant q, \tag{31}$$

and let each sequence of (one-dimensional) operators $\{R_{i,n}\}_{n \in \mathbb{Z}}$ form an NSR. If we take for the corresponding reconstructions $\{P_{i,n}\}_{n \in \mathbb{Z}}$ the natural injection, then the tensor product restrictions $\{R_{\boldsymbol{n}}\}$

and prolongations $\{P_n\}$ form a CSR and a CSP, respectively. (So, together they form a commutative set of transfer operators.)

Proof. By the nesting of the subspaces it is immediate that for each i the set $\{P_{i,n}\}_{n\in\mathbb{Z}}$ forms an NSP (Section 2.1). Further, using the fact that the prolongation is the natural injection, we can identify restrictions R_i with the corresponding projections Π_i, and by Lemma 2.7 we see that for each i we have $\Pi_{i,p}\Pi_{i,q} = \Pi_{i,q}\Pi_{i,p} = \Pi_{i,p}$ if $p \leqslant q$. Thus, for each i it follows that the sequences $\{P_{i,n}\}_{n\in\mathbb{Z}}$ and $\{R_{i,n}\}_{n\in\mathbb{Z}}$ are CSP and SCR. Now we prove that $\{R_n\}$ is an CSR by showing $R_m R_n = R_{\min(m,n)}$ as follows.

Let $l = \min(m,n)$ and let $v \in X(\Omega)$ be arbitrary, then we can write

$$v(x) = \sum_{a\in A} c_a u_a(x),$$

$$R_m v(x) = \sum_a c_a R_m u_a(x) = \sum_a c_a \prod_i R_{i,m_i} u_{i,a_i}(x_i).$$

Similarly,

$$R_n R_m v(x) = \sum_a c_a R_n \prod_i R_{i,m_i} u_{i,a_i}(x_i) = \sum_a c_a \prod_i R_{i,n_i} R_{i,m_i} u_{i,a_i}(x_i)$$

$$= \sum_a c_a \prod_i R_{i,\min(m_i,n_i)} u_{i,a_i}(x_i) = \sum_a c_a R_{\min(m,n)} \prod_i u_{i,a_i}(x_i)$$

$$= R_{\min(m,n)} \sum_a c_a \prod_i u_{i,a_i}(x_i) = R_{\min(m,n)} v(x).$$

Hence $R_n R_m = R_{\min(m,n)}$, and, thus, $\Pi_n \Pi_m = \Pi_{\min(m,n)}$. Now Theorem 2.14 shows that $\{R_n\}$ is an CSR and $\{P_n\}$ is an NSP.

To prove that $\{P_n\}$ is an CSP, we have to show $V_{\min(m,n)} = V_m \cap V_n$, or

$$\bigotimes_i V_{i,\min(m_i,n_i)} = \bigotimes_i V_{i,m_i} \cap \bigotimes_i V_{i,n_i}. \tag{32}$$

As for each i we know that $\{V_{i,n}(\Omega_i)\}_{n\in\mathbb{Z}}$ is a nested sequence of subspaces of $X_i(\Omega_i)$, we can construct a sequentially ordered set of basis functions $B_i = \{u_{i,b}\}_b$ in $X_i(\Omega_i)$, such that $n_i < m_i$ implies $u_{i,b} \in V_{i,n_i} \Rightarrow u_{i,b} \in V_{i,m_i}$. It follows that we have $V_{i,m_i} = \text{Span}(\{u_{i,l} \in B_i \mid u_{i,l} \in V_{i,m_i}\})$ and similarly

$$V_{i,\min(n_i,m_i)} = \text{Span}(\{u_{i,l} \in B_i \mid u_{i,l} \in V_{i,n_i} \cap V_{i,m_i}\}).$$

So we see

$$\underset{u_{i,l_i}\in B_i\cap V_{i,\min(m_i,n_i)}}{\text{Span}} \prod_i u_{i,l_i} = \underset{u_{i,m_i}\in B_i\cap V_{i,m_i}}{\text{Span}} \prod_i u_{i,m_i} \cap \underset{u_{i,n_i}\in B_i\cap V_{i,n_i}}{\text{Span}} \prod_i u_{i,n_i},$$

which is equivalent with (32). □

Example 2.24 (Piecewise constant approximation). If we consider $L_2^{\text{loc}}(\Omega) = \overline{X(\Omega)}$ and we choose for $R_{i,n}$ the one-dimensional L_2-projection $R_{i,n}: X(\Omega_i) \to V_n(\Omega_i) \subset X(\Omega_i)$, where $\Omega_i \subset \mathbb{R}$ and

$V_n(\Omega_i)$ is the space of piecewise constant functions on dyadic intervals (i.e., if $R_{i,n}$ denotes taking mean values over intervals $[j2^{-n}, (j+1)2^{-n}]$ in the ith coordinate direction), then, for each i, the restrictions $\{R_{i,n}\}_n$ form a one-dimensional NSR. The corresponding reconstructions $P_{i,n}$ represent piecewise constant interpolation over dyadic intervals. This makes the prolongations $\{P_{i,n}\}_n$ an NSP.

Then, as a consequence of the above theorem, $\{R_k\}$ and $\{P_k\}$ are commutative transfer operators, i.e., $\{R_k\}$ is a CSR and $\{P_k\}$ a CSP.

Example 2.25 (Piecewise d-linear approximation). If we select the restriction $R_{i,n} : C^0(\Omega_i) \longrightarrow V_n(\Omega_i) \subset \mathbb{R}^{\mathbb{Z}}$ to be taking function values at dyadic points $j2^{-n}$ in the interval $\Omega_i \subset \mathbb{R}$, then $\{R_{i,n}\}_n$ is an NSR. Corresponding reconstructions $P_{i,n}$, defined by piecewise linear interpolation over dyadic intervals, make the prolongations $\{P_{i,n}\}_n$ an NSP.

As a consequence of the previous theorem, with $C^0(\Omega) = \overline{X(\Omega)}$, the tensor product operators $\{R_k\}$ and $\{P_k\}$, defined on $X(\Omega)$, are commutative (are a CSR and a CSP respectively). The restriction R_k takes the function values at grid points Ω_n^+, and the prolongation P_k makes a multi-linear interpolation over cells in Ω_n.

In the above examples, with $V_n \subset X$ we took for the reconstruction P_n the natural injection (the identity in X). In this way we may identify R_n and Π_n. It appears that in both cases, i.e., for the piecewise constant and the piecewise linear approximation, we have a projection Π_n of the form

$$\Pi_n = \prod_{j=1}^{d} \Pi_{n_j e_j} = \prod_{j=1}^{d} R_{n_j e_j}.$$

Here $R_{n_j e_j} : X(\Omega) \to X(\Omega)$ is the operator on the tensor product space $X(\Omega)$ such that

$$R_{n_j e_j} u_a(x) = R_{j,n_j} u_{j,a_j}(x_j) \cdot \prod_{i \neq j}^{d} u_{i,a_i}(x_i).$$

In the following section we consider the case of nested subspaces $\{V_n\}$ with $V_n \subset X$ and $X = \bigcup_n V_n$, and where all spaces V_n are spanned by dilations of a single function $\phi(x)$, together with all its dyadic translations. This leads to the more-dimensional *multiresolution analysis* or *MRA*. In this case the spaces $W_m = \text{Ran}(Q_m)$ correspond with more-dimensional wavelet spaces.

2.6. More-dimensional MRA and wavelets

It will be convenient if
  (i) we can make an arbitrarily accurate approximation of any function $u \in X$ by taking the multi-integer n large enough.
Moreover, it will be convenient if
  (ii) all spaces $\{\text{Ran}(P_n)\}$ or $\{V_n\}$ have a similar structure, and
  (iii) there is a clear relation between the spaces in $\{\text{Ran}(P_n)\}$ or $\{V_n\}$.

In order to create such a structure, in this section we introduce the multidimensional multiresolution analysis. For this purpose we will restrict ourselves to Hilbert spaces. First we introduce the important notion of frame.

Definition 2.26. A sequence $\{x_n\}$ in a Hilbert space H is a *frame* if there exist numbers $A, B > 0$ such that for all $x \in H$ we have

$$A\|x\|^2 \leqslant \sum_n |(x, x_n)|^2 \leqslant B\|x\|^2. \tag{33}$$

The numbers A, B are called *frame bounds*. The frame is *tight* if $A = B$. The frame is *exact* if it ceases to be a frame whenever any single element is deleted from the sequence. If the sequence $\{x_n\}$ satisfies (only) the second part of the inequality (33) then the sequence is called a *Bessel sequence*.

Having introduced the exact frame, we can define the partially ordered, more-dimensional multiresolution analysis. Notice that this is different from the more-dimensional multiresolution analysis introduced in [1], which considers a sequentially ordered nested set of approximating spaces.

Definition 2.27. Let $\Omega = \mathbb{R}^d$ and let $X(\Omega)$ be a Hilbert space of functions defined on Ω. A *multidimensional multiresolution analysis* of $X(\Omega)$, is a *partially ordered* set of *closed* linear subspaces

$$\left\{ V_n \mid V_n \subset X(\Omega) \right\}_{n \in \mathbb{Z}^d}$$

with the four properties:

$$\bigcap_n V_n = \{0\}; \qquad \overline{\bigcup_n V_n} = X(\Omega), \tag{34a}$$

$$f(x) \in V_n \Leftrightarrow f(2^m x) \in V_{n+m} \quad \forall n, m \in \mathbb{Z}^d, \tag{34b}$$

$$f(x) \in V_n \Leftrightarrow f(x - 2^{-n} k) \in V_n \quad \forall n, k \in \mathbb{Z}^d, \tag{34c}$$

$$\exists \phi \in V_0: \quad \left\{ \phi(x - k) \right\}_{k \in \mathbb{Z}^d} \text{ is an exact frame for } V_0. \tag{34d}$$

The function $\phi(x)$ in (34d) is called the *father function* or the *scaling function* of the multiresolution analysis.

For $\Omega = \mathbb{R}^d$ the tensor product Examples 2.24 and 2.25 in Section 2.5 also yield examples of a multidimensional MRA.

For piecewise constant interpolation we take $X(\Omega) = L^2(\mathbb{R}^d)$ as the starting point. The characteristic function on the unit cube (the more-dimensional Haar function) is the scaling function ϕ. The set $\{V_n\}$ contains the spaces of piecewise constant functions on Ω_n, and a CSR is obtained by $R_n : X(\Omega) \to V_n$, the L^2-projection. It is obvious that in this case the set $\{\phi(x - k)\}_k$ is an orthonormal basis and hence an exact frame with bounds $A = B = 1$.

For piecewise linear interpolation we take $X(\Omega) = H^e(\mathbb{R}^d)$ as the Hilbert space. The set $\{V_n\}$ contains the space of piecewise d-linear functions, determined by their nodal values at Ω_n^+. A CSR is obtained by $R_n : X(\Omega) \to V_n$, the piecewise d-linear interpolation at Ω_n^+. Here, the d-linear finite-element basis function is the scaling function ϕ. By [9, Theorem 2.1.3] it is easily seen that in this case $\{\phi(x - k)\}_k$ is an exact frame, with as frame bounds the extreme eigenvalues of the frame

operator $S: L^2(\Omega) \to L^2(\Omega)$ defined by $Su = \sum_n (u, \phi(\boldsymbol{x} - \boldsymbol{k})) \phi(\boldsymbol{x} - \boldsymbol{k})$. Bounds for these extreme eigenvalues are $A = 3^{-d}$ and $B = 1$, respectively.

As in the tensor product case, we take for the reconstruction the natural injection $P_n: V_n \to X$ so that $R_n \equiv \Pi_n$ for all $n \in \mathbb{Z}^d$.

More-dimensional wavelets

A *wavelet space* $W_n \subset V_n$, a closed subspace of V_n which contains those functions in V_n that cannot be represented in any of the function spaces on the next coarser level, i.e., these functions are in V_n but not in $\mathrm{Span}(V_{n-e_1}, \ldots, V_{n-e_d})$. Thus $W_n \subset V_n$ is a closed subspace so that

$$V_n = W_n \oplus \mathrm{Span}(V_{n-e_1}, \ldots, V_{n-e_d}). \tag{35}$$

This means that W_n contains the 'difference information' that is available in the fine grid V_n but not in the span of the coarser grids $V_{n-e_1}, V_{n-e_2}, \ldots, V_{n-e_d}$.

The space W_n is the *complement* of $\mathrm{Span}(V_{n-e_1}, \ldots, V_{n-e_d})$ in V_n. Of course, this complement is not uniquely determined. If we want we can make use of the Hilbert space structure and consider the (unique) *orthogonal complement*

$$W_n \perp \mathrm{Span}(V_{n-e_1}, \ldots, V_{n-e_d}). \tag{36}$$

This choice corresponds with $R_n: X \to V_n$ being the orthogonal projection. However, in many cases we will use spaces W_n that don't satisfy this orthogonality property!

As soon as we have selected a CSR $\{R_n\}$, then corresponding pre-wavelet spaces are defined as in Section 2.4. These pre-wavelet spaces on an MRA are wavelet spaces.

In the case of an MRA no coarsest grid exists, so that (28) gives

$$V_n = \bigoplus_{j \leqslant n} W_j. \tag{37}$$

Because of property (34a) we can decompose the space $X(\Omega)$ in

$$X(\Omega) = \bigoplus_{j \in \mathbb{Z}^d} W_j \tag{38}$$

so that we can write any $u \in X(\Omega)$ as $u = \sum_{j \in \mathbb{Z}^d} w_j$ with $w_j \in W_j$. A restriction $R_n: X(\Omega) \to V_n$ is now determined by

$$R_n \left(\sum_j w_j \right) = \sum_{j \leqslant n} w_j. \tag{39}$$

By Definition 2.19 we recognize the direct hierarchical surplus

$$Q_n: X(\Omega) \to \mathrm{Ran}(Q_n) = W_n. \tag{40}$$

We see that there is no coarsest grid and we can decompose R_n as

$$R_n = \sum_{k \leqslant n} Q_k. \tag{41}$$

The four relations (34a)–(34d) imply that also the spaces W_n are scaled versions of one space W_0,

$$f(\boldsymbol{x}) \in W_n \Leftrightarrow f(2^{-n}\boldsymbol{x}) \in W_0, \quad \forall n \in \mathbb{Z}^d, \tag{42}$$

and, moreover, that they are translation invariant for the discrete translations $2^{-n}\mathbb{Z}^d$,

$$f(\boldsymbol{x}) \in W_0 \Leftrightarrow f(\boldsymbol{x} - \boldsymbol{k}) \in W_0, \quad \forall \boldsymbol{n} \in \mathbb{Z}^d. \tag{43}$$

As soon as we find a function $\psi(\boldsymbol{x})$ with the property that $\psi(\boldsymbol{x} - \boldsymbol{k})$, $k \in \mathbb{Z}^d$, is a basis of W_e, then by a simple rescaling we see that $\psi(2^n \boldsymbol{x} - \boldsymbol{k})$, yields a basis of W_{n+e}. Such a function is the more-dimensional generalization of a *wavelet* [3]. Because of (38) the full collection

$$\left\{ \psi_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}) \mid \psi_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}) = \psi(2^n \boldsymbol{x} - \boldsymbol{k}), \ \boldsymbol{n}, \boldsymbol{k} \in \mathbb{Z}^d \right\}$$

is a basis of $X(\mathbb{R}^d)$.

3. Piecewise approximation in d dimensions

In this section, we first describe the approximating function spaces for piecewise constant and piecewise multilinear approximation. For the spaces V_n and for the pre-wavelet spaces W_n we describe basis functions. Using these bases, in Sections 3.2 and 3.3 we give proofs for error estimates on regular and sparse grids. Most of the estimates are essentially also found in [2,13,16], but here the proofs are more general, simpler and given in a unified treatment.

3.1. Piecewise approximation

3.1.1. Piecewise constant approximation

First, let $\Omega = \mathbb{R}^d$ or $\Omega = (0,1)^d$. We approximate $u \in X(\Omega) = L^p_{\text{loc}}(\Omega)$ by $u_n \in V_n$, in the space of piecewise constant functions on Ω_n, i.e., in

$$V_n = \text{Span}\left(\{\phi_{nj}\}\right), \tag{44}$$

with, for some $q \geqslant 1$ or $q = \infty$,

$$\phi_{\boldsymbol{n}\boldsymbol{j}}(\boldsymbol{x}) = 2^{|n|/q}\phi(2^n \boldsymbol{x} - \boldsymbol{j}),$$

$$\phi(\boldsymbol{x}) = \prod_{j=1}^{d} \chi_{[0,1]}(x_j), \quad \text{with } \chi_{[0,1]}(x) \text{ the characteristic function on the unit interval.} \tag{45}$$

This clearly describes a basis for a tensor product space, and we may write

$$V_{\boldsymbol{n}} = V_{\boldsymbol{n}}(\Omega) = \bigotimes_{j=1}^{d} V_{n_j}(\Omega_j), \tag{46}$$

the tensor product of spaces $V_{n_j}(\Omega_j)$. These V_{n_j} are the spaces of piecewise constant functions with meshwidth $h_j = 2^{-n_j}$ on $\Omega_j \subset \mathbb{R}$. The corresponding grid of cells on the Cartesian product of $\{\Omega_j\}$, is denoted by Ω_n. The cell centers are denoted by by Ω_n^*.

We define the restriction R_n as the projection

$$R_n : X \to V_n \subset X,$$

$$u \mapsto u_n = R_n u, \quad \text{with } u_{n,i} = u_n\big((\boldsymbol{i} + \boldsymbol{e}/2)h\big) = 2^{+|n|} \int_{\Omega_{ni}} u(\xi)\,\mathrm{d}\Omega. \tag{47}$$

This restriction is of type (30), and $R_{\boldsymbol{n}} = R_{n_1,\dots,n_d}$ can be decomposed as

$$R_{\boldsymbol{n}} = \prod_{j=1}^{d} R_{n_j \boldsymbol{e}_j}, \tag{48}$$

where $R_{n_j \boldsymbol{e}_j} u(\boldsymbol{x})$ is the function, piecewise constant in the jth coordinate direction on a partitioning $\Omega_{\boldsymbol{n}}$, so that

$$R_{n_j \boldsymbol{e}_j} u(\boldsymbol{x}) = 2^{n_j} \int\limits_{\boldsymbol{x}-(h/2)\boldsymbol{e}_j}^{\boldsymbol{x}+(h/2)\boldsymbol{e}_j} u(\xi_1,\dots,\xi_d)\, \mathrm{d}\xi_j$$

for all $\boldsymbol{x}$ with $(x_j 2^{n_j} \pm 1/2) \in \mathbb{Z}$.

In the special case $X = L^2(\Omega)$, the space $X = X(\Omega)$ is a Hilbert space, and $\{\phi_{\boldsymbol{nj}}\}$ is an orthogonal (orthonormal if $q = 2$) basis in $V_{\boldsymbol{n}}$. In this case $R_{\boldsymbol{n}}$ is the orthogonal projection $L^2(\Omega) \to V_{\boldsymbol{n}}$. For $\Omega = \mathbb{R}^d$, the set $\{V_{\boldsymbol{n}}\}$ as defined in (44)–(45) is a typical MRA. This is no longer the case if we consider a bounded domain Ω, but the decomposition as treated in Section 2 still can be used in the case of a bounded domain.

It is easily checked that the more-dimensional wavelet $\psi(\boldsymbol{x}) \in W_{\boldsymbol{e}}$, corresponding with the piecewise constant scaling function $\phi(\boldsymbol{x}) \in V_{\boldsymbol{0}}$, from the previous section, is the more-dimensional *elementary checkerboard function* given by

$$\psi(\boldsymbol{x}) = \begin{cases} 0, & \text{if } \boldsymbol{x} \notin \Omega_{\boldsymbol{0},\boldsymbol{0}}, \\ (-1)^{|\boldsymbol{k}|}, & \text{if } \boldsymbol{x} \in \Omega_{\boldsymbol{0},\boldsymbol{0}} \text{ and } \boldsymbol{x} \in \Omega_{\boldsymbol{e},\boldsymbol{k}}. \end{cases} \tag{49}$$

This function is the tensor product of the Haar wavelet.

In wavelet theory the spaces $W_{\boldsymbol{n}}$ are labeled *channels*, and the distinct channels are linearly independent. The first decomposition of an arbitrary function from $X(\Omega)$ consists in writing $u(\boldsymbol{x}) = \sum_{\boldsymbol{n}} w_{\boldsymbol{n}}(\boldsymbol{x})$, where $w_{\boldsymbol{n}} \in W_{\boldsymbol{n}}$ with $\boldsymbol{n} \in \mathbb{Z}^d$, according to (38).

Each subspace $W_{\boldsymbol{n}}$, has its natural basis, the *standard basis*,[3]

$$\left\{ \psi_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}) \mid \psi_{\boldsymbol{n},\boldsymbol{k}}(\boldsymbol{x}) = 2^{|\boldsymbol{n}-\boldsymbol{e}|/q} \psi\big(2^{\boldsymbol{n}-\boldsymbol{e}} \boldsymbol{x} - \boldsymbol{k}\big),\ \boldsymbol{k} \in \mathbb{Z}^d \right\}, \tag{50}$$

of functions with a minimal support. We see that $\psi = \psi_{\boldsymbol{e},\boldsymbol{0}} \in V_{\boldsymbol{e}}$ is a function with the unit cube $\Omega_{\boldsymbol{0},\boldsymbol{0}}$ as support. The basis function $\psi_{\boldsymbol{n},\boldsymbol{k}}$ is a scaled, elementary checkerboard function, that may be characterized either by its support, which is a single cell in $\Omega_{\boldsymbol{n}-\boldsymbol{e}}$, or by the centerpoint of this cell, $\boldsymbol{z}_{\boldsymbol{n}-\boldsymbol{e},\boldsymbol{k}} = 2^{-|\boldsymbol{n}-\boldsymbol{e}|}(\boldsymbol{k} + \boldsymbol{e}/2)$.

On the open unit cube $\Omega = (0,1)^d$ we consider the $2^{|\boldsymbol{n}|}$-dimensional spaces $V_{\boldsymbol{n}} = \bigotimes_{j=1}^{d} V_{n_j}$, the tensor product of $V_{n_j}((0,1))$, the spaces of piecewise constant functions with meshwidth $h_j = 2^{-n_j}$ in the jth coordinate direction. For functions defined on $\Omega = (0,1)^d$ we can write relation (38) as

$$X(\Omega) = \bigoplus_{\boldsymbol{n} \geqslant 0} W_{\boldsymbol{n}}, \tag{51}$$

---

[3] Notice that in more dimensions we use the indexing $\psi_{\boldsymbol{n}+\boldsymbol{e},\boldsymbol{k}}$, whereas in the one-dimensional case one usually writes ψ_k^n.

and make a decomposition in channels correspondingly. Each subspace W_{n+e}, now with $n \geqslant 0$, has its standard basis $\psi_{n+e,k}$:

$$\left\{ \psi_{n+e,k}(x) \mid \psi_{n+e,k}(x) = 2^{|n|/q} \psi(2^n x - k), \ 0 \leqslant k < 2^n \right\}. \tag{52}$$

For $\Omega = (0,1)^d$, the exceptions related with the boundary are found in the spaces W_n with a zero index (i.e., $\|n\| = 0$). These W_n have basis functions with different shapes. They are derived from the corresponding functions for the unbounded case, but their support is restricted to Ω. Their corresponding nodal points $z_{n-e,k}$, are found on the boundary $\partial\Omega = \overline{\Omega}\backslash\Omega$ instead of in the interior. For $\|n\| = 0$ we have W_n spanned by a basis

$$\left\{ \psi_{n,k}(x) \ \middle| \ \begin{array}{ll} 0 \leqslant k_j < 2^{n_j - 1}, & \text{if } n_j \neq 0, \\ k_j = 0, & \text{if } n_j = 0, \end{array} \ j = 1, \ldots, d \right\}. \tag{53}$$

Taking such modifications into account, both for $\Omega = (0,1)^d$ and for $\Omega = \mathbb{R}^d$, for each $u \in L^p_{\text{loc}}(\Omega)$ we may write a *hierarchical expansion* (a wavelet expansion) according to (38) or (51), as

$$u(x) = \sum_{n \in \mathbb{Z}} w_n = \sum_{n,k} c_{nk} \psi_{n,k}(x) = \sum_{n,k} c_{nk} \widehat{\psi}(2^n x - k), \tag{54}$$

where $\widehat{\psi}$ is simply

$$\widehat{\psi}(x) = \begin{cases} \text{sign} \|x\|, & \text{if } \max(x_1, \ldots, x_d) < 1, \\ 0, & \text{if } \max(x_1, \ldots, x_d) > 1, \end{cases}$$

and $c_{nk} = 0$ for all k with $\|k\|$ even.

3.1.2. Piecewise d-linear approximation

We approximate $u \in X = C^0(\Omega)$ by $u_n \in V_n$, in the space of piecewise d-linear functions on Ω_n, i.e., in

$$V_n = \text{Span}\left(\{\phi_{nj}\}\right), \tag{55}$$

with, for some $q \geqslant 1$ or $q = \infty$,

$$\phi_{nj}(x) = 2^{|n|/q} \phi(2^n x - j),$$
$$\phi(x) = \prod_{j=1}^{d} \Lambda(x_j), \quad \text{with } \Lambda(x) = \max(0, 1 - |x|) \text{ the usual hat function.} \tag{56}$$

Clearly this is a basis for a tensor product space as (46), where V_{n_j} are spaces of piecewise linear functions on a partitioning of Ω_j with meshwidth $h_j = 2^{-n_j}$. The set of nodal points $\{j2^n\}_{j \in \mathbb{Z}^d}$ in Ω is denoted by Ω_n^+.

Here we define the restriction R_n as the projection

$$R_n : X \to V_n \subset X,$$
$$u \mapsto u_n = R_n u, \quad u_n(x) = u(x) \ \forall x \in \Omega_n^+. \tag{57}$$

The restriction is also of type (30) and the operator $R_n = R_{n_1,\dots,n_d}$ can be decomposed as (48) where $R_{n_j e_j} u(x)$ is a function, piecewise linear in the jth coordinate direction, on a partitioning Ω_n, such that $R_{n_j e_j} u(x) = u(x)$ for all x with $x_j / h_j \in \mathbb{Z}$.

It is clear that there exists a basis-function in V_n for each nodal point x_{nk} in $\overline{\Omega}_n^+$. If and only iff $\|k\|$ is even, there exists a 'parent grid' $\overline{\Omega}_m^+$ with $m \leqslant n$ and $m \neq n$, for which $x_{nk} \in \overline{\Omega}_m^+$. Hence, in this case each wavelet space W_n has its natural basis

$$\left\{ \phi_{nk} \mid k \in \mathbb{Z}^d \text{ with } \|k\| \text{ odd} \right\}. \tag{58}$$

On the closed cube $\overline{\Omega} = [0,1]^d$ we consider the $\prod_{j=1}^d (2^{n_j} + 1)$-dimensional spaces $V_n = V_n(\Omega) = \bigotimes_{j=1}^d V_{n_j}([0,1])$. With *homogeneous Dirichlet boundary conditions*, the dimension of the corresponding space $V_n^0 \subset V_n$ is $\prod_{j=1}^d (2^{n_j} - 1)$. It is immediately clear that typical FE-basis functions for V_n are the d-dimensional hat-functions: functions that vanish on all but one point of $\overline{\Omega}_n^+$. Each such FE basis function of V_n^0 is characterized by an interior point from Ω_n^+.

We notice that for $\overline{\Omega} = [0,1]^d$ we have

$$V_n = \{0\} \quad \text{except for } n \geqslant 0, \tag{59}$$
$$V_n^0 = \{0\} \quad \text{except for } n \geqslant e. \tag{60}$$

With W_n (or W_n^0) we denote the subspace of V_n (respectively V_n^0) of functions that vanish at the gridpoints of all $\overline{\Omega}_{n-e_j}^+$ (respectively $\Omega_{n-e_j}^+$), $j = 1, \dots, d$. From (55) we see that for $\Omega = \mathbb{R}^d$

$$W_n = \mathrm{Span}\left(\left\{ \phi_{nj} \mid \|j\| \text{ odd} \right\} \right) \tag{61}$$

and for $\Omega = [0,1]^d$ we see that $0 \leqslant j \leqslant 2^n$ and

$$W_n = \mathrm{Span}\left(\left\{ \phi_{nj} \; \middle| \; \begin{array}{ll} j_i \text{ odd}, \ 0 < j_i < 2^{n_i}, & \text{if } n_i > 0, \\ j_i = 0, 1, & \text{if } n_i = 0, \end{array} \ i = 1, \dots, d \right\} \right). \tag{62}$$

Clearly $W_n = W_n^0 = \{0\}$, except for $n \geqslant 0$. If $\|n\| = 0$ we see that $V_n^0 = \{0\}$ and W_n is spanned by FE basis functions that are characterized by boundary points on the unit cube. Thus, the trace of a function on the boundary is exclusively approximated by elements of W_n with $\|n\| = 0$. Further we see $W_n^0 = W_n$ if $n \geqslant e$. Apparently

$$V_n = \bigoplus_{0 \leqslant k \leqslant n} W_k, \quad \text{and} \quad V_n^0 = \bigoplus_{e \leqslant k \leqslant n} W_k. \tag{63}$$

3.2. Error estimates for regular grids

The decompositions of type (38) allow the approximation of a sufficiently smooth function in $X(\Omega)$ by a series with elements in W_j. To obtain an impression of the quality of these expansions in the following sections we derive error estimates.

3.2.1. Estimates for piecewise constant approximation

As the case where domain boundaries are present, is the more general one, we study the case $\Omega = (0, 1)^d$. To quantify the error of approximation on Ω, we introduce for $u \in C^e(\overline{\Omega})$ the seminorm

$$
|u| = \max_{x \in \Omega} \left| D^e u(x) \right| + \max_{\substack{0 \leqslant p \leqslant e \\ 0 < |p| < d}} \max_{x \in \overline{\Omega} \setminus \Omega} \left| D^p u(x) \right|. \tag{64}
$$

Now we derive the following

Theorem 3.1. *If we consider an expansion of a $C^e(\overline{\Omega})$-function, u, in piecewise constant functions on the grid Ω_n, for an arbitrary $n \in \mathbb{Z}^d$, $n > 0$, and if we write*

$$
R_n u = \sum_{0 \leqslant m \leqslant n} w_m, \tag{65}
$$

with $w_m \in W_m$, $0 \leqslant m \leqslant n$, then, for $m \neq 0$ we have

$$
\|w_m\|_{L_2(\Omega)} \leqslant 2^{-d/2} 2^{-|m|} |u|, \tag{66}
$$

and an estimate for the approximation error

$$
\|u - R_n u\|_{L_2(\Omega)} \leqslant (2/3)^{d/2} \|h_n\| |u|. \tag{67}
$$

Proof. We take $\{\psi_{mk}\}$ as a basis in W_m, $e \leqslant m \leqslant n$. All these functions form an $L_2(\Omega)$-orthonormal set (orthonormal Haar basis) and they are orthogonal to all functions in W_n, $n \neq m$.

Thus, we find (65) with $w_m = \sum_k a_{mk} \psi_{mk}$, where

$$
a_{mk} = (u, \psi_{m,k}) = \int_{\Omega} u\, \psi_{m,k}\, \mathrm{d}\Omega = \int_{\Omega_{m-e,k}} u \psi_{m,k}\, \mathrm{d}\Omega.
$$

For $m \geqslant e$ the point $z_{m-e,k}$ lies in the interior of Ω and the estimate holds with

$$
|u| = \max_x \left| D^e u(x) \right|.
$$

Viz., by Taylor expansion around $z_{m-e,k}$, we have

$$
|a_{mk}| = \left| \int_{\Omega_{m-e,k}} u \psi_{m,k}\, \mathrm{d}\Omega \right|
$$

$$
\leqslant \left| \int_{\Omega_{m-e,k}} \|x - z_{m-e,k}\| |u| \psi_{m,k}\, \mathrm{d}\Omega \right|
$$

$$
= |u| 2^{|m-e|/2} 2^d \prod_{j=1}^{d} \int_0^{2^{-m_j}} \xi_j\, \mathrm{d}\xi_j
$$

$$
= |u| 2^{-d/2} 2^{-3|m|/2}. \tag{68}
$$

For $m \neq 0$, $m \not\geqslant e$, i.e., for ψ_{mk} with an m-component equal to zero, the point $z_{m-e,k}$ lies on the boundary and the function $\psi_{m,k}$ is constant in one direction over the whole domain Ω, and it is of

Haar-wavelet type for the non-zero indices (or index). In this case the orthonormal basis functions are $2^{\beta/2}\psi_{m,k}$, where β is the number of zeroes in m. Nevertheless, provided that $m \neq 0$, in this situation the same estimate (68) holds with, e.g., if $m_1 = 0$ and $m_j \geqslant 1$ for $j = 2, \ldots, d$,

$$|u| = \max_{x} \left| \frac{\partial^{d-1} u(x)}{\partial x_2 \cdots \partial x_d} \right|.$$

Hence, the estimate (68) holds for $m \geqslant 0$, $m \neq 0$, if we use the seminorm (64), and we find

$$\|w_m\|^2 = \sum_k |a_{mk}|^2 \leqslant \sum_k 2^{-d} 2^{-3|m|} |u|^2 = 2^{-d} 2^{-2|m|} |u|^2, \tag{69}$$

so that $\|w_m\| \leqslant 2^{-d/2} 2^{-|m|} |u|$, which leads to (66) and (67) because

$$\|u - R_n u\|^2 = \sum_{\substack{m_1 > n_1 \\ \text{or} \ldots \text{or} \\ m_d > n_d}} \|w_m\|^2 \leqslant \sum_{j=1}^{d} \sum_{\substack{m \geqslant 0 \\ m_j > n_j}} 2^{-d} 2^{-2|m|} |u|^2$$

$$\leqslant 2^{-d} |u|^2 \sum_{j=1}^{d} \sum_{\substack{m \geqslant 0 \\ m_j > n_j}} (1/4)^{m_1 + \cdots + m_d}$$

$$\leqslant |u|^2 \, (2/3)^d \sum_{j=1}^{d} h_{n_j}^2 \leqslant |u|^2 \, (2/3)^d \|h_n\|^2. \qquad \square$$

If we have no further a priori knowledge about u, the most efficient approximation will be one with $h_1 = \cdots = h_d$ because this equalizes the main terms in the error bound. We see that the truncation error for $u - R_n u$ is not particularly promising or surprising: the major part of the error is produced by the largest meshwidth: $(\max(h_1, \ldots, h_d))^{d/2}$, whereas the total number of degrees of freedom for an element in V_n is $2^{|n|}$.

3.2.2. Estimates for piecewise d-linear approximation

For a function $u \in C^0(\Omega)$ we consider piecewise linear approximation as in Section 3.1.2. We approximate u by $u_n \in V_n$, where V_n is the space of piecewise d-linear functions on Ω_n. We take u_n such that $u_n(x) = u(x)$ for all $x \in \Omega_n^+$ and we write

$$u_n(x) = \sum_j d_{nj} \phi_{nj}(x), \tag{70}$$

where $\phi_{nj}(x)$ is defined by (56).

With $u_n \in V_n$ the piecewise linear approximation on Ω_n of the function $u \in C^0(\Omega)$, we make the hierarchical decomposition $V_n = \bigoplus_{k \leqslant n} W_k$, and write

$$u_n = \sum_{k \leqslant n} w_k, \quad w_k \in W_k, \tag{71}$$

where

$$w_k(x) = \sum_j c_{kj} \phi_{kj}(x), \tag{72}$$

with $c_{kj} = 0$ for all j with $\|j\|$ even.

In practice the coefficients c_{kj}, $\|j\|$ odd, are computed as *hierarchical surplus coefficients*, by taking the difference between the value $u(jh_k)$ and the interpolant from coarser grids. This is most conveniently formulated by introducing stencil notation. Therefore, we introduce the *difference operator*

$$\nabla_h u(z) = u(z + h) - u(z),$$

$$(73)$$

and the usual central difference approximation for the second derivative by stencil notation, as

$$\left[\tfrac{1}{2}, -1, \tfrac{1}{2}\right]_{h_j e_j} u(z) = \tfrac{1}{2} \nabla^2_{h_j e_j} u(z - h_j e_j).$$

With this notation we write an expression for the hierarchical coefficients in a piecewise linear approximation. We see that d-linear interpolation leads to the following expression for the hierarchical surplus coefficient

$$c_{kj} \|h_k\|^{-1/q} = \prod_{j=1,d} \left[-\tfrac{1}{2}, 1, -\tfrac{1}{2}\right]_{h_j e_j} u(jh_k).$$

$$(74)$$

Notice that the factor $\|h_k\|^{1/q}$ cancels the scaling factor $2^{|k|/q}$ in the definition of ϕ_{kj}, so that the function $u(x)$ is expanded as

$$u(x) \approx \sum_{k,j} c_{kj} \|h_k\|^{-1/q} \phi\left(2^k x - j\right).$$

An expression for the coefficient c_{kj} is found in the following lemma.

Lemma 3.2. *Let $u \in C^{e+m}$, for a given m with $0 \leqslant m \leqslant e$, and let*

$$L_{nj}(x) = 2^{-|n|} \phi\left(2^n x - j\right),$$

$$(75)$$

then, for each $\phi_{nj} \in W_n$, $\|n\| \neq 0$, $\|j\|$ odd, we have

$$\|h_n\|^{-1/q} |c_{nj}| = \prod_{i=1}^{d} \left[-\tfrac{1}{2}, 1, -\tfrac{1}{2}\right]_{h_i e_i} u(jh_n)$$

$$= (-1)^{|e+m|} 2^{-d} \int_\Omega D^{e+m} u(x) \, D^{e-m} L_{nj}(x) \, \mathrm{d}\Omega.$$

Proof. We see that for $\|n\| \neq 0$, $\|j\|$ odd, each L_{nj} has a support in the interior of Ω. Taking this into account, we give the proof after a coordinate translation with $2^m j$ then we see that for all i, $0 \leqslant i \leqslant d$,

$$\sum_{s_i=-1,1} \int_0^{s_i h_i} D^{e_i} u(z) \, \mathrm{d}z_i = \sum_{s_i=-1,1} u(z) \Big|_{z_i=0}^{s_i h_i} = \sum_{s_i=-1,1} \nabla_{s_i h_i e_i} u(z) \Big|_{z_i=0} = [1, -2, 1]_{h_i e_i} u(z)|_{z_i=0}$$

and hence

$$\prod_{i=1,d} \left[-\tfrac{1}{2}, 1, -\tfrac{1}{2} \right]_{h_i e_i} u(\mathbf{0})$$

$$= \left(-\tfrac{1}{2}\right)^d \prod_{i=1,d} \left[\sum_{s_i=-1,1} \int_0^{s_i h_i} D^{e_i} \bullet \mathrm{d}x_i \right] u(x)$$

$$= \left(-\tfrac{1}{2}\right)^d \sum_{s_1,\ldots,s_d=-1,1} \int_0^{s_1 h_1} \cdots \int_0^{s_d h_d} D^e u(\mathbf{x}) \, \mathrm{d}x_1 \cdots \mathrm{d}x_d$$

$$= \left(-\tfrac{1}{2}\right)^d \int_{-h_1}^{h_1} \cdots \int_{-h_d}^{h_d} D^e u(\mathbf{x}) \prod_{i=1}^d \left(2^{-n_i} \frac{\mathrm{d}}{\mathrm{d}x_i} \Lambda\big(2^{n_i} x_i\big) \right) \mathrm{d}x_1 \cdots \mathrm{d}x_d$$

$$= \left(-\tfrac{1}{2}\right)^d \int D^e u(\mathbf{x}) 2^{-|n|} D^e \phi\big(2^{|n|} \mathbf{x}\big) \, \mathrm{d}\Omega$$

$$= (-1)^{|e+m|} \, 2^{-d} \int D^{e+m} u(\mathbf{x}) D^{e-m} L_{n\mathbf{0}}(\mathbf{x}) \, \mathrm{d}\Omega. \qquad \square$$

Remarks.

(1) For $\|n\| = 0$ (i.e., for boundary points), the same formula holds, provided that the formula is restricted to the lower dimensional boundary manifold (e.g., the face or the edge of the unit cube).

(2) For an m with $0 \leqslant m \leqslant e$ we derive an expression for $\|D^m \phi\|_p$, with ϕ given by (56) as follows:

$$\|D^m \phi\|_p^p = \int_\Omega \prod_i |D^{m_i} \Lambda(x_i)|^p \, \mathrm{d}\Omega = 2^d \, (p+1)^{-|e-m|}.$$

So that

$$\|D^m \phi\|_p = 2^{d/p} (p+1)^{-|e-m|/p}. \tag{76}$$

(3) In (56) we have $\phi_{nj} = 2^{|n|/q} \phi(2^n \mathbf{x} - \mathbf{j})$, and hence

$$\|D^m \phi_{nj}\|_p^p = \int 2^{|n|p/q} |D^m \phi(2^n \mathbf{x} - \mathbf{j})|^p \, \mathrm{d}\Omega = 2^{|n|(p/q-1)} 2^{|m \circ n|p} \|D^m \phi\|_p^p.$$

Here,

$$2^{m \circ n} = \prod_{i=1}^d 2^{m_i n_i} = \prod_{i=1}^d h_{n_i}^{-m_i} = h_n^{-m},$$

so that for arbitrary j,

$$\|D^m \phi_{nj}\|_p = \|h_n\|^{(1/p-1/q)} h_n^{-m} \|D^m \phi\|_p. \tag{77}$$

This means that the norm $\|\phi_{nj}\|_p$ is independent of the level n iff we take $q = p$.

(4) To obtain error estimates in the approximation Theorem 3.3 we compute an expression for $\|D^m L_{nj}(x)\|_p$, with L_{nj} as introduced in (75), in particular for $p = 1, 2, \infty$.

$$\|D^m L_{nj}(x)\|_p^p = \|h_n\|^{p+1} h_n^{-pm} \|D^m \phi\|_p^p = \|h_n\|^{p+1} h_n^{-pm} 2^d (p+1)^{-|e-m|}.$$

So that we may conclude, also considering the special case $p = \infty$,

$$\|D^m L_{nj}(x)\|_1 = \|h_n\|^2 (h_n/2)^{-m},$$

$$\|D^m L_{nj}(x)\|_2 = (2/3)^{d/2} \|h_n\|^{3/2} (h_n/\sqrt{3})^{-m},$$

$$\|D^m L_{nj}(x)\|_\infty = \|h_n\| (h_n)^{-m}.$$

Using the above expressions and Lemma 3.2, we can derive the error estimates in the following theorem.

Theorem 3.3. *Let* $u \in C_0^{e+m}(\Omega)$ *be given for some* m *with* $0 \leqslant m \leqslant e$, *and let* $u_n \in V_n$ *be the piecewise linear approximation on* Ω_n *of* u, *such that* $u_n(x) = u(x)$ *for all* $x \in \Omega_n^+$. *If we make the hierarchical decomposition* $V_n = \bigoplus_{k \leqslant n} W_k$, *and write*

$$u_n = \sum_{k \leqslant n} w_k, \quad w_k \in W_k,$$

then we have the estimates

$$\|w_k\|_2 \leqslant \|D^{e+m} u\|_2 \|h_k\|^2 2^{-d/2} 3^{-|m|/2} h_k^{-(e-m)},$$

$$\|w_k\|_\infty \leqslant \|D^{e+m} u\|_\infty \|h_k\|^2 (h_k/2)^{-(e-m)},$$

$$\|u - u_n\|_2 \leqslant \|D^{e+m} u\|_2 2^{-d} 3^{-3|m|/2} \sum_{i=1}^d h_{n_i}^{(1+m_i)},$$

$$\|u - u_n\|_\infty \leqslant \|D^{e+m} u\|_\infty 6^{-|m|} \sum_{i=1}^d h_{n_i}^{(1+m_i)}.$$

Proof. Using (74) and Lemma 3.2 we can obtain estimates for the hierarchical coefficients c_{kj}. We fix k and we derive, writing $h = h_k$,

$$\|h\|^{-1/q} |c_{kj}| = 2^{-d} \left| \int_\Omega D^{e+m} u(x) \chi_{kj}(x) D^{e-m} L_{kj}(x) \, dx \right|$$

$$\leqslant 2^{-d} \|D^{e+m} u \chi_{kj}\|_\infty \|D^{e-m} L_{kj}\|_1$$

$$\leqslant 2^{-d} \|D^{e+m} u \chi_{kj}\|_\infty \|h\|^2 (h/2)^{-(e-m)},$$

where χ_{kj} is the characteristic function for the support of $L_{kj}(x)$, or similarly

$$\|h\|^{-1/q} |c_{kj}| = 2^{-d} \left| \int_\Omega D^{e+m} u(x) \chi_{kj} D^{e-m} L_{kj}(x) \, dx \right|$$

$$\leqslant 2^{-d} \|D^{e+m} u \chi_{kj}\|_2 \|D^{e-m} L_{kj}\|_2$$

$$\leqslant 6^{-d/2} \|D^{e+m} u \chi_{kj}\|_2 \|h\|^{3/2} (h/\sqrt{3})^{-(e-m)}.$$

We write $w_k = \sum_j c_{kj} \phi_{kj}$ with $\|j\|$ odd, and we know that these functions $\{\phi_{kj}\}_j$, for fixed k have disjoint supports. Hence, for the hierarchical contribution,

$$\|w_k\|_2^2 = \left\| \sum_i c_{ki} \phi_{ki} \right\|_2^2 = \sum_{i,j} c_{ki} c_{kj} \int \phi_{ki} \phi_{kj} \, d\Omega = \sum_i c_{ki}^2 \|\phi_{ki}\|_2^2$$

$$\leqslant \|h\|^{2/q} \|\phi_{ki}\|_2^2 \sum_j \left\| D^{e+m} u \chi_{kj} \right\|_2^2 6^{-d} \|h\|^3 \left(h/\sqrt{3} \right)^{-2(e-m)}$$

$$\leqslant \|h\|^4 \, 2^{-d} \left\| D^{e+m} u \right\|_2^2 h^{-2(e-m)} 3^{-|m|}.$$

For the other norm $\|w_k\|_\infty$ we obtain similarly

$$\|w_k\|_\infty = \left\| \sum_j c_{kj} \phi_{kj} \right\|_\infty \leqslant \max_j |c_{kj}| \|\phi_{kj}\|_\infty$$

$$\leqslant \|h\|^{1/q} \max_j 2^{-d} \left\| D^{e+m} u \, \chi_{kj} \right\|_\infty \|h\|^2 (h/2)^{-(e-m)} \|h\|^{-1/q}$$

$$\leqslant 2^{-d} \left\| D^{e+m} u \right\|_\infty \|h\|^2 (h/2)^{-(e-m)}.$$

For the error, for $p = 2$ or $p = \infty$, we get

$$\|u - u_n\|_p = \left\| \sum_k w_k - \sum_{k \leqslant n} w_k \right\|_p \leqslant \sum_{k \not\leqslant n} \|w_k\|_p$$

$$\leqslant C_p \left\| D^{e+m} u \right\|_p \sum_{k \not\leqslant n} \|h_k\|^2 h_k^{-(e-m)}$$

$$= C_p \left\| D^{e+m} u \right\|_p \sum_{k \not\leqslant n} h_k^{e+m},$$

with $C_2 = 2^{-d/2} 3^{-|m|/2}$ or $C_\infty = 2^{-d} 2^{|e-m|}$. This yields the above mentioned estimates, by taking into account that

$$\sum_{k \not\leqslant n} h_k^{e+m} = \sum_k h_k^{e+m} - \sum_{k \leqslant n} h_k^{e+m}$$

$$= \prod_{i=1}^d \frac{1}{1 - \left(\frac{1}{2} \right)^{1+m_i}} \left[1 - \prod_{i=1}^d \left(1 - \left(\tfrac{1}{2} \right)^{(1+n_i)(1+m_i)} \right) \right]$$

$$\leqslant 2^d \left(\tfrac{2}{3} \right)^{|m|} \left[1 - \prod_{i=1}^d \left(1 - 2^{-(1+n_i)(1+m_i)} \right) \right] = 3^{-|m|} \sum_{i=1}^d h_{n_i}^{(1+m_i)}.$$

For $m = e$ this simply reads $\sum_{k \not\leqslant n} \|h_k\|^2 \leqslant 3^{-d} \|h_n\|^2.$ □

Corollary 3.4. *As a direct corollary we find*

$$\|u - u_n\|_2 \leqslant 54^{-d/2} \left\| D^{2e} u \right\|_2 \|h_n\|^2, \qquad \|u - u_n\|_\infty \leqslant 6^{-d} \left\| D^{2e} u \right\|_\infty \|h_n\|^2,$$

and, for $p = 2$ or $p = \infty$, and $0 \leqslant m \leqslant e$,

$$\|w_{\boldsymbol{k}}\|_p \leqslant C \|D^{\boldsymbol{e}+\boldsymbol{m}} u\|_p \boldsymbol{h}_{\boldsymbol{k}}^{\boldsymbol{e}+\boldsymbol{m}}, \tag{78}$$

$$\|u - u_{\boldsymbol{n}}\|_p \leqslant C \|D^{\boldsymbol{e}+\boldsymbol{m}} u\|_p \sum_{i=1}^{d} h_{n_i}^{1+m_i}, \qquad \|u - u_{\boldsymbol{n}}\|_p \leqslant C \|D^{2\boldsymbol{e}} u\|_p \|\boldsymbol{h}_{\boldsymbol{n}}\|^2. \tag{79}$$

From (78) also follows a bound by a $W_p^{\boldsymbol{e},\ell}$-norm. We immediately see that, for $0 \leqslant \ell \leqslant d$,

$$\|w_{\boldsymbol{k}}\|_p \leqslant C \|u\|_{W_p^{\boldsymbol{e},\ell}} \min_{|\boldsymbol{m}|=\ell} \boldsymbol{h}_{\boldsymbol{k}}^{\boldsymbol{e}+\boldsymbol{m}} \leqslant C \|u\|_{W_p^{\boldsymbol{e},\ell}} \|\boldsymbol{h}_{\boldsymbol{k}}\|^{1+\ell/d}. \tag{80}$$

We gave the proof of Theorem 3.3 for functions that vanish at the boundary of Ω. Taking into account the remark following Lemma 3.2, it is clear that similar estimates (with different constants and with terms including derivatives of u that are restricted to the boundary planes as in (64)) also hold for functions with non-homogeneous boundary conditions.

3.3. Error estimates for sparse grids

3.3.1. Estimates for piecewise constant approximation

For piecewise constant approximation we use a *sparse (box) grid* $\Omega_n = \bigcup_{|\boldsymbol{k}| \leqslant n} \Omega_{\boldsymbol{k}} \cap \Omega$. A sparse grid approximation is obtained by interpolation on this grid by means of the space spanned by all $W_{\boldsymbol{k}}$ with $|\boldsymbol{k}| \leqslant n$.

Theorem 3.5. *Let $\widehat{R}_n u$ be the piecewise constant approximation of a function $u \in L_2(\Omega)$ on a sparse grid on level n:*

$$\widehat{R}_n u = \sum_{|\boldsymbol{k}| \leqslant n} w_{\boldsymbol{k}}, \qquad w_{\boldsymbol{k}} \in W_{\boldsymbol{k}}, \tag{81}$$

then, with $\|\boldsymbol{h}\| = 2^{-n}$, the volume of the finest cells, we have the estimate

$$\left\| u - \widehat{R}_n u \right\|_{L_2(\Omega)} \leqslant C |u| \|\boldsymbol{h}\| \log^{(d-1)/2} \|\boldsymbol{h}\|. \tag{82}$$

Proof. To prove the theorem for the $L_2(\Omega)$-norm, we use (69), and the orthogonality of the hierarchical basis functions, to obtain

$$\left\| u - \widehat{R}_n u \right\|_{L_2(\Omega)}^2 \leqslant \sum_{|\boldsymbol{k}|>n} \|w_{\boldsymbol{k}}\|_{L_2(\Omega)}^2 \leqslant \sum_{|\boldsymbol{k}|>n} 2^{-d} 4^{-|\boldsymbol{k}|} |u|^2 = 2^{-d} |u|^2 \sum_{|\boldsymbol{k}|>n} 2^{-2|\boldsymbol{k}|}. \tag{83}$$

We know

$$\sum_{l>n} 2^{-2l} \binom{l+d-1}{d-1} = 2^{-2(n+1)} \binom{n+d}{d-1} F(1, 1+n+d; 2+n; 1/4) = G(n,d). \tag{84}$$

Here F is the hypergeometric function. It follows that

$$G(n,d) \sim \frac{n^{d-1} 2^{-2n}}{3(d-1)!} \quad \text{for } n \to \infty,$$

where the asymptotic value is reached soon for small values of d. Hence

$$\|u - R_n u\|^2 \leqslant 2^{-d}|u|^2 \sum_{l > n} 2^{-2l} \binom{l+d-1}{d-1} \leqslant 2^{-d}|u|^2 G(n,d) \leqslant 2^{-d}|u|^2 C_{nd} \frac{n^{d-1}2^{-2n}}{3(d-1)!}, \quad (85)$$

where C_{nd} is a constant that tends to one for large n ($C_{nd} < d$ for $n \geqslant 4$). So, we conclude that

$$\|u - R_n u\|_{L_2(\Omega)} \leqslant C|u|n^{(d-1)/2}2^{-n},$$

which is equivalent with (82). $\quad\square$

To guarantee a small error on a regular grid, in (67) all cell edges h_j need to be small, but in (82) for the sparse grid only the volume $\|h\|$ has to be small. Further, in the two-dimensional case, the estimate (82) is of a similar order of accuracy as (67), except for a logarithmic small factor. However, the number of degrees of freedom for the approximation (82) is significantly less. Namely, in the unit cube, for $R_n u$ the number of degrees of freedom is $2^{|n|}$, whereas for $\widehat{R}_n u$ it is $O(n^{d-1}2^n)$, viz. $2n2^n + 1$ in the 2D-case, and in the 3D-case, e.g., $(n^2 + n + 2)2^n - 1$. Because significantly less degrees of freedom are involved in the approximation $\widehat{R}_n u$ than in the approximation of $R_{(n,n,n)}u$, i.e., less coefficients $a_{j,k}$ and less gridpoints $z_{j,k}$, in analogy to [7], we call the approximation $\widehat{R}_n u$ the *sparse grid approximation* and

$$\Omega_n^\star = \left\{ z_{j,k} \mid z_{j,k} \in Z_n, \ |n| \leqslant n \right\}$$

is the *sparse (box) grid* for this approximation on level n.

3.3.2. Estimates for piecewise linear approximation

For piecewise linear approximation we use a *sparse (vertex) grid* $\Omega_n^+ = \bigcup_{|k| \leqslant n} \Omega_k^+ \cap \Omega$. A sparse grid approximation is obtained by interpolation on this grid by means of the space spanned by all W_k with $|k| \leqslant n$.

Theorem 3.6. *Let $\widehat{R}_n u$ be the piecewise d-linear approximation of a function $u \in C_0^{e+m}(\Omega)$, with $0 \leqslant m \leqslant e$, on a sparse grid on level n:*

$$\widehat{R}_n u = \sum_{|k| \leqslant n} w_k, \quad w_k \in W_k, \quad (86)$$

then, with $\|h\| = 2^{-n}$, the volume of the finest cells, we have for $p = 2, \infty$, with $m = e$ the estimates

$$\left\| u - \widehat{R}_n u \right\|_p \leqslant C \left\| D^{2e} u \right\|_p \|h\|^2 \log^{d-1} \|h\|^{-1}, \quad (87)$$

and with $|m| < d$ the estimates

$$\left\| u - \widehat{R}_n u \right\|_p \leqslant C \left\| D^{e+m} u \right\|_p \|h\| \log^{d-1-|m|} \|h\|^{-1}, \quad (88)$$

and with $0 \leqslant \ell \leqslant d$

$$\left\| u - \widehat{R}_n u \right\|_p \leqslant C \|u\|_{W_p^{e,\ell}} \|h\|^{1+\ell/d} \log^{d-1} \|h\|^{-1}. \quad (89)$$

Proof. Using the estimates for $\|w_k\|_p$ from Theorem 3.3, we prove, more generally, for some m with $0 \leqslant m \leqslant e$, and for $p = 2$ or $p = \infty$,

$$\left\| u - \widehat{R}_n u \right\|_p \leqslant \sum_{|\boldsymbol{k}|>n} \left\| w_{\boldsymbol{k}} \right\|_p$$

$$\leqslant \sum_{|\boldsymbol{k}|>n} C \left\| D^{e+m} u \right\|_p h_{\boldsymbol{k}}^{e+m}$$

$$= C \left\| D^{e+m} u \right\|_p \sum_{|\boldsymbol{k}|>n} h_{\boldsymbol{k}}^{e+m}$$

$$\leqslant C \left\| D^{e+m} u \right\|_p 2^{-n} \left(C_1 2^{-n} n^{|\boldsymbol{m}|-1} + C_2 n^{d-|\boldsymbol{m}|-1} \right), \tag{90}$$

with $C_1 = 0$ if $|\boldsymbol{m}| = 0$, and $C_2 = 0$ if $|\boldsymbol{m}| = d$. Hence, for $\boldsymbol{m} \neq \boldsymbol{e}$ we have

$$\left\| u - \widehat{R}_n u \right\|_p \leqslant C \left\| D^{e+m} u \right\|_p \|\boldsymbol{h}\| \log^{d-1-|\boldsymbol{m}|} \|\boldsymbol{h}\|^{-1}.$$

Moreover, (90) yields, for $\boldsymbol{m} = \boldsymbol{e}$,

$$\left\| u - \widehat{R}_n u \right\|_p \leqslant C \left\| D^{2e} u \right\|_p \|\boldsymbol{h}\|^2 \log^{d-1} \|\boldsymbol{h}\|^{-1}.$$

Further, using the estimate (80) we obtain, similar to the proof for Theorem 3.5,

$$\left\| u - \widehat{R}_n u \right\|_p \leqslant \sum_{|\boldsymbol{k}|>n} \| w_{\boldsymbol{k}} \|_p \sum_{|\boldsymbol{k}|>n} C \| u \|_{W_p^{e,\ell}} \| h_{\boldsymbol{k}} \|^{1+\ell/d}$$

$$= C \| u \|_{W_p^{e,\ell}} \sum_{|\boldsymbol{k}|>n} 2^{-|\boldsymbol{k}|(1+\ell/d)}$$

$$= C \| u \|_{W_p^{e,\ell}} \binom{n+d}{d-1} F\left(1, 1+n+d; 2+n; 2^{-(1+\ell/d)}\right) 2^{-n(1+\ell/d)}$$

$$\leqslant C \| u \|_{W_p^{e,\ell}} \|\boldsymbol{h}\|^{1+\ell/d} \log^{d-1} \|\boldsymbol{h}\|^{-1}. \qquad \Box \tag{91}$$

Theorem 3.7. *Let $\widehat{R}_n u$ be the piecewise d-linear approximation of a function $u \in C_0^{e,1}(\Omega)$ on a sparse grid on level n, as in Theorem 3.6, then, with $\|\boldsymbol{h}\| = 2^{-n}$, the volume of the finest cells, we have, for $p = 2$, $p = \infty$, the estimates*

$$\left\| u - \widehat{R}_n u \right\|_{W_p^1} \leqslant C \|\boldsymbol{h}\| \log^{d-1} \|\boldsymbol{h}\|^{-1} \| u \|_{W_p^{e,1}}. \tag{92}$$

If, moreover, we know $u \in C^{2e}$, then

$$\left\| u - \widehat{R}_n u \right\|_{W_p^1} \leqslant C \|\boldsymbol{h}\| \left\| D^{2e} u \right\|_p. \tag{93}$$

Proof. Let u be sufficiently differentiable and let $0 \leqslant \boldsymbol{m} \leqslant \boldsymbol{e}$ and $|\boldsymbol{m}| \geqslant 1$, then:
  Part 1:

$$\left\| D^m w_{\boldsymbol{k}} \right\|_p = \left\| D^m \prod_{j=1}^d (R_{\boldsymbol{k}} - R_{\boldsymbol{k}-e_j}) u \right\|_p$$

$$\leqslant C^d \left\| D^m \prod_{j=1}^d h_{k_j} D^{e_j} u \right\|_p$$

$$\leqslant C^d \|\boldsymbol{h}_{\boldsymbol{k}}\| \left\| D^m D^e u \right\|_p, \tag{94}$$

$$\left\| D^m \big(u - \widehat{R}_n u \big) \right\|_p \leqslant \left\| \sum_{|k|>n} D^m w_k \right\|_p$$

$$\leqslant C^d \left\| D^{m+e} u \right\|_p \frac{n^{d-1} 2^{-n}}{(d-1)!} \overline{G}(n,d)$$

$$\leqslant C \left\| D^{m+e} u \right\|_p \|h\| \log^{d-1} \|h\|. \tag{95}$$

Part 2:

$$\left\| D^m \big(u - \widehat{R}_n u \big) \right\|_p \leqslant \left\| \sum_{|k|>n} D^m w_k \right\|_p$$

$$\leqslant C \left\| D^{2e} u \right\|_p 2^{-n} \big(C_1 2^{-n} n^{d-|m|-1} + C_2 n^{|m|-1} \big), \tag{96}$$

with $C_1 = 0$ if $|m| = d$, and $C_2 = 0$ if $|m| = 0$.
Because

$$\|v\|_{\mathcal{W}_p^1} = \left(\|v\|_p^p + \sum_{|m|=1, 0 \leqslant m \leqslant e} \|D^m v\|_p^p \right)^{1/p},$$

we consider the case $|m| = 1$ and we find

$$\left\| u - \widehat{R}_n u \right\|_{W_p^1} \leqslant C \left\| D^{2e} u \right\|_p \|h\|.$$

Together with the result of Theorem 3.6 this proves the theorem. $\square$

References

[1] G. Beylkin, *Wavelets and Fast Numerical Algorithms*, Proceedings of Symposia in Applied Mathematics 47 (American Mathematical Society, Providence, RI, 1993).

[2] H.J. Bungartz, Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson–Gleichung, Ph.D. Thesis, Institut für Informatik, TU München (1992).

[3] C.K. Chui, *An Introduction to Wavelets* (Academic Press, 1992).

[4] P.M. de Zeeuw, Development of semi-coarsening techniques, *Applied Numerical Mathematics* 19 (1996) 433–465.

[5] J.E. Dendy, M.P. Ida and J.M. Rutledge, A semicoarsening multigrid algorithm for SIMD machines, *SIAM J. Sci. Statist. Comput.* 13 (1992) 1460–1469.

[6] M. Griebel, W. Huber, U. Rüde and T. Störtkuhl, The combination technique for parallel sparse-grid-preconditioning or -solution of PDEs on workstation networks, in: *Parallel Processing: CONPAR 92—VAPP V, Proceedings of the Second Joint International Conference on Vector and Parallel Processing*, Lyon, France, September 1–3, 1992, Lecture Notes in Computer Science 634 (Springer, Berlin, 1992) 217–228.

[7] M. Griebel, M. Schneider and C. Zenger, A combination technique for the solution of sparse grid problems, in: *Proceedings of the IMACS International Symposium on Iterative Methods in Linear Algebra* (Elsevier, Amsterdam, 1992) 263–281.

[8] M. Griebel, C. Zenger and S. Zimmer, Multilevel Gauss–Seidel algorithms for full and sparse grid problems, *Computing* 50 (1993) 127–148.

[9] C.E. Heil and D.F Walnut, Continuous and discrete wavelet transforms, *SIAM Rev.* 31 (1989) 628–666.

[10] P.W. Hemker, Sparse-grid finite-volume multigrid for 3D-problems, *Adv. Comput. Math.* 4 (1995) 83–110.

[11] P.W. Hemker and P.M. de Zeeuw, BASIS3: A data structure for 3-dimensional sparse grids, Technical Report NM-R9321, CWI, Amsterdam (1993).

[12] W.A. Mulder, A high resolution Euler solver based on multigrid, semi coarsening, and defect correction, *J. Comput. Phys.* 100 (1992) 91–104.

[13] C. Pflaum, Diskretisierung elliptischer Differentialgleichungen mit dünnen Gittern, Ph.D. Thesis, Technische Universität München, Munich (1995).

[14] R.A. Smith and A. Weiser, Semicoarsening multigrid on a hypercube, *SIAM J. Sci. Statist. Comput.* 13 (1992) 1314–1329.

[15] S.A. Smoljak, Quadrature and interpolation formulas for tensor product of certain classes of functions, *Dokl. Akad. Nauk* 148 (1963) 1042–1045; translation: Soviet Math. 4 (1963) 240–243.

[16] C. Zenger, Sparse grids, in: *Parallel Algorithms for Partial Differential Equations: Proceedings of the Sixth GAMM-Seminar, Kiel, January 1990*, Notes on Numerical Fluid Mechanics 31 (Vieweg, Braunschweig, 1991).

[17] L.-B. Zhang, Semi-coarsening in multigrid solution of steady incompressible Navier–Stokes equations, *J. Comput. Math.* 8 (1990) 92–97.