ELSEVIER

Applied Numerical Mathematics 19 (1996) 433–465

APPLIED NUMERICAL MATHEMATICS

Development of semi-coarsening techniques *

Paul M. de Zeeuw [1]

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

Abstract

Departing from Mulder's semi-coarsening technique for first order PDEs, the notion of a grid of grids is introduced and a multi-level finite-volume technique for second order elliptic PDEs is developed. Various grid transfer operators are investigated, in combination with damped Jacobi relaxation. Convergence rates as they are predicted by Fourier local mode analysis are compared with practical measurements. The wide variety of grids at our disposal leads to the notion of coherent representations of a function on different grids. A sawtooth multi-level algorithm is proposed for the case of multiple semi-coarsening. A hierarchical set of basis functions for finite volumes on sparse grids is briefly discussed.

Keywords: Coherence; Finite volumes; Grid of grids; Hierarchical basis; Linear systems; Mesh aspect ratio; Multigrid; Multi-level; Multiple semi-coarsening; Semi-coarsening; Sparse grids; Wavelets

1. Introduction

In multigrid methods we have to take care of obtaining adequate coarse grid corrections to accelerate an iterative solution process. The standard procedure of grid coarsening, i.e. doubling the mesh size in each space dimension, is known to be not robust in more-dimensional cases where *flow alignment* or *anisotropic diffusion* occurs (see [2,14]). These are examples of phenomena, defined in more space dimensions, that are locally one-dimensional in essence and do not really allow for coarsening in all directions. Here, a classical coarse grid correction (CGC) fails to yield proper corrections, simply because these cannot be represented on the standard coarsened grid. This has to be compensated for by powerful smoothing procedures. E.g. in 2D one applies linewise instead of pointwise relaxation methods, or one resorts to incomplete factorizations. Indeed, in [25, §7.12] the best smoothing methods that can handle both the (rotated) anisotropic Poisson equation and convection–diffusion equations are exactly of this type. A similar story goes for the solution of Navier–Stokes equations

* This research was supported by the Brite-EuRam II Aeronautics R & D Programme of the European Communities (Contract No. AER2-CT92-0040).

[1] Paul.de.Zeeuw@cwi.nl.

SSDI 0168-9274(95)00095-X

when the unknowns are strongly coupled in one direction due to high mesh aspect ratios in for instance high Reynolds boundary layer and wake flows and also in far-field flows.

Where multigrid standard coarsening in 2D lacks the possibility of representing components in the error that are low frequent in one direction but very high frequent in the other, this holds even more in three space dimensions.

One rigorous remedy is to apply semi-coarsening, i.e. coarsening merely in the direction of the strongest coupling [2,14]. Thus, at the expense of a larger number of grid points on the coarse grid, we can represent components in the correction which are of low frequency in the direction of the strong coupling and of high frequency in the other directions. In this manner, the CGC is effective and we do not need to put too high demands to the smoother and can therefore rely on simple procedures. In Fig. 1 a piecewise constant grid function is shown that is of the highest frequency in one direction and of the lowest in the other (a so-called washboard function). Restriction (by integration) to the standard coarsened grid annihilates the function whereas restriction to the semi-coarsened grid is clearly representative.

Another, but related, remedy is the frequency decomposition multigrid method (FDMGM) of Hackbusch (see [15,16]). In 2D, through shifting the standard coarse grid one obtains three further grids for the representation of corrections corresponding to four different types of frequencies. Depending on the type, the grids are recursively coarsened. In 3D this approach can be generalized to the creation of seven further coarse grids, shifting the standard coarse grid. Evaluation of the defect equations on all the coarse grids leads to a multiple, intertwined, coarse grid correction. Again, at the expense of a larger complexity on the coarse grid, one obtains an effective CGC which clears the way for a simple smoothing operator.

In 2D, Mulder proposed to perform semi-coarsening in two directions simultaneously [19,20]. A fine grid is coarsened in the x- and y-direction respectively. Vice versa, each coarse grid is linked to two finer grids, refined in the x- and y-direction respectively. Similar information from different coarse grids is combined in order to limit the complexity to $O(N)$, where N is the number of cells on the finest grid. This combining of information implies the averaging of residuals. For the transfer of corrections from coarse grids to fine grids Mulder proposes alternating prolongations, in the x- and y-direction respectively. Naik and Van Rosendale [21] propose a weighted average of the interpolated corrections in the x- and y-direction.

A related, but different, approach can be found in [10] where on different subspaces, in parallel, a PDE is discretized and solved (on e.g. a four color division of the space of grid functions).

In 3D, semi-coarsening in three directions takes an extra storage of $7N$ cells for the coarse grids. Though this amount of storage is only proportional to the number of unknowns, it still may become prohibitive in practice. Zenger [26] launched the idea of defining a specific set of hierarchical basis functions within an FE-space. The set is chosen in a way that N reduces significantly with only a slight deterioration of accuracy of (smooth) solutions. The supports of the hierarchical basis functions relate to each other by semi-coarsening (and semi-refinement) in the same way as the semi-coarsened finite volumes.

In this paper we consider finite volumes rather than finite elements. We confine ourselves to linear problems, but we indicate where a generalization is possible. First, in Section 2 we describe a multi-level method on a complete grid of grids, this is in contrast with the classical approach of a sequence of standard coarsened grids. The use of some particular grid transfer operators is discussed and numerical results are given for the (an)isotropic diffusion equation. In Section 3 we

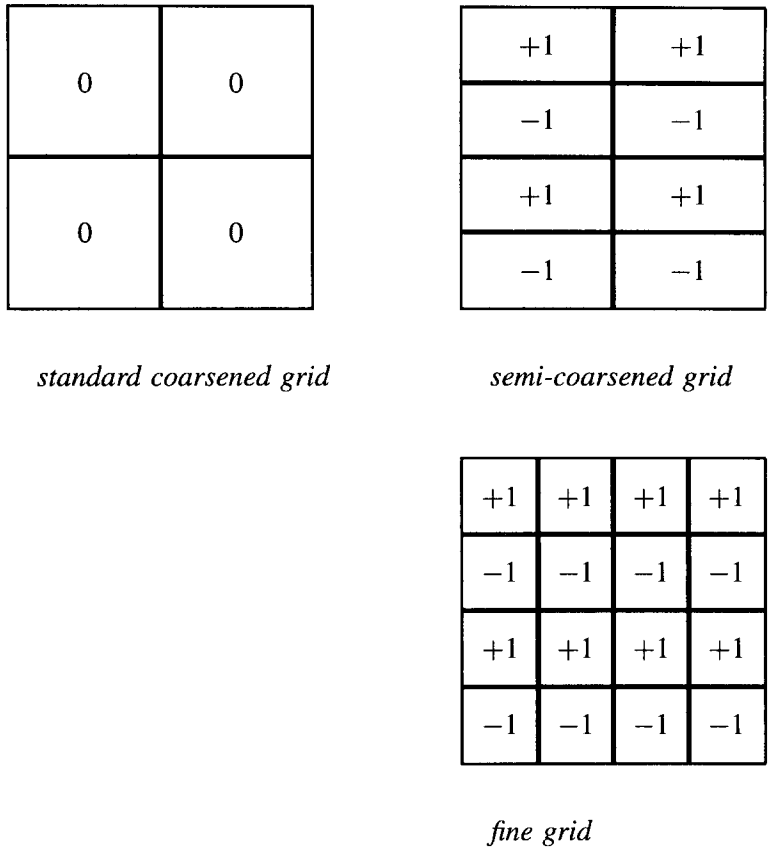

Fig. 1. Coarsening of a washboard grid function.

touch upon the possibility of developing a multi-level method for sparse grids which should be the finite-volume counterpart of algorithms already developed for finite elements (see e.g. [13]). The use of hierarchical basis functions in the context of finite volumes, is also described in this section. In Section 4 conclusions are summarized.

2. The multi-level method on a grid of grids

Before we arrive at the proposed multi-level algorithm and its results (in the Sections 2.6–2.9), we first introduce general notions (and notations) in Section 2.1 and, specifically, grid transfer operators in Sections 2.2–2.3. In Section 2.4 we describe the notion of what will be called "coherence" of grid functions. Hereby we can describe an important possible difference between a representation of a function on multiple semi-coarsened grids on one hand and on sparse grids on the other. In

Section 2.5 we investigate the Galerkin approach for the discretization on coarser grids; for linear problems with constant coefficients we can give a detailed analysis of the stencils resulting from the Galerkin approach.

2.1. Grids of grids

For convenience we introduce the notation used for the two-dimensional case. The analogous notation is used for three space dimensions, but the explanation would, possibly, be less clear by the abundance of indices. The set of natural numbers, supplied with zero, is written as $\mathbb{N}$. By $\boldsymbol{n}$ we denote a pair of integers (n_1, n_2) in $\mathbb{N}^2$. By $\boldsymbol{n} \leqslant \boldsymbol{m}$ we mean that $n_j \leqslant m_j$ for $j = 1, 2$. The inequalities $<$, $\geqslant$, $>$ between $\boldsymbol{n}$ and $\boldsymbol{m}$ are defined analogously. The domain of definition, Ω, is assumed to be the open unit square.

Now we introduce the following notation:

$\mathbf{0} = (0,0) \in \mathbb{N}^2$;

$\boldsymbol{e}_1 = (1,0) \in \mathbb{N}^2$;

$\boldsymbol{e}_2 = (0,1) \in \mathbb{N}^2$;

$\boldsymbol{e} = \boldsymbol{e}_1 + \boldsymbol{e}_2 \in \mathbb{N}^2$;

$|\boldsymbol{n}| = n_1 + n_2 \in \mathbb{N}$;

$\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$;

$h_{n_k} = 2^{-n_k} \in \mathbb{R}$ for $k = 1, 2$;

$N_{\boldsymbol{n}} = \{(x_1, x_2) \mid x_k = i_k h_{n_k},\ (i_1, i_2) \in \mathbb{N}^2\}$;

$\Omega_{\boldsymbol{n}} = \overline{\Omega} \cap N_{\boldsymbol{n}}$;

$\boldsymbol{G} = \{\Omega_{\boldsymbol{n}} \mid \boldsymbol{n} \in \mathbb{N}^2\}$;

$\boldsymbol{G}_{\boldsymbol{n}} = \{\Omega_{\boldsymbol{m}} \mid \boldsymbol{m} \leqslant \boldsymbol{n},\ \boldsymbol{m} \in \mathbb{N}^2\}$;

$C_{\boldsymbol{n}} = \{(x_1, x_2) \mid x_k = (i_k + \frac{1}{2}) h_{n_k},\ (i_1, i_2) \in \mathbb{N}^2\}$;

$\Omega^c_{\boldsymbol{n}} = \Omega \cap C_{\boldsymbol{n}}$;

$g_{\boldsymbol{n}} : \Omega^c_{\boldsymbol{n}} \to \mathbb{R}$;

$g_{\boldsymbol{n},\boldsymbol{i}} = (g_{\boldsymbol{n}})_{\boldsymbol{i}} = g_{\boldsymbol{n}}(\boldsymbol{x})$ with $\boldsymbol{x} = ((i_1 + \frac{1}{2}) h_{n_1}, (i_2 + \frac{1}{2}) h_{n_2}) \in \Omega^c_{\boldsymbol{n}}$;

$S_{\boldsymbol{n}} = \{g_{\boldsymbol{n}} \mid g_{\boldsymbol{n}} : \Omega^c_{\boldsymbol{n}} \to \mathbb{R}\}$;

$R_{\boldsymbol{m},\boldsymbol{n}} : S_{\boldsymbol{n}} \to S_{\boldsymbol{m}}$ $(\boldsymbol{m} < \boldsymbol{n})$ a linear surjection;

$P_{\boldsymbol{n},\boldsymbol{m}} : S_{\boldsymbol{m}} \to S_{\boldsymbol{n}}$ $(\boldsymbol{m} < \boldsymbol{n})$ a linear injection;

$\Omega_{\boldsymbol{n},\boldsymbol{i}}$ = the interior of an elementary rectangle with vertices defined on $\Omega_{\boldsymbol{n}}$, with center $((i_1 + \frac{1}{2}) h_{n_1}, (i_2 + \frac{1}{2}) h_{n_2}) \in \Omega^c_{\boldsymbol{n}}$ and dimensions: h_{n_k}, $k = 1, 2$.

Table 1
Substructures in the grid of grids

Type	Definition	Meaning
1	$n_2 \geqslant 0 \wedge n_1 = c \geqslant 0$	x_1 *semi-coarsening*
2	$n_1 \geqslant 0 \wedge n_2 = c \geqslant 0$	x_2 *semi-coarsening*
3	$n_1, n_2 \geqslant 0 \wedge n_1 - n_2 = c$	*standard coarsening*
4	$n_1, n_2 \geqslant 0 \wedge n_1 + n_2 = l$	*grid level* l

Here $\Omega_{n,i}$ is called a *cell* and $g_n \in S_n$ is called a *grid function*. The set of *values* of g_n can be interpreted as components of a vector in $\mathbb{R}^{2^{|n|}}$. The set Ω_n^c is called the set of *cell centers* of Ω_n. The symbol S_n denotes the linear space of real-valued functions on Ω_n^c. The 2-tuple $\boldsymbol{n}$ is called the index of Ω_n. The integer $l = |\boldsymbol{n}|$ is called a *grid level*. $R_{m,n}$ is a *restriction* and $P_{n,m}$ is a *prolongation*. $\boldsymbol{G}$ is called the *infinite grid of grids*. A finite subset $\boldsymbol{F}$ of $\boldsymbol{G}$ is called a *(finite) grid of grids*. $\boldsymbol{G}_n$ is called a *complete* grid of grids (it follows at once that such a complete grid of grids is finite). A specific grid in the grid of grids is identified as Ω_n. See Fig. 2 for an illustration of a complete grid of grids. Within some of the coarser grids, in the upper left corner of this diagram, the cells, covering the grids, are indicated. Enumerated in Table 1, we recognize four types of straight lines in the grid of grids, each with their own meaning. We observe that semi-coarsening in either direction (Type 1, 2) and standard coarsening (Type 3) are included. Grids with the same grid level correspond to lines of Type 4. Grid levels will be of use in the description of algorithms to come. An example of a set of grids belonging to the same grid level l is shown in Fig. 2 by the dashed line. Note that grids on the same grid level l count the same number of cells.

When a grid of grids $\boldsymbol{F}$ can be written as $\boldsymbol{F} = \{\Omega_{n_1}, \ldots, \Omega_{n_\gamma}\}$, then the *least common multiple* (LCM($\boldsymbol{F}$)) of $\boldsymbol{F}$, is the grid Ω_n with

$$n_k = \max_{j=1,\ldots,\gamma} \{k\text{th component of } \boldsymbol{n}_j\},$$

and the *greatest common divisor* (GCD($\boldsymbol{F}$)) of $\boldsymbol{F}$ is the grid Ω_m with

$$m_k = \min_{j=1,\ldots,\gamma} \{k\text{th component of } \boldsymbol{n}_j\}.$$

An *incomplete* grid of grids is a grid of grids that is not complete. An *enclosure* $\boldsymbol{E}$ of a finite grid $\boldsymbol{F}$ of grids is a complete grid of grids that includes $\boldsymbol{F}$. The *smallest enclosure* $\boldsymbol{E}_n$ of a finite grid $\boldsymbol{F}$ of grids is an enclosure of $\boldsymbol{F}$ such that for no $\boldsymbol{m} < \boldsymbol{n}$ another enclosure exists. One can show there exists exactly one smallest enclosure, viz. $\boldsymbol{E}_n$ such that Ω_n is the least common multiple of $\boldsymbol{F}$. The grid Ω_n is then also called the *finest grid* of the smallest enclosure. An incomplete grid of grids *of the first kind* is an incomplete grid of grids which nevertheless includes the finest grid of its smallest enclosure. An incomplete grid of grids *of the second kind* is a grid of grids which does not include the finest grid of its smallest enclosure. See Fig. 12 for a specific example of a grid of grids that is incomplete of the second kind. It corresponds to Zenger's sparse grids [26].

The equation and its discretization

We investigate the two-dimensional case of the general (single) second order (elliptic) equation

$$Lu = f \quad \text{on } \Omega, \tag{2.1}$$

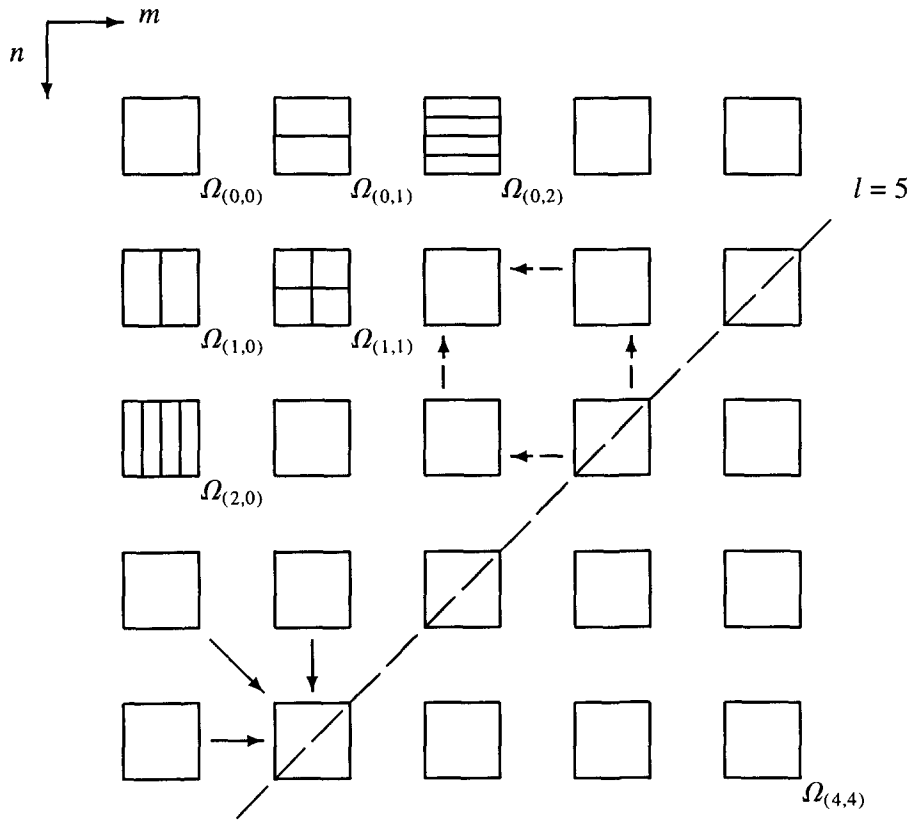

Fig. 2. A complete grid, $G_{(4,4)}$, of grids in $\mathbb{R}^2$.

with suitable boundary conditions. For a complete grid of grids G_n we can obtain discrete versions of equation (2.1) on each grid Ω_m with $m \leqslant n$,

$$L_m u_m = f_m \quad \text{on } \Omega_m. \tag{2.2}$$

In the rest of Section 2 we consider a multi-level approach for the solution of this equation on the finest grid.

2.2. *Restriction operators*

Let $f \in L^2(\Omega)$ be a square-integrable function, then we define the operator R_n by

$$R_n : L^2(\Omega) \to S_n, \tag{2.3a}$$

$$f_{n,i} = (R_n f)_i = \int_{\Omega_{n,i}} f \, \mathrm{d}\Omega, \tag{2.3b}$$

i.e. the function f is integrated over each cell $\Omega_{n,i}$ (see Section 2.1). Thus, grid function f_n corresponds to the finite-volume discretization of function f on Ω_n. Definition (2.3) leads to a natural definition of the restriction operation between grid functions within this context of finite-volume discretization. The definition reads:

$$R^{\mathrm{I}}_{n,n+e_k} : S_{n+e_k} \to S_n, \tag{2.4a}$$

$$f_{n,i} = (R^{\mathrm{I}}_{n,n+e_k} f_{n+e_k})_i = f_{n+e_k,2i} + f_{n+e_k,2i+e_k}. \tag{2.4b}$$

By the definition of integration, the integral of f over a box on the coarse grid is the sum of integrals of f over the two constituting boxes on the semi-refined grid. This explains definition (2.4): this restriction can be seen as a Riemann sum over subdomains. In the multi-level method to be described in Section 2.6 this restriction will be applied to grid functions that represent the right-hand side of (2.2).

The grids Ω_n are nested in the sense that $\Omega_n \supset \Omega_m$ when $\boldsymbol{n} \geqslant \boldsymbol{m}$. In Fig. 2 nesting takes place horizontally and vertically. Grids Ω_n which are on the same grid level $|\boldsymbol{n}| = l$ cannot be nested. Suppose we have some complete grid of grids $\boldsymbol{G}_n$. We note that for all grids $\Omega_m \in \boldsymbol{G}_n$ it holds true that $\Omega_m \subset \Omega_n$. Due to this nesting, restriction (2.4) (based on integration) is commutative w.r.t. the x_k-directions:

$$R^{\mathrm{I}}_{n+e,n+e_1} R^{\mathrm{I}}_{n+e_1,n} = R^{\mathrm{I}}_{n+e,n+e_2} R^{\mathrm{I}}_{n+e_2,n}. \tag{2.5}$$

See the dashed arrows in Fig. 2.

Definition 2.1. Let $S = \{R^{(\mathrm{S})}_{n-e_k,n} \mid \boldsymbol{n} \in \mathbb{N}^2, \boldsymbol{e}_k \in \{\boldsymbol{e}_1, \boldsymbol{e}_2\}\}$ be a set of 1D restriction operators. If the restriction

$$R^{(\mathrm{S})}_{n_\gamma,n_1} : S_{n_1} \to S_{n_\gamma}, \tag{2.6a}$$

$$R^{(\mathrm{S})}_{n_\gamma,n_1} = R^{(\mathrm{S})}_{n_\gamma,n_{\gamma-1}} R^{(\mathrm{S})}_{n_{\gamma-1},n_{\gamma-2}} \cdots R^{(\mathrm{S})}_{n_2,n_1}, \tag{2.6b}$$

with

$$\boldsymbol{n}_\gamma \leqslant \boldsymbol{n}_{\gamma-1} \leqslant \cdots \leqslant \boldsymbol{n}_1$$

is uniquely defined (i.e. independent of $\boldsymbol{n}_{\gamma-1} \cdots \boldsymbol{n}_2$ for $\gamma > 2$), the restriction $R^{(\mathrm{S})}_{n_\gamma,n_1}$ is called *path-independent.*

We note that $R^{\mathrm{I}}_{n_\gamma,n_1}$ is an example of a path-independent restriction because of (2.5). When f is a rapidly varying function, the restriction (2.4) yields an appropriate discretization procedure for f on the coarser grids [5].

2.3. Prolongation operators

Some prolongation operators are indicated as undashed arrows in Fig. 2.

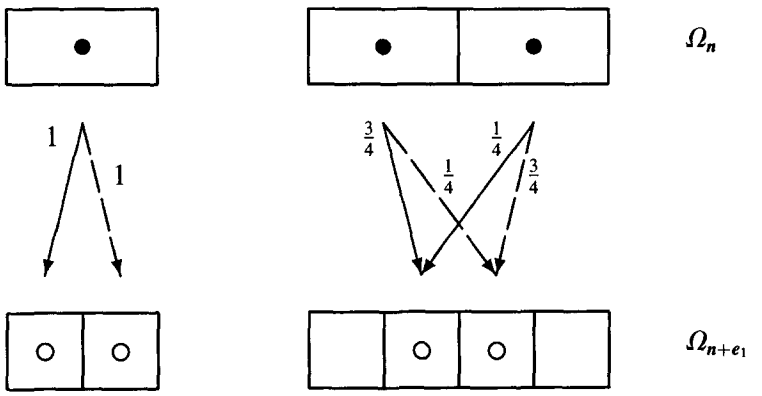

Fig. 3. Prolongations in the x_1-direction in $\mathbb{R}^2$.

Piecewise constant prolongation

The definition of piecewise constant prolongation in the x_k-direction reads:

$$P^1_{n+e_k,n} : S_n \to S_{n+e_k}, \tag{2.7a}$$

$$u_{n+e_k,2i} = (P^1_{n+e_k,n}u_n)_{2i} = u_{n,i}, \tag{2.7b}$$

$$u_{n+e_k,2i+e_k} = (P^1_{n+e_k,n}u_n)_{2i+e_k} = u_{n,i}. \tag{2.7c}$$

We note that the standard piecewise constant prolongation in $\mathbb{R}^2$ can be seen as the subsequent application of 1D piecewise constant prolongations. We also note that

$$P^1_{n+e_k,n} = (R^1_{n,n+e_k})^{\mathrm{T}}. \tag{2.8}$$

Second order prolongation

The definition of second order prolongation in the x_k-direction reads:

$$P^2_{n+e_k,n} : S_n \to S_{n+e_k}, \tag{2.9a}$$

$$u_{n+e_k,2i} = (P^2_{n+e_k,n}u_n)_{2i} = \tfrac{1}{4}u_{n,i-e_k} + \tfrac{3}{4}u_{n,i}, \tag{2.9b}$$

$$u_{n+e_k,2i+e_k} = (P^2_{n+e_k,n}u_n)_{2i+e_k} = \tfrac{3}{4}u_{n,i} + \tfrac{1}{4}u_{n,i+e_k}. \tag{2.9c}$$

This is shown in Fig. 3. The definition can also be applied (with cyclic numbering of i) at the boundary of Ω_n when periodic boundary conditions are prescribed.

Bilinear prolongation

The definition of the standard bilinear prolongation in $\mathbb{R}^2$ reads:

$$P^{\mathrm{bi}}_{n+e,n} : S_n \to S_{n+e}, \tag{2.10a}$$

$$u_{n+e,e+2i} = (P^{\mathrm{bi}}_{n+e,n}u_n)_{e+2i} = \tfrac{9}{16}u_{n,i} + \tfrac{3}{16}u_{n,i+e_1} + \tfrac{3}{16}u_{n,i+e_2} + \tfrac{1}{16}u_{n,i+e}, \tag{2.10b}$$

$$u_{n+e,e+2i+e_1} = (P^{\mathrm{bi}}_{n+e,n}u_n)_{e+2i+e_1} = \tfrac{3}{16}u_{n,i} + \tfrac{9}{16}u_{n,i+e_1} + \tfrac{1}{16}u_{n,i+e_2} + \tfrac{3}{16}u_{n,i+e}, \tag{2.10c}$$

$$u_{n+e,e+2i+e_2} = (P^{\mathrm{bi}}_{n+e,n}u_n)_{e+2i+e_2} = \tfrac{3}{16}u_{n,i} + \tfrac{1}{16}u_{n,i+e_1} + \tfrac{9}{16}u_{n,i+e_2} + \tfrac{3}{16}u_{n,i+e}, \tag{2.10d}$$

$$u_{n+e,e+2i+e} = (P^{\mathrm{bi}}_{n+e,n}u_n)_{e+2i+e} = \tfrac{1}{16}u_{n,i} + \tfrac{3}{16}u_{n,i+e_1} + \tfrac{3}{16}u_{n,i+e_2} + \tfrac{9}{16}u_{n,i+e}. \tag{2.10e}$$

This is symbolically shown in Fig. 4. Here, the cell centers of the respective coarser grids Ω_n, Ω_{n+e_1}, Ω_{n+e_2} are indicated by $\bullet$, whereas cell centers of Ω_{n+e} are indicated by $\circ$. The latter are also depicted in the coarser grids in order to demonstrate how the second order prolongations are determined. The definition can also be applied (with cyclic numbering of i) at the boundary of Ω_n when periodic boundary conditions are prescribed. We note that the bilinear prolongation can be decomposed into the second order 1D prolongations defined in (2.9):

$$P^{\mathrm{bi}}_{n+e,n} = P^2_{n+e,n+e_2}P^2_{n+e_2,n} = P^2_{n+e,n+e_1}P^2_{n+e_1,n}. \tag{2.11}$$

Definition 2.2. Let $S = \{P^{(\mathrm{S})}_{n,n-e_k} \mid n \in \mathbb{N}^2, e_k \in \{e_1, e_2\}\}$ be a set of 1D prolongation operators. If the prolongation

$$P^{(\mathrm{S})}_{n_\gamma,n_1} : S_{n_1} \to S_{n_\gamma}, \tag{2.12a}$$

$$P^{(\mathrm{S})}_{n_\gamma,n_1} = P^{(\mathrm{S})}_{n_\gamma,n_{\gamma-1}}P^{(\mathrm{S})}_{n_{\gamma-1},n_{\gamma-2}}\cdots P^{(\mathrm{S})}_{n_2,n_1}, \tag{2.12b}$$

with

$$n_1 \leqslant n_2 \leqslant \cdots \leqslant n_\gamma$$

is uniquely defined (i.e. independent of $n_2 \cdots n_{\gamma-1}$ for $\gamma > 2$), then the prolongation $P^{(\mathrm{S})}_{n_\gamma,n_1}$ is called *path-independent.*

Of course, $P^1_{n_\gamma,n_1}$ is a path-independent prolongation. Also $P^2_{n_\gamma,n_1}$ is a path-independent prolongation because of (2.11).

Combination of coarse grid corrections

We want to solve (2.2) on Ω_n by a multi-level approach. We define the residual:

$$r_n = f_n - L_n u_n. \tag{2.13}$$

We consider the grids Ω_{n-q},

$$q \in Q \equiv \{e_1, e_2, e\} \tag{2.14}$$

for acceleration of convergence by coarse grid correction. A picture is shown in Fig. 4 of prolongations stemming from the various coarser grids. We assume that the correction equations are solved exactly on the coarser grids, which means that we compute

$$c_{n-q} = L^{-1}_{n-q}R_{n-q,n}r_n. \tag{2.15}$$

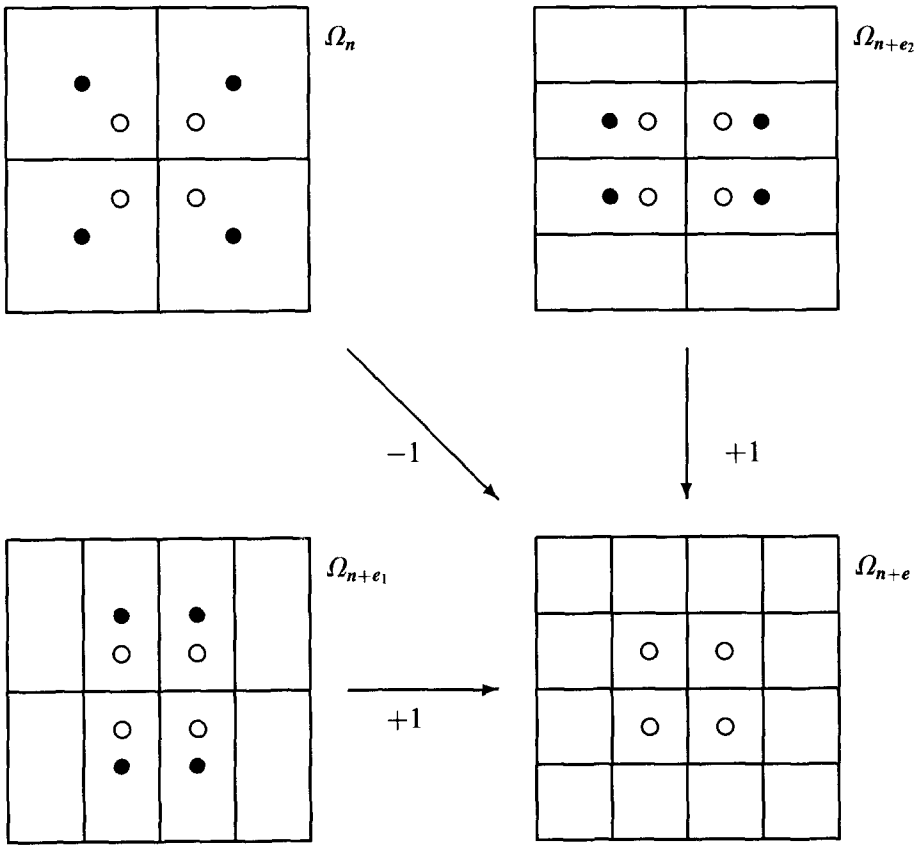

Fig. 4. Some second order prolongations in $\mathbb{R}^2$.

At first we might consider the following coarse grid correction

$$\tilde{u}_n = u_n + P_{n,n-q}c_{n-q} \tag{2.16}$$

for the three different possibilities given by (2.14). When $q = e$ we have the classical coarse grid correction in 2D for standard coarsened grids; when $q = e_k$ ($k = 1$ or 2) we have a coarse grid correction going with semi-coarsening. However, instead of (2.16) we apply the following coarse grid correction

$$\tilde{u}_n = u_n + \omega_{e_1}P_{n,n-e_1}c_{n-e_1} + \omega_{e_2}P_{n,n-e_2}c_{n-e_2} + \omega_e P_{n,n-e}c_{n-e}, \tag{2.17}$$

where a weighted combination of the corrections is chosen with weights:

$$\omega_{e_1}, \omega_{e_2}, \omega_e \in \mathbb{R}.$$

The correction as given by (2.17) is an example of additive subspace correction. We demand at least first order accuracy, i.e. when the c_{n-q} represent one and the same constant function then the weighted combination of the corrections should represent the same constant function, hence:

$$\omega_{e_1} + \omega_{e_2} + \omega_e = 1. \tag{2.18}$$

We note that both the approach of Naik and Van Rosendale [21]

$$\omega_{e_1} + \omega_{e_2} = 1, \ \omega_e = 0 \tag{2.19}$$

and the proposal of Rüde [22, p. 290] and Hemker [18]

$$\omega_{e_1} = 1, \ \omega_{e_2} = 1, \ \omega_e = -1 \tag{2.20}$$

fit within this framework. We examine how the multiple corrections reduce the residual, also involving the choice of the weights in (2.17). We write the amplification matrices for the residual due to the coarser grid corrections:

$$M_{n,n-q,n} : S_n \to S_n, \tag{2.21a}$$

$$M_{n,n-q,n} = I_n - L_n(P_{n,n-q}L_{n-q}^{-1}R_{n-q,n}), \tag{2.21b}$$

where I_n is the identity operator for grid functions in S_n. When we define

$$\tilde{r}_n = f_n - L_n\tilde{u}_n, \tag{2.22}$$

i.e. the residual after the coarse grid correction (2.17), it follows from (2.15) that

$$\tilde{r}_n = \left(\sum_{q\in Q} \omega_q M_{n,n-q,n}\right) r_n. \tag{2.23}$$

If we employ the Galerkin coarse grid approximation (GCA, see e.g. [25]) on all coarser grids, i.e.

$$L_{n-q} = R_{n-q,n}L_nP_{n,n-q}, \tag{2.24}$$

then it follows at once that

$$R_{n-q,n}M_{n,n-q,n} = \emptyset_{n-q}, \tag{2.25}$$

i.e. the nil-operator that annihilates all grid functions in S_{n-q}. We apply this result for the examination of the following grid functions:

$$R_{n-e_1,n}\tilde{r}_n, \qquad R_{n-e_2,n}\tilde{r}_n, \qquad R_{n-e,n}\tilde{r}_n,$$

i.e. the transfer by restriction of the new residual onto the coarser grids from which the coarser grid corrections originate. We assume the restriction operators to be path-independent (see Definition 2.1). Firstly, we easily establish that

$$R_{n-e,n}\tilde{r}_n = 0_{n-e}, \tag{2.26}$$

i.e. the zero grid function in the space S_{n-e}. Secondly, we can prove that

$$\begin{aligned} R_{n-e_1,n}\tilde{r}_n = R_{n-e_1,n}((\omega_{e_2} + \omega_e)I_n \\ - L_nP_{n,n-e_2}(\omega_{e_2}L_{n-e_2}^{-1} + \omega_eP_{n-e_2,n-e}L_{n-e}^{-1}R_{n-e,n-e_2}))R_{n-e_2,n}r_n \end{aligned} \tag{2.27}$$

under the additional assumption that the prolongation operators are path-independent (see Definition 2.2). We define the following operator, associated with the approximation property (see Hackbusch [14]):

$$A_{n,m,n} : S_n \to S_n \quad (\boldsymbol{m} < \boldsymbol{n}), \tag{2.28a}$$

$$A_{n,m,n} = L_n^{-1} - P_{n,m} L_m^{-1} R_{m,n}. \tag{2.28b}$$

When $\omega_e = -\omega_{e_2}$, the equality (2.27) reduces to:

$$R_{n-e_1,n}\tilde{r}_n = -\omega_{e_2} R_{n-e_1,n} L_n P_{n,n-e_2} A_{n-e_2,n-e,n-e_2} R_{n-e_2,n} r_n. \tag{2.29}$$

When the approximation property would hold between L_{n-e_2} and L_{n-e} up to a high order this results into a "small" grid function. Analogously, when $\omega_e = -\omega_{e_1}$, we derive

$$R_{n-e_2,n}\tilde{r}_n = -\omega_{e_1} R_{n-e_2,n} L_n P_{n,n-e_1} A_{n-e_1,n-e,n-e_1} R_{n-e_1,n} r_n. \tag{2.30}$$

It follows from (2.18) and $\omega_e = -\omega_{e_1}$, $\omega_e = -\omega_{e_2}$ that

$$\omega_{e_1} = 1, \qquad \omega_{e_2} = 1, \qquad \omega_e = -1.$$

Hence, equations (2.29) and (2.30) apply specifically to proposal (2.20).

2.4. Coherence

Definition 2.3. A grid function f_m is called a *coherent* right-hand side representation of f_n, $\boldsymbol{n} > \boldsymbol{m}$, when by means of a path-independent restriction $R_{m,n}$ it holds that

$$R_{m,n} f_n = f_m. \tag{2.31}$$

For $R_{m,n}$ we only consider $R^1_{m,n}$, see (2.4) and (2.6). For the elementary example $\boldsymbol{m} = \boldsymbol{n} - \boldsymbol{e}$ we observe that apparently

$$f_m = R_{n-e,n-e_1} R_{n-e_1,n} f_n = R_{n-e,n-e_2} R_{n-e_2,n} f_n$$

when f_m is a coherent right-hand side representation of f_n. Note that when in $\mathbb{R}^3$, instead of $\mathbb{R}^2$, there would be six instead of two ways of determining f_m which, again, should all yield the same grid function.

A set of grid functions $\{f_m\}$, defined on a grid of grids V, is called a *coherent set* when all f_m are, simultaneously, coherent representations of one grid function defined on the finest grid of the smallest enclosure of V.

For a complete grid of grids the coherence of right-hand sides (and of the residual and its transfers to coarser grids) can always be enforced if we choose to do so. This holds also true for an incomplete grid of the first kind (see Section 2.1). This is because all the coarser grid functions can be derived from one and only grid, namely the finest grid of the smallest enclosure.

The grid functions $f_{n_1}, \ldots, f_{n_\gamma}$ are called *mutually coherent* (*on the right*) when for the least common multiple of $\{\Omega_{n_1}, \ldots, \Omega_{n_\gamma}\}$ it holds that a grid function f_n exists such that all f_{n_j}, $j =$

$1, \ldots, \gamma$, are coherent right-hand representations of f_n. Note that by this definition grid functions f_{n_1}, f_{n_2} might be mutual coherent while neither $n_1 < n_2$, nor $n_2 < n_1$, nor $n_1 = n_2$.

Example 2.4. Suppose that $f_{n-e_1} = R_{n-e_1,n} f_n$ and $f_{n-e_2} = R_{n-e_2,n} f_n$ (with path-independent $R_{m,n}$) then f_{n-e_1} and f_{n-e_2} are mutually coherent.

We derive the following

Proposition 2.5. *Let Ω_n be the common divisor of Ω_{n_1} and Ω_{n_2}, then the grid functions f_{n_1} and f_{n_2} are mutually coherent if and only if*

$$R_{n,n_1} f_{n_1} = R_{n,n_2} f_{n_2}.$$

Proof.

(1) Let f_{n_1} and f_{n_2} be mutually coherent, then, by definition, for some m with $m \geqslant n_1$, $m \geqslant n_2$ a grid function f_m exists such that

$$f_{n_1} = R_{n_1,m} f_m,$$

$$f_{n_2} = R_{n_2,m} f_m.$$

It follows that

$$R_{n,n_1} f_{n_1} = R_{n,n_1} R_{n_1,m} f_m = R_{n,m} f_m = R_{n,n_2} R_{n_2,m} f_m = R_{n,n_2} f_{n_2}.$$

(2) We assume

$$R_{n,n_1} f_{n_1} = R_{n,n_2} f_{n_2}.$$

First we consider the canonical grids, i.e. $n_1 = n + e_1$ and $n_2 = n + e_2$. Then the least common multiple of Ω_{n_1} and Ω_{n_2} is Ω_m with $m = n + e$. We consider merely the canonical cells, see Fig. 5. Geometrically, the four subfigures are at the same location in $\mathbb{R}^2$. The values of the respective grid functions f_{n_1} and f_{n_2} at the canonical cells are given in this picture. Either grid function f_{n_k}, $k = 1, 2$, yields the value Σ on the cell at the common divisor Ω_n after application of R_{n,n_k}. We show that f_{n_k}, $k = 1, 2$, are mutually coherent w.r.t. some f_m to be constructed. In order to satisfy

$$f_{n_1} = R_{n_1,m} f_m,$$

$$f_{n_2} = R_{n_2,m} f_m$$

it is sufficient that the values $p, q, r, s \in \mathbb{R}$ of f_m satisfy the following system of linear equations:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \\ s \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & +1 & 0 \\ \frac{1}{2} & -1 & 0 \\ \frac{1}{2} & 0 & -1 \\ \frac{1}{2} & 0 & +1 \end{pmatrix} \begin{pmatrix} \Sigma \\ a \\ b \end{pmatrix}. \tag{2.32}$$

The general solution reads:

$$\begin{pmatrix} p \\ q \\ r \\ s \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & +\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{4} & +\frac{1}{2} & +\frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} & +\frac{1}{2} \end{pmatrix} \begin{pmatrix} \Sigma \\ a \\ b \end{pmatrix} + \mu \begin{pmatrix} +1 \\ -1 \\ -1 \\ +1 \end{pmatrix}, \quad \mu \in \mathbb{R}. \tag{2.33}$$

In this manner we construct f_m.

Other cases than the canonical grids follow by induction for $\boldsymbol{n}_1$, $\boldsymbol{n}_2$ in $\mathbb{N}^2$, as follows. Suppose that $\boldsymbol{n}_k > \boldsymbol{n} + \boldsymbol{e}_k$, $k = 1, 2$ (the most general case, special cases can be treated analogously). We start by considering the grid functions

$$f_{n+e_k} = R_{n+e_k,n_k} f_{n_k}, \quad k = 1, 2$$

on the corresponding grids with the common divisor Ω_n. It holds that

$$R_{n,n+e_1} f_{n+e_1} = R_{n,n+e_2} f_{n+e_2}$$

because of the basic assumption. By means of the construction (2.33) we create a grid function f_{n+e} w.r.t. to which f_{n+e_k} are mutually coherent. By repeating this procedure, we construct f_{n+e+e_k}, etc. until we have filled the smallest enclosure of Ω_{n_k}, $k = 1, 2$, one by one. W.r.t. the grid function f_m on the finest grid Ω_m of this smallest enclosure the f_{n_k}, $k = 1, 2$, are now mutually coherent. □

Remark 2.6. Note the degrees of freedom in constructing f_m because of the possible choices of μ. If we choose $\mu = 0$ in (2.33) then we choose the vector with the smallest 2-norm. Thus, when we apply $\mu = 0$ throughout at Ω_m, we construct the f_m with the smallest 2-norm.

The second part of Proposition 2.5 implies that also for a set of grid functions defined on an incomplete grid of grids of the second kind, we can easily establish (looking at the common divisors) whether those grid functions are a coherent set.

Coherence on the left-hand side

A similar notion of coherence can be defined as well for grid functions that e.g. represent a solution. Firstly, we define the restriction operator

$$\widehat{I}_n : L^2(\Omega) \to S_n, \tag{2.34a}$$

$$u_{n,i} = (\widehat{I}_n u)_i = \frac{1}{|\Omega_{n,i}|} (R_n u)_i, \tag{2.34b}$$

with $u \in L^2(\Omega)$ an integrable function; R_n is as defined by (2.3); $|\Omega_{n,i}|$ denotes the area in $\mathbb{R}^2$ (the volume in $\mathbb{R}^3$) of the cell $\boldsymbol{i}$ at the grid Ω_n. We can interpret $u_{n,i}$ as the average value of u at cell $\boldsymbol{i}$. Secondly, we define a restriction operator

$$I_{n,n+e_k} : S_{n+e_k} \to S_n, \tag{2.35a}$$

$$u_{n,i} = \frac{|\Omega_{n+e_k,2i}|}{|\Omega_{n,i}|} u_{n+e_k,2i} + \frac{|\Omega_{n+e_k,2i+e_k}|}{|\Omega_{n,i}|} u_{n+e_k,2i+e_k}. \tag{2.35b}$$

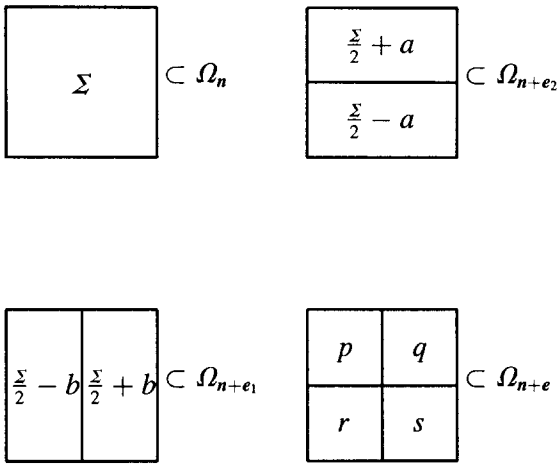

Fig. 5. Cells in the canonical case in $\mathbb{R}^2$.

Again we integrated the grid function, but we divided by the cell area (the cell volume) for each cell in order to find the average value. Then, analogously to (2.6), we can uniquely define

$$I_{n_\gamma,n_1} : S_{n_\gamma} \to S_{n_1}, \tag{2.36a}$$

$$I_{n_\gamma,n_1} = I_{n_\gamma,n_{\gamma-1}} I_{n_{\gamma-1},n_{\gamma-2}} \cdots I_{n_2,n_1} \tag{2.36b}$$

for

$$\boldsymbol{n}_\gamma < \boldsymbol{n}_{\gamma-1} < \cdots < \boldsymbol{n}_1.$$

A grid function u_m is called a *coherent* left-hand side representation of u_n, $\boldsymbol{n} > \boldsymbol{m}$, when by means of the path-independent restriction, defined by (2.35) and (2.36), it holds that

$$I_{m,n} u_n = u_m. \tag{2.37}$$

Now that we have defined coherent left- and right-hand side representations we pose and answer the question whether coherence remains after application of the operator L_n (or its inverse). Firstly, we define the following projection (a "high pass" filter):

$$H_{n,m,n} : S_n \to S_n \ (\boldsymbol{m} < \boldsymbol{n}), \tag{2.38a}$$

$$H_{n,m,n} = I_n - P_{n,m} I_{m,n}. \tag{2.38b}$$

Secondly, we define the following (operator-dependent) restriction operator

$$\tilde{I}_{n-q,n} = L_{n-q}^{-1} R_{n-q,n} L_n. \tag{2.39}$$

This restriction operator depends on the discretization operators and is of theoretical value only. We observe that, like some other restriction operators, it is the left inverse of a prolongation operator:

$$I_{n-q} = \tilde{I}_{n-q,n} P_{n,n-q}. \tag{2.40}$$

Thirdly, we define the grid functions r_{n-q} (see (2.13)) which represent residuals:

$$r_{n-q} = f_{n-q} - L_{n-q}u_{n-q}, \quad q \in \mathcal{Q}, \tag{2.41}$$

where f_{n-q} are coherent representations of f_n. We now state the following

Proposition 2.7. *Let us assume that the coarse grid discrete operators are defined by GCA* (2.24) *and that both restrictions and prolongations are path-independent;* $\mathcal{Q} = \{e_1, e_2, e\}$ *(see* (2.14)*). We assume that the inverse of* L_{n-q} *exists for all* $q \in \mathcal{Q}$.

(1) *If* $u_{n-q}, q \in \mathcal{Q}$ *are coherent (left-hand side) representations of* u_n, *then*

$$r_{n-q} - R_{n-q,n}r_n = L_{n-q}\tilde{I}_{n-q,n}H_{n,n-q,n}u_n. \tag{2.42}$$

(2) *If* $r_{n-q}, q \in \mathcal{Q}$ *are coherent (right-hand side) representations of* r_n, *then*

$$u_{n-q} - I_{n-q,n}u_n = \tilde{I}_{n-q,n}H_{n,n-q,n}u_n. \tag{2.43}$$

Proof.

(1) From the coherence of the u_{n-q} it follows by definition that

$$u_{n-q} - I_{n-q,n}u_n = 0_{n-q}.$$

The derivation of (2.42) is straightforward.

(2) From the coherence of the r_{n-q} it follows by definition that

$$r_{n-q} - R_{n-q,n}r_n = 0_{n-q}.$$

The derivation of (2.43) is straightforward. □

When we consider the set of discretizations (2.2) with $m \leqslant n$ for a complete grid of grids G_n, we deduce from Proposition 2.7 that when the grid functions f_m on the right-hand side are a coherent set, it does not follow that the corresponding solutions u_m are a coherent set; nor the other way round.

2.5. Galerkin approximations

When we disregard the use of mixed derivatives we can confine ourselves to discretization stencils not larger than five-point ones (seven-point in 3D). We consider the following typical equation:

$$Lu \equiv \left(-\varepsilon\frac{\partial^2}{\partial x_1^2} - \mu\frac{\partial^2}{\partial x_2^2} + \alpha\frac{\partial}{\partial x_1} + \beta\frac{\partial}{\partial x_2} + \sigma I\right)u = f. \tag{2.44}$$

On structured and rectangular grids in 2D the discretization that we employ, boils down to:

$$h_n h_m \begin{bmatrix} 0 & -\mu h_m^{-2} & 0 \\ -\varepsilon h_n^{-2} & 2(\varepsilon h_n^{-2} + \mu h_m^{-2}) & -\varepsilon h_n^{-2} \\ 0 & -\mu h_m^{-2} & 0 \end{bmatrix} \quad \text{(diffusion stencil)}, \tag{2.45a}$$

$$h_n h_m \begin{bmatrix} 0 & +\beta(2h_m)^{-1} & 0 \\ -\alpha(2h_n)^{-1} & 0 & +\alpha(2h_n)^{-1} \\ 0 & -\beta(2h_m)^{-1} & 0 \end{bmatrix} \quad \text{(convection stencil)}, \tag{2.45b}$$

$$h_n h_m \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{(identity stencil)} \tag{2.45c}$$

for $\boldsymbol{n} = (n, m)$. Actually, this symbolizes the discretization in use on all grids. In this section we compare (2.45) to the use of GCA (2.24). If R_{n_γ, n_1} and P_{n_1, n_γ} are path-independent then we observe that

$$L_{n_\gamma} = R_{n_\gamma, n_1} L_{n_1} P_{n_1, n_\gamma} \tag{2.46}$$

is uniquely defined (path-independent). In the particular case that

$$P_{n_1, n_\gamma} \equiv P^1_{n_1, n_\gamma}, \qquad R_{n_\gamma, n_1} \equiv R^1_{n_\gamma, n_1} \tag{2.47}$$

the Galerkin approximation generates the following stencils at the grid $\Omega^c_{(n-1,m)}$ (with $h_{n-1} = 2h_n$):

$$h_{n-1} h_m \begin{bmatrix} 0 & -\mu h_m^{-2} & 0 \\ -2\varepsilon h_{n-1}^{-2} & 2(2\varepsilon h_{n-1}^{-2} + \mu h_m^{-2}) & -2\varepsilon h_{n-1}^{-2} \\ 0 & -\mu h_m^{-2} & 0 \end{bmatrix} \quad \text{(diffusion stencil)}, \tag{2.48a}$$

$$h_{n-1} h_m \begin{bmatrix} 0 & +\beta(2h_m)^{-1} & 0 \\ -\alpha(2h_{n-1})^{-1} & 0 & +\alpha(2h_{n-1})^{-1} \\ 0 & -\beta(2h_m)^{-1} & 0 \end{bmatrix} \quad \text{(convection stencil)}, \tag{2.48b}$$

$$h_{n-1} h_m \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{(identity stencil)}. \tag{2.48c}$$

We observe that the convection stencil (2.48b) is consistent with (2.45b) and that the identity stencil (2.48c) is consistent with (2.45c), but also that the diffusion stencil (2.48a) is not consistent with (2.45a). This observation is in accordance with an accuracy condition for transfer operators that needs to be satisfied:

$$m_P + m_R > 2m, \tag{2.49}$$

(see [1,14,17,25]) where $2m$ is the order of the PDE, and m_P, m_R the highest order plus one of polynomials that are interpolated exactly by the prolongations P and sR^{T} where s is a scaling factor. When the prolongation and restriction are given by (2.47) then $m_P = m_R = 1$ and rule (2.49) is violated for the (second order) diffusion term in (2.44). For the convection (first order) and the identity term (zeroth order) the prolongation and restriction are sufficiently accurate.

In order to mend the consistency of the Galerkin approximation on the coarse grid of the second order part of equation (2.44) we may employ the same restriction operator but the prolongation operator of second order (see Section 2.3):

$$P_{n_1,n_\gamma} \equiv P^2_{n_1,n_\gamma}, \qquad R_{n_\gamma,n_1} \equiv R^1_{n_\gamma,n_1}. \tag{2.50}$$

Hereby rule (2.49) is now satisfied. By choosing (2.50), the Galerkin approximation generates the following stencils at the grid $\Omega^c_{(n-1,m)}$:

$$h_{n-1}h_m \begin{bmatrix} 0 & 0 & 0 \\ -\varepsilon h_{n-1}^{-2} & 2\varepsilon h_{n-1}^{-2} & -\varepsilon h_{n-1}^{-2} \\ 0 & 0 & 0 \end{bmatrix} \quad (x_1\text{-diffusion stencil}), \tag{2.51a}$$

$$h_{n-1}h_m \begin{bmatrix} -\frac{1}{8}\mu h_m^{-2} & -\frac{3}{4}\mu h_m^{-2} & -\frac{1}{8}\mu h_m^{-2} \\ +\frac{1}{4}\mu h_m^{-2} & \frac{3}{2}\mu h_m^{-2} & +\frac{1}{4}\mu h_m^{-2} \\ -\frac{1}{8}\mu h_m^{-2} & -\frac{3}{4}\mu h_m^{-2} & -\frac{1}{8}\mu h_m^{-2} \end{bmatrix} \quad (x_2\text{-diffusion stencil}), \tag{2.51b}$$

$$h_{n-1}h_m \begin{bmatrix} 0 & +\beta(2h_m)^{-1} & 0 \\ -\alpha(2h_{n-1})^{-1} & 0 & +\alpha(2h_{n-1})^{-1} \\ 0 & -\beta(2h_m)^{-1} & 0 \end{bmatrix} \quad (\text{convection stencil}), \tag{2.51c}$$

$$h_{n-1}h_m \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{8}\sigma & \frac{3}{4}\sigma & \frac{1}{8}\sigma \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{identity stencil}). \tag{2.51d}$$

Stencil (2.45a) is turned into the sum of stencils (2.51a) and (2.51b). We observe that for the diffusion in the x_2-direction a nine-point stencil (2.51b) comes into being. When we perform a lumping procedure for the stencil (2.51b) (thus averaging out the x_1-dependence) we observe consistency between Galerkin approximation (2.51) and discretization (2.45). A disadvantage of this central differencing type of discretization is that each time the grid is x_1-coarsened by the factor 2 the mesh Péclet number in the x_1-direction is multiplied by the same factor, which is reflected by the Galerkin coarse grid approximation (2.51) (it can be observed most clearly from the derivation below of (2.56)). For a substantial number of grid levels this may cause divergence for a multilevel algorithm, as was already observed in [7]. A remedy may be to use upwind differencing for the separate discretization on each coarse grid individually or the use of Galerkin coarse grid approximation (2.24) in connection with upwind prolongation (see e.g. [6]).

When we employ discretization (2.45) also on all coarser grids together with the prolongation as given by (2.17) and (2.20) where the said first and second order prolongation can be plugged into, then both experiments and Fourier local mode analysis [18] show that first order prolongation is sufficiently accurate even for the case of a second order PDE.

Further analysis of the Galerkin approach for constant coefficients

We perform an analysis of the behaviour of the Galerkin coarse grid approximations for the advection–diffusion equation (2.44). This analysis is the analogue of the one introduced and performed in [4] for bilinear finite elements. In the analysis we confine ourselves to constant coefficients

and therefore the matrix L_n is represented by a single nine-point stencil only. Note that here we consider a five-point stencil as a special case of a nine-point stencil. With the one choice for the restriction and the respective two choices for the prolongation (first order and second order) in the x_1-direction we obtain a coarse grid matrix L_{n-e_1} which is also represented by a nine-point stencil. Because of the constant coefficients the construction (2.24) can be seen as the linear transformation

$$Gf = c, \quad \text{with } f, c \in \mathbb{R}^9, \tag{2.52}$$

where G can be represented by a 9×9 matrix. The vectors f and c correspond with the stencils

$$f^\star \sim \begin{bmatrix} f_7 & f_8 & f_9 \\ f_4 & f_5 & f_6 \\ f_1 & f_2 & f_3 \end{bmatrix}, \qquad c^\star \sim \begin{bmatrix} c_7 & c_8 & c_9 \\ c_4 & c_5 & c_6 \\ c_1 & c_2 & c_3 \end{bmatrix}. \tag{2.53}$$

The stencil $f^\star$ is defined on the finer, the stencil $c^\star$ is defined on the coarser grid. The matrix G describes what stencil $c^\star$ is obtained on the semi-coarsened grid, from an arbitrary stencil $f^\star$, when stencil $c^\star$ is constructed by GCA, see (2.24). Let G_1 correspond to the case of (2.47) i.e. first order prolongation and G_2 correspond to the case of (2.50) i.e. second order prolongation. An eigenvalue decomposition of G_1 exists and reads:

$$G_1 = V_1 D_1 V_1^{-1}, \quad G_1, V_1, D_1 \in \mathbb{R}^9 \times \mathbb{R}^9, \tag{2.54}$$

where D_1 is a diagonal matrix showing the eigenvalues of G_1 and

$$V_1 = \begin{pmatrix} -1 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & -1 & -\frac{1}{2} & 0 \\ -1 & 0 & 0 & +\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & -1 & 0 & 0 & +\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & -1 & +\frac{1}{2} & 0 \\ 0 & 0 & -1 & 0 & 0 & +\frac{1}{2} & 0 & 0 & 0 \end{pmatrix}, \tag{2.55a}$$

$$V_1^{-1} = \begin{pmatrix} -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ -1 & 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & +1 \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -1 & -1 & -1 & 0 & 0 & 0 & +1 & +1 & +1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \tag{2.55b}$$

$$D_1 = \mathrm{diag}(\,1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 2 \quad 2 \quad 2\,). \tag{2.55c}$$

The matrix G_1 and its decomposition follow from a straightforward evaluation. The column vectors of V_1 are the right eigenvectors of G_1, the row vectors of V_1^{-1} are the left eigenvectors of G_1. When we write the column vectors of V_1 as stencils (2.53) we immediately recognize some standard central differences. Thus, by G_1^q, the Galerkin coarse grid approximation after q times coarsening in the x_1-direction is now fully described for the case of constant coefficients.

A similar eigenvalue decomposition of G_2 exists with:

$$V_2 = \begin{pmatrix}
-1 & 0 & 0 & -\frac{1}{2} & 0 & 0 & -\frac{1}{6} & -\frac{1}{12} & \frac{1}{36} \\
2 & 0 & 0 & 0 & 0 & 0 & -\frac{2}{3} & -\frac{1}{3} & \frac{1}{9} \\
-1 & 0 & 0 & +\frac{1}{2} & 0 & 0 & -\frac{1}{6} & -\frac{1}{12} & \frac{1}{36} \\
0 & -1 & 0 & 0 & -\frac{1}{2} & 0 & \frac{1}{3} & 0 & \frac{1}{9} \\
0 & 2 & 0 & 0 & 0 & 0 & \frac{4}{3} & 0 & \frac{4}{9} \\
0 & -1 & 0 & 0 & +\frac{1}{2} & 0 & \frac{1}{3} & 0 & \frac{1}{9} \\
0 & 0 & -1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{6} & \frac{1}{12} & \frac{1}{36} \\
0 & 0 & 2 & 0 & 0 & 0 & -\frac{2}{3} & \frac{1}{3} & \frac{1}{9} \\
0 & 0 & -1 & 0 & 0 & +\frac{1}{2} & -\frac{1}{6} & \frac{1}{12} & \frac{1}{36}
\end{pmatrix}, \tag{2.56a}$$

$$V_2^{-1} = \begin{pmatrix}
-\frac{1}{3} & +\frac{1}{6} & -\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{3} & +\frac{1}{6} & -\frac{1}{3} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{3} & +\frac{1}{6} & -\frac{1}{3} \\
-1 & 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & +1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & +1 \\
-\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & +\frac{1}{6} & +\frac{1}{6} & +\frac{1}{6} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\
-1 & -1 & -1 & 0 & 0 & 0 & +1 & +1 & +1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}, \tag{2.56b}$$

$$D_2 = \mathrm{diag}(\,\tfrac{1}{2} \quad \tfrac{1}{2} \quad \tfrac{1}{2} \quad 1 \quad 1 \quad 1 \quad 2 \quad 2 \quad 2\,). \tag{2.56c}$$

The multiplicity of all the different eigenvalues of both decompositions is larger than 1, so clearly these decompositions are not quite uniquely defined.

We observe that with (2.56) the mesh Péclet number in the x_1-direction is multiplied by the factor 2 and the mesh Péclet number in the x_2-direction remains unchanged; with (2.55) the mesh Péclet numbers in both x_1- and x_2-direction remain unchanged.

2.6. *Sawtooth multi-level method*

We make a deliberate choice for sawtooth multigrid. This type of multigrid employs a V-cycle without pre-relaxation, hence there is no interfering of smoothing when the residual is transferred

to subsequent coarser grids. This guarantees the coherence wished for w.r.t. the residuals (see Section 2.4). For an introduction to sawtooth multigrid see [24]. An additional advantage is that this cycle can be programmed in a simple way without recursion. Here, the algorithm is written in a fashion of the FAS and/or NMGM algorithms [2,14]. This is immaterial for the linear problems that we have under consideration, it merely indicates a possible generalization for nonlinear problems. Due to this nonlinear approach we need to store an (old) approximation of the solution on each coarser grid: $u_{\boldsymbol{n}}^{\text{old}}$. These coarse solutions are chosen to be fixed throughout execution of the algorithm. The finest grid, i.e. the grid with the highest level of refinement in each x_i-direction, is essentially the one grid of interest where we want to obtain a solution. This one grid is denoted by $\boldsymbol{n}'$ (with $\boldsymbol{n}' > \boldsymbol{0}$). We want to solve the following linear system stemming from the discretization of a PDE on the finest grid

$$L_{\boldsymbol{n}'} u_{\boldsymbol{n}'} = f_{\boldsymbol{n}'}. \tag{2.57}$$

We have some starting solution $u_{\boldsymbol{n}'}$ and employ the following scheme to improve it (we use the grids $\Omega_{\boldsymbol{n}}$ with $\boldsymbol{n} \leqslant \boldsymbol{n}'$ to accelerate convergence):

Sawtooth multi-level (*SML*).

Stage $\mathcal{A}$:

$r_{\boldsymbol{n}'} := f_{\boldsymbol{n}'} - L_{\boldsymbol{n}'} u_{\boldsymbol{n}'}$ (1)
$d_{\boldsymbol{n}'} := r_{\boldsymbol{n}'}$ (2)

`for` l `from` $|\boldsymbol{n}'| - 1$ `to` 1 `by` -1 (3)
`do` (4)
  `for all` $\boldsymbol{n} \leqslant \boldsymbol{n}'$ `with` $|\boldsymbol{n}| = l \wedge \boldsymbol{n} \geqslant \boldsymbol{0}$ (5)
  `do` (6)
    `choose an arbitrary` $\boldsymbol{e}_{j_0} \in \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ `with` $\Omega_{\boldsymbol{n}+\boldsymbol{e}_{j_0}} \neq \emptyset$ (7)
    $d_{\boldsymbol{n}} := R_{\boldsymbol{n},\boldsymbol{n}+\boldsymbol{e}_{j_0}} d_{\boldsymbol{n}+\boldsymbol{e}_{j_0}}$ (8)
    $f_{\boldsymbol{n}}^{\text{temp}} := L_{\boldsymbol{n}} u_{\boldsymbol{n}}^{\text{old}} + d_{\boldsymbol{n}}$ (9)
    $u_{\boldsymbol{n}} := u_{\boldsymbol{n}}^{\text{old}}$ (10)
  `end do` (11)
`end do` (12)

Stage $\mathcal{B}$:

`to` 2 (13)
`do` (14)
  $RELAX(L_{\boldsymbol{0}}, u_{\boldsymbol{0}}, f_{\boldsymbol{0}}^{\text{temp}})$ (15)
`end do` (16)
$c_{\boldsymbol{0}} := u_{\boldsymbol{0}} - u_{\boldsymbol{0}}^{\text{old}}$ (17)

Stage $\mathcal{C}$:

`for` l `from` 1 `to` $|\boldsymbol{n}'|$ (18)
`do` (19)
  `for all` $\boldsymbol{n} \leqslant \boldsymbol{n}'$ `with` $|\boldsymbol{n}| = l \wedge \boldsymbol{n} \geqslant \boldsymbol{0}$ (20)
  `do` (21)
    `if` $\boldsymbol{n} > \boldsymbol{0}$ `then` (22)
      $u_{\boldsymbol{n}} := u_{\boldsymbol{n}} + \sum_{\boldsymbol{q} \in Q} \omega_{\boldsymbol{q}} P_{\boldsymbol{n},\boldsymbol{n}-\boldsymbol{q}} c_{\boldsymbol{n}-\boldsymbol{q}}$ (23)

```
        else if  n_1 > 0  then                                   (24)
            u_n := u_n + P_{n,n-e_1} c_{n-e_1}                   (25)
        else                                                     (26)
            u_n := u_n + P_{n,n-e_2} c_{n-e_2}                   (27)
        end if                                                   (28)
        RELAX(L_n, u_n, f_n^temp)                                (29)
        if  l < |n'|  then                                       (30)
            c_n := u_n - u_n^old                                 (31)
        end if                                                   (32)
    end do                                                       (33)
end do                                                           (34)
```

For the meaning of $\boldsymbol{Q}$, ω_q see (2.14) and (2.17). We recognize three stages within this scheme; $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. In stage $\mathcal{A}$ the residual on the finest grid $\Omega_{n'}$ is transferred to all subsequent coarser grids. In stage $\mathcal{B}$ we determine the coarsest grid correction. In stage $\mathcal{C}$ corrections are transferred to finer grids. Post-relaxation follows. The three stages together constitute one (sawtooth) multigrid cycle.

The stages $\mathcal{A}, B$ are built up such that the various operations on the grid functions involved are partitioned per grid level l. The designation "`for all ...`" in the lines (5) and (20) of the description of the algorithm means that the ranking order is arbitrary. Consequently the operations involved are suitable for parallelization. When we use Jacobi-type iterations for *RELAX*() we can moreover vectorize each parallel process.

In the lines (7)–(8) we notice that a freedom of choice exists for the restriction of a fine grid residual onto a coarser grid. This is because the d_n ($\mathbf{0} \leqslant \boldsymbol{n} \leqslant \boldsymbol{n}'$) constitute, by definition, a coherent set of right-hand side grid functions (see also Section 2.4). E.g. when in 2D and $\boldsymbol{n} > \mathbf{0}$, the following holds:

$$R_{n,n+e_1} d_{n+e_1} = R_{n,n+e_1} R_{n+e_1,n+e} d_{n+e} = R_{n,n+e_2} R_{n+e_2,n+e} d_{n+e} = R_{n,n+e_2} d_{n+e_2}.$$

Hence we observe that in the lines (7)–(8) the grid function d_n is uniquely defined, though the result can be obtained in (two) different ways.

When we would have chosen a multigrid algorithm which includes pre-relaxation on lower levels, then because of the smoothing of the solutions in stage $\mathcal{A}$, this would involve the updating of d_{n+e_k} for $k = 1, 2$:

$$d_{n+e_k} := f^{\text{temp}}_{n+e_k} - L_{n+e_k} u_{n+e_k}$$

before the transfer to coarser grids. But then, in general, the grid function d_n is not uniquely defined anymore because

$$R_{n,n+e_1} d_{n+e_1} \neq R_{n,n+e_2} d_{n+e_2}$$

and some weighted averaging of these restricted residuals would have to be introduced for the computation of d_n. It is not clear in advance that equal weighting

$$d_n := \tfrac{1}{2} R_{n,n+e_1} d_{n+e_1} + \tfrac{1}{2} R_{n,n+e_2} d_{n+e_2}$$

(see e.g. [19,21]) would be the appropriate choice in all cases possible. E.g. consider the particular situation that $\|d_{n+e_1}\| \gg \|d_{n+e_2}\|$ due to some odd behaviour of the pre-relaxation method. Equal weighting for the restricted residuals in stage $\mathcal{A}$ then results in a too large correction for u_{n+e_2} and a too small correction for u_{n+e_1} in stage $\mathcal{C}$. Anyway, such an algorithm would become far less transparent than in its present form. Speaking in terms of coherence (see Section 2.4), we conclude that in the multi-level method without pre-relaxation we do obtain coherent representations of the finest grid residual on all levels, but with pre-relaxation we do not.

In a direct line with our approach we prefer the combination of prolongations with the weights (2.20) (see line (23) of the description of the algorithm). By our approach of handling restricted residuals and interpolated corrections we avoid that we have to determine weights for providing a way of switching to an appropriate coarse grid as in [21]. Obviously, all previous arguments for the 2D case hold for the 3D case as well.

The domain reduction method

The SML-algorithm can be conceived as an algorithm that solves the problem (2.57) in the space $S_{n'}$ in parallel in the subspaces S_n with $|\boldsymbol{n}| = l < |\boldsymbol{n}'|$ (and $\mathbf{0} \leqslant \boldsymbol{n}$) for subsequent l. I.e. we are (approximately) solving in parallel

$$R_{n,n'} L_{n'} P_{n',n} u_n = R_{n,n'} f_{n'}$$

for the above mentioned $\boldsymbol{n}$ (with zero initial guess and using GCA, see (2.24)). This shows a similarity to the domain reduction method [8–10] which uses a finite group of symmetries of the system of linear equations (2.57) to obtain a decomposition into independent subproblems, which can be solved in parallel. This decomposition involves the concept of additive subspace correction, though it is not stated as such. However, the grids are chosen quite differently from the ones in this paper and are not nested (compare to Section 2.2). Correspondingly, the weights within the additive subspace corrections are chosen differently (compare (2.17)–(2.18) to the CGC of Algorithm 3 in [8, §2]).

2.7. *The accuracy condition for grid transfer operators*

It is well known for multigrid methods with standard coarsening that we have to satisfy the accuracy condition (2.49) for grid transfer operators on penalty of lack of convergence (see e.g. [1,14,17,25]). Therefore, in the context of multiple semi-coarsening we address the question of the order of accuracy when prolongations from various grids are combined (see Section 2.3).

Proposition 2.8. *Consider the function*

$$v : \mathbb{R}^2 \to \mathbb{R},$$

$$v(x_1, x_2) = a_{00} + a_{10} x_1 + a_{01} x_2.$$

Let $v_n = \widehat{I}_n v$ *and* $v_{n-q} = \widehat{I_{n-q}} v$ *for* $\boldsymbol{q} \in \boldsymbol{Q} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}\}$.

(1) *When* (2.20) *is valid, then*

$$v_n = \sum_{q \in Q} \omega_q P^{\mathrm{I}}_{n,n-q} v_{n-q}. \tag{2.58}$$

(2) *When* (2.19) *is valid, then* (2.58) *holds* (*generally*) *only for* $a_{10} = a_{01} = 0$.

Proof. Both parts of the proof follow from straightforward evaluation for the functions 1, x, y separately. □

Remark 2.9.

(1) The first part of the proposition states that the combined piecewise constant prolongation is of second order accuracy for (2.20), the second part of the proposition states that the combined piecewise constant prolongation is of first order accuracy for (2.19).

(2) For neither (2.19) nor (2.20) functions of type

$$v(x_1, x_2) = a_{11} x_1 x_2$$

are interpolated exactly.

Along the boundary of the grid of grids (i.e. $n_1 = 0$ or $n_2 = 0$) the SML-algorithm acts differently than somewhere amidst the grid of grids (i.e. $\boldsymbol{n} > \boldsymbol{0}$). The weighted averaging of corrections stemming from different grids does not take place, simply for lack of coarse grids across the boundary of the grid of grids. Thus, a correction stems from one coarse grid only (see lines (25) and (27) of SML). When we study the transfer of defects and corrections for $n_1 = 0$, $n_2 = 1, 2, \ldots, l$ we observe that, along this boundary of the grid of grids, the SML-algorithm degenerates to multigrid for an essentially one-dimensional problem (during one sweep, results from grids with $n_1 > 0$ have no influence whatsoever). This leads to the conjecture that for elliptic problems and with the use of a restriction operator as defined in Section 2.2 we need a second order prolongation for the correction (at lines (25) and (27) of SML). However, in practice we did not perceive any difference in convergence rate for the SML-algorithm as a whole when such a second order prolongation was applied at the boundary of the grid of grids. For purely convection problems piecewise constant prolongation should be sufficiently accurate anyway.

2.8. The smoothing method

For a smoothing procedure within the sawtooth multi-level procedure we employ damped pointwise Jacobi(α) relaxation

$$u_n^{\text{new}} = \alpha D_n^{-1}(f_n - (L_n - D_n)u_n^{\text{old}}) + (1 - \alpha)u_n^{\text{old}} \tag{2.59}$$

with α the damping parameter and D_n the main diagonal of the discrete operator L_n. The amplification matrix $J_n(\alpha)$ for the error reads

$$e_n^{\text{new}} = J_n(\alpha)e_n^{\text{old}} = (I_n - \alpha D_n^{-1} L_n)e_n^{\text{old}}. \tag{2.60}$$

This simple procedure is not fit for standard MG-methods, see e.g. [25, § 7.6]. In the case of anisotropy for standard MG-methods one needs to resort to linewise relaxation or incomplete factorization for smoothing in order to obtain satisfactory MG-convergence. For MG-methods which use multiple semi-coarsening such as SML, we may expect that damped Jacobi becomes an appropriate smoother again. When washboard functions are present in the error they cannot hamper convergence as with standard multigrid for now they are resolved on the semi-coarsened grids. Indeed, two-level

Fourier analysis, see [18], shows that this is the case. For $\alpha = \frac{1}{2}$ the Jacobi method is known to annihilate completely the pure chess-board component in the error. Hemker [18] favours $\alpha = \frac{2}{3}$ due to Fourier analysis. When we apply two subsequent Jacobi(α) relaxation sweeps we can use different α_1, α_2 as damping parameters. A proper choice of such a combination of two different α's may prove to be more effective than the application of two Jacobi(α)-sweeps with one and the same α. We propose to use as subsequent values $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{2}{3}$. This proposal is examined in the next subsection.

2.9. Numerical results

As test problem we consider the following

Model problem

$$Lu \equiv (-\varepsilon\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2})u = f \text{ on } \Omega,$$

$$1 \geqslant \varepsilon \geqslant 0,$$

$$\Omega \equiv (0,1) \times (0,1)$$

with periodic boundary conditions and right-hand side

$$\text{case (a): } f = +\delta_{(a,b)} - \delta_{(b,a)},$$

$$\text{case (b): } f = +\delta_{(a,b)} + \delta_{(b,a)} - \delta_{(a,a)} - \delta_{(b,b)}$$

where

$$a, b \in \mathbb{R},\ a = \frac{1}{8} - \nu,\ b = \frac{7}{8} - \nu,\ \nu > 0$$

with ν a small positive real number. We treat the diffusion coefficient ε as a parameter that determines the degree of anisotropy in our test problems. The δ_x denotes the well-known Dirac δ_x-distribution which can be defined by

$$\langle \delta_x, \varphi \rangle = \varphi(x), \quad x \in \mathbb{R}^2,$$

see e.g. [23]. We may consider the Dirac distribution δ_x as a function with a support around x that vanishes and with an integral that is equal to 1. In this way we model sinks and sources in the right-hand side f of the model problem.

After discretization in the manner of (2.45a) the corresponding stencils on the grids $\Omega_{m,n}$ read as follows:

$$L_{(n,m)} = \begin{bmatrix} 0 & -2^{m-n} & 0 \\ -\varepsilon 2^{n-m} & \varepsilon 2^{n-m+1} + 2^{m-n+1} & -\varepsilon 2^{n-m} \\ 0 & -2^{m-n} & 0 \end{bmatrix}. \tag{2.61}$$

Also if $\varepsilon = 1$ we observe an anisotropic appearance of this stencil as soon as $n \neq m$. Because ν is a small positive number, the sinks and sources are not located at either boundary of the cells but well in the inside.

First we discuss the (discretized) model problem. Solvability requires that $f \in \mathcal{R}(L)$, i.e. the right-hand side f is within the range of the operator L (in the discrete case $f_n \in \mathcal{R}(L_n)$ for f_n in S_n). For $\varepsilon > 0$ both case (a) and case (b) satisfy this requirement. In the discrete case the requirement boils down to that the sum of the elements of f_n needs to vanish. For $\varepsilon = 0$ only case (b) is solvable. For this value of ε a complete decoupling takes place of the solution in the x_1-direction. Here, in the discrete case, the requirement boils down to that the sum of the elements of f_n needs to vanish along each individual grid line with $x_1 = \textit{constant}$. Only case (b) satisfies this requirement. Of course, with Dirichlet boundary conditions, case (a) is solvable as well. The foregoing demonstrates that for $\varepsilon \downarrow 0$ case (a) becomes a difficult test problem to solve.

Apart from these preliminary remarks we note that iterative methods will converge less easily for the model problem (both case (a) and case (b)) than for the same problem with Dirichlet boundary conditions instead where only a substantial lower number of frequency components can hamper the convergence rate.

Description of results

At grid level $l = 12$ we perform experiments for the sawtooth multi-level procedure with $\boldsymbol{n}' = (6,6), (7,5), (8.4), (9,3), (10,2), (11,1)$ respectively (for each such $\boldsymbol{n}'$ the number of grid points is 4096). For prolongation and restriction we fix upon (2.47). For *RELAX*() in SML we use two damped Jacobi iterations, namely Jacobi(α_1), Jacobi(α_2) with $\alpha_1 = \alpha_2 = \frac{1}{2}$. The convergence histories for the different $\boldsymbol{n}'$ are shown in Fig. 6. Along the horizontal axis the number of SML-cycles is written, along the vertical axis the 10-logarithm of the maximum norm of the residual $r_{\boldsymbol{n}'}$. The test problem is the model problem with $\varepsilon = 1$, and f defined by case (a).

We repeat the experiments, but now instead of $l = 12$ (Fig. 6) we perform at grid level $l = 14$, see Fig. 7. In this way we can check, grid by grid, whether the convergence rates do not slow down after reducing the mesh size in both directions. Indeed, in this sense the convergence rates turn out to be perfectly grid independent.

The foregoing experiments (see Figs. 6 and 7) are repeated but for different values of the α, namely the combination $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{2}{3}$. The results are shown in Figs. 8 and 9. We observe a general improvement of the convergence rate, and less variance between convergence rates at different grids $\Omega_{\boldsymbol{n}'}$ on the same grid level $l = |\boldsymbol{n}'|$. Experiments were also performed for $\alpha_1 = \alpha_2 = \frac{2}{3}$ (results not shown). However, the results with $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{2}{3}$ exhibit far better convergence rates.

The next experiment involves the anisotropic ($1 > \varepsilon > 0$) case of model problem (a). We use SML with the combination $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{2}{3}$ as damping parameters for Jacobi. We perform experiments for fixed $\boldsymbol{n} = (7,7)$ and vary the x_1-diffusion coefficient from 1 to 10^{-10}. For decreasing ε the solution develops increasing gradients. The respective convergence histories are given by Fig. 10. The convergence behaviour proves to be satisfying. The sudden flattening of the curves for small ε after several iterations is due to the finite machine precision and the growing magnitude of the solution which is inversely proportional to ε.

The last experiment is repeated for right-hand side (b). For this right-hand side the solution does not grow with decreasing ε. In Fig. 11 we observe that the convergence behaviour is without complication and independent of ε, however small.

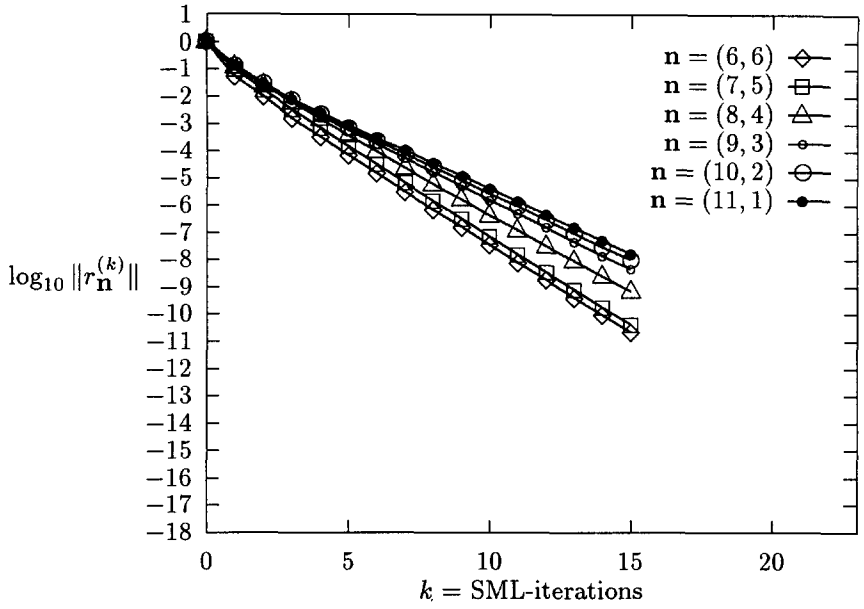

Fig. 6. Convergence history of SML for Poisson ($\varepsilon = 1$), right-hand side (a), periodic b.c., $|\boldsymbol{n}| = 12$, $\alpha_1 = \alpha_2 = \frac{1}{2}$.

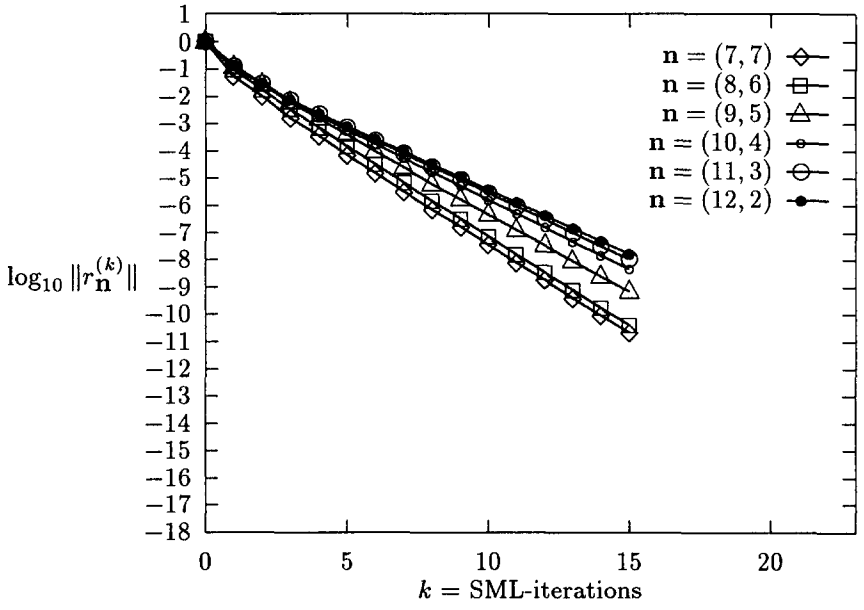

Fig. 7. Convergence history of SML for Poisson ($\varepsilon = 1$), right-hand side (a), periodic b.c., $|\boldsymbol{n}| = 14$, $\alpha_1 = \alpha_2 = \frac{1}{2}$.

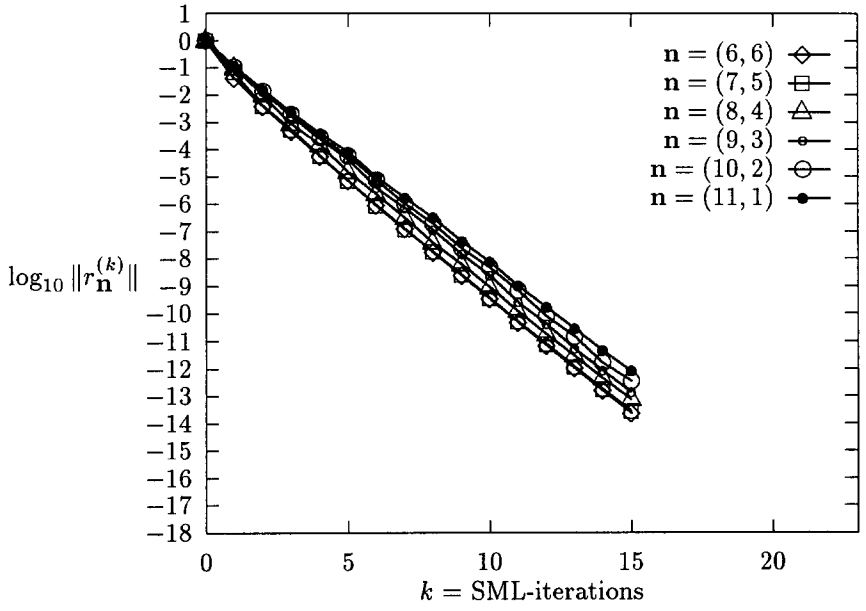

Fig. 8. Convergence history of SML for Poisson ($\varepsilon = 1$), right-hand side (a), periodic b.c., $|\boldsymbol{n}| = 12$, $\alpha_1 = \frac{1}{2}$; $\alpha_2 = \frac{2}{3}$.

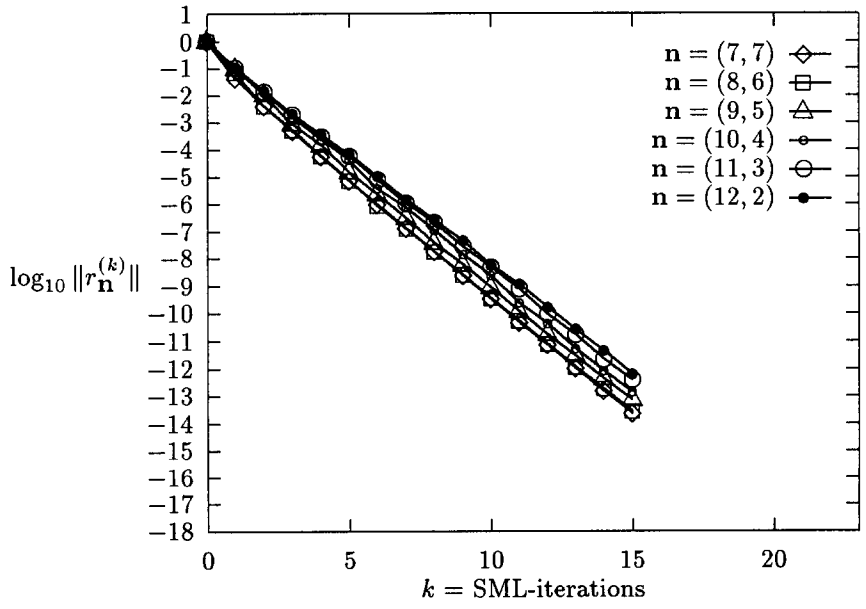

Fig. 9. Convergence history of SML for Poisson ($\varepsilon = 1$), right-hand side (a), periodic b.c., $|\boldsymbol{n}| = 14$, $\alpha_1 = \frac{1}{2}$; $\alpha_2 = \frac{2}{3}$.

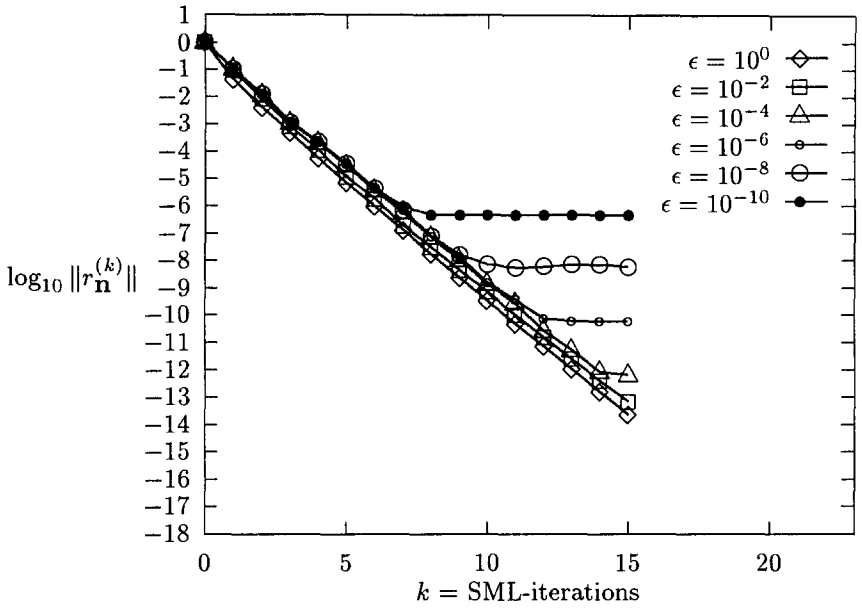

Fig. 10. Convergence history of SML for anisotropic Poisson (varying ε), right-hand side (a), periodic b.c., $\boldsymbol{n} = (7,7)$, $\alpha_1 = \frac{1}{2}$; $\alpha_2 = \frac{2}{3}$.

3. Incomplete grid of grids

In Section 2.6 we described the SML-algorithm which is designed for a complete grid of grids. In this section we discuss briefly the possibility of an algorithm suitable for an incomplete grid of grids of the second kind, namely Zenger's sparse grids [26]. We define the latter by

$$Z_\Lambda = \{\Omega_{\boldsymbol{m}} \mid |\boldsymbol{m}| \leqslant \Lambda \wedge \boldsymbol{m} \geqslant \boldsymbol{0}\}, \quad \Lambda \in \mathbb{N}, \tag{3.1}$$

see Fig. 12. Λ is called the highest grid level. Note that here we consider finite volumes rather than finite elements.

3.1. Hierarchical basis for finite volumes

Zenger [26], Griebel [12] and Bungartz [3] applied a hierarchical basis for sparse grids in the context of finite element methods. Using a variational formulation this leads for linear elliptic PDEs to linear systems that can be solved efficiently by cycling sequentially through the sparse grid of grids, see [13].

Here we formulate an orthogonal set of hierarchical basis functions which might be used in the search for the finite-volume counterpart of the said approach. First we define the function

$$\varphi : \mathbb{R} \to \mathbb{R}, \tag{3.2a}$$

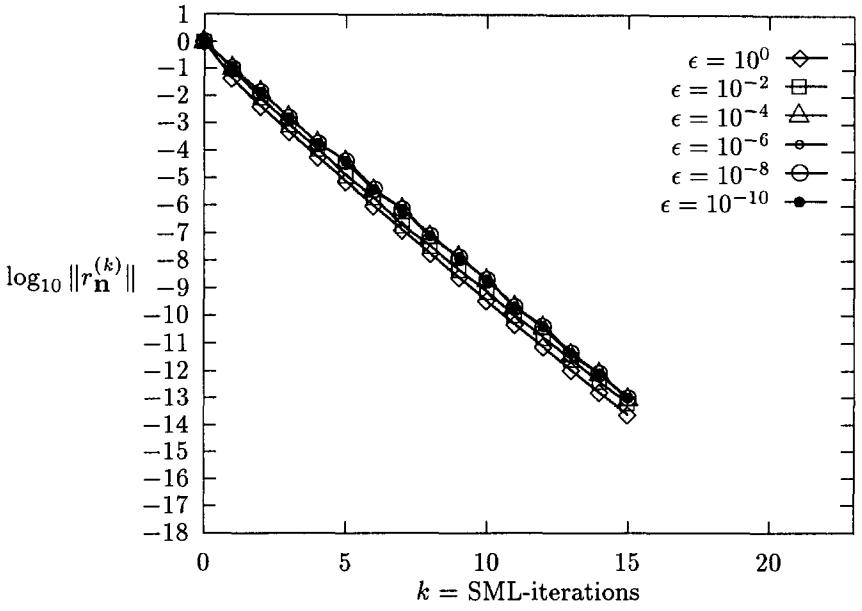

Fig. 11. Convergence history of SML for anisotropic Poisson (varying ε), right-hand side (b), periodic b.c., $\boldsymbol{n} = (7,7)$, $\alpha_1 = \frac{1}{2}$; $\alpha_2 = \frac{2}{3}$.

$$\varphi(x) = \begin{cases} -1, & \text{if } 0 < x < \frac{1}{2}, \\ +1, & \text{if } \frac{1}{2} < x < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.2b}$$

Next we define the function

$$\psi_{(0,0)}^{(0,0)} : \Omega \to \mathbb{R}, \tag{3.3a}$$

$$\psi_{(0,0)}^{(0,0)}(\boldsymbol{x}) = 1. \tag{3.3b}$$

Further we define

$$\psi_{(n_1,0)}^{(i,0)} : \Omega \to \mathbb{R}, \tag{3.4a}$$

$$\psi_{(n_1,0)}^{(i,0)}(\boldsymbol{x}) = \varphi(2^{n_1-1}x_1 - i), \quad i = 0, \ldots, (2^{n_1-1} - 1) \tag{3.4b}$$

for $n_1 \in \mathbb{N}$ and

$$\psi_{(0,n_2)}^{(0,j)} : \Omega \to \mathbb{R}, \tag{3.5a}$$

$$\psi_{(0,n_2)}^{(0,j)}(\boldsymbol{x}) = \varphi(2^{n_2-1}x_2 - j), \quad j = 0, \ldots, (2^{n_2-1} - 1) \tag{3.5b}$$

for $n_2 \in \mathbb{N}$. Using definitions (3.3), (3.4), (3.5) we now define the functions

$$\psi_{(n_1,n_2)}^{(i,j)} : \Omega \to \mathbb{R}, \tag{3.6a}$$

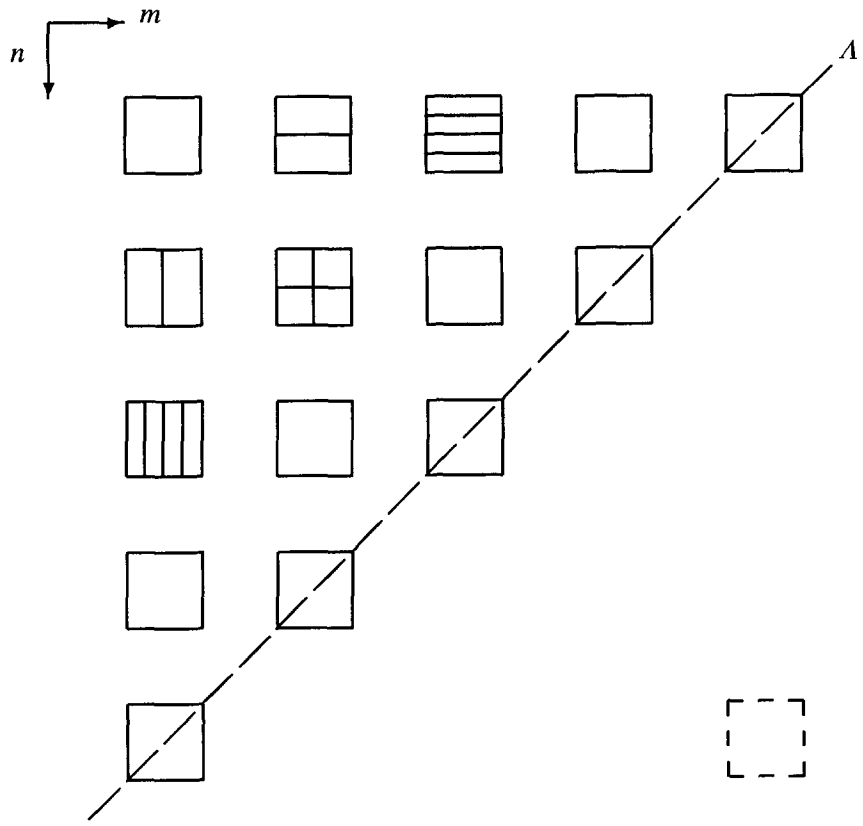

Fig. 12. Z_Λ: sparse grids in $\mathbb{R}^2$.

$$\psi^{(i,j)}_{(n_1,n_2)}(x) = \psi^{(i,0)}_{(n_1,0)}(x) \cdot \psi^{(0,j)}_{(0,n_2)}(x), \quad i = 0,\ldots,(2^{n_1-1}-1);\ j = 0,\ldots,(2^{n_2-1}-1) \tag{3.6b}$$

for $(n_1, n_2) \in \mathbb{N}^2$. We define the space

$$W_n(\Omega) = \{\psi^{(i,j)}_n\}, \quad i = 0,\ldots,(2^{n_1-1}-1);\ j = 0,\ldots,(2^{n_2-1}-1). \tag{3.7}$$

Finally, we define the space of hierarchical basis functions on Z_Λ:

$$V_\Lambda(\Omega) = \{\psi^{(0,0)}_{(0,0)}\} \oplus \bigoplus_{n\neq 0,|n|\leqslant\Lambda} W_n. \tag{3.8}$$

4. Conclusions

We examined the feasibility of multi-level methods based on multiple semi-coarsening. Within a grid of grids we have seen examples of path-independent prolongations and restrictions. This leads also to path-independent Galerkin coarse grid approximations (GCA) for the discretizations. For a discretized linear second order elliptic PDE with constant coefficients in two space dimensions the outcome of GCA was fully analysed. For a (standard) second order prolongation it turned out that with each GCA coarsening the corresponding mesh Péclet number is multiplied by the factor 2. We applied additive subspace correction by weighted averaging of the corrections for the solution which stem from multiple semi-coarsened grids. Various choices are proposed by different authors, yet all fitted within the same framework. One such choice appeared to increase the order of accuracy of the underlying separate 1D prolongations. We introduced the notion of coherence: it shows a relation that may hold between grid functions that represent the same continuous function. When we consider the discretizations on a grid of grids, we showed that coherent grid functions at the right-hand side do not imply coherent solutions, nor the other way round. We formulated a sawtooth multi-level algorithm (SML) which relies on simple Jacobi smoothing and additive subspace correction by multiple semi-coarsening. This algorithm is amenable to parallelization and vectorization. For SML averaging of residuals at coarser grids is not material (and superfluous). Coherence appears to be important for convergence of SML. For discrete problems with anisotropic stencils SML showed a satisfactory and grid independent convergence, as had been predicted by two-level Fourier analysis. Especially the usage of Jacobi with alternating damping parameters $\frac{1}{2}$ and $\frac{2}{3}$ exhibited good convergence rates.

Multi-level methods for sparse grids within the context of finite elements already exist in the literature. The finite-volume counterpart of such methods may be subject to future research. For this purpose a set of hierarchical basis functions within the context of finite volumes was formulated.

Acknowledgement

The referees of this article are gratefully acknowledged for their constructive comments.

References

[1] A. Brandt, Multi-level adaptive techniques (MLAT) for partial differential equations: ideas and software, in: J. Rice, ed., *Mathematical Software* (Academic Press, New York, 1977) 277–318.

[2] A. Brandt, Guide to multigrid development, in: W. Hackbusch and U. Trottenberg, eds., *Multigrid Methods*, Lecture Notes in Mathematics 960 (Springer, Berlin, 1982) 220–312.

[3] H.J. Bungartz, Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung, Dissertation, Fakultät für Mathematik und Informatik, Technische Universität München, München (1992).

[4] P.M. de Zeeuw, Matrix-dependent prolongations and restrictions in a blackbox multigrid solver, *J. Comput. Appl. Math.* 33 (1990) 1–27.

[5] P.M. de Zeeuw, Nonlinear multigrid applied to a one-dimensional stationary semiconductor model, *SIAM J. Sci. Statist. Comput.* 13 (1992) 512–530.

[6] P.M. de Zeeuw, Multigrid and advection, in: C.B. Vreugdenhil and B. Koren, eds., *Numerical Methods for Advection-Diffusion Problems*, Notes on Numerical Fluid Mechanics 45 (Vieweg, Braunschweig, 1993) Ch. 14, 335–351.

[7] P.M. de Zeeuw and E.J. van Asselt, The convergence rate of multi-level algorithms applied to the convection-diffusion equation, *SIAM J. Sci. Statist. Comput.* 6 (1985) 492–503.

[8] C.C. Douglas, A tupleware approach to domain decomposition methods, *Appl. Numer. Math.* 8 (1991) 353–373.
[9] C.C. Douglas and J. Mandel, An abstract theory for the domain reduction method, *Computing* 48 (1992) 73–96.
[10] C.C. Douglas and B.F. Smith, Using symmetries and antisymmetries to analyze a parallel multigrid algorithm: the elliptic boundary value problem case, *SIAM J. Numer. Anal.* 26 (1989) 1439–1461.
[11] M. Griebel, A parallelizable and vectorizable multi-level algorithm on sparse grids, in: W. Hackbusch, ed., *Parallel Algorithms for Partial Differential Equations*, Notes on Numerical Fluid Mechanics 31 (Vieweg, Braunschweig, 1991) 94–100.
[12] M. Griebel, M. Schneider and C. Zenger, A combination technique for the solution of sparse grid problems, in: R. Beauwens and P. de Groen, eds., *Iterative Methods in Linear Algebra* (North-Holland, Amsterdam, 1992) 263–281.
[13] M. Griebel, C. Zenger and S. Zimmer, Multilevel Gauss–Seidel-algorithms for full and sparse grid problems, *Computing* 50 (1993) 127–148.
[14] W. Hackbusch, *Multi-Grid Methods and Applications* (Springer, Berlin, 1985).
[15] W. Hackbusch, A new approach to robust multigrid solvers, in: J. McKenna and R. Temam, eds., *ICIAM '87* (SIAM, Philadelphia, PA, 1988) 114–126.
[16] W. Hackbusch and S. Hagemann, Frequency decomposition multi-grid methods for hyperbolic problems, in: J. Ballmann and R. Jeltsch, eds., *Nonlinear Hyperbolic Equations: Theory, Computation Methods and Applications*, Notes on Numerical Fluid Mechanics 24 (Vieweg, Braunschweig, 1989) 209–217.
[17] P.W. Hemker, On the order of prolongations and restrictions in multigrid procedures, *J. Comput. Appl. Math.* 32 (1990) 423–429.
[18] P.W. Hemker, Remarks on sparse-grid finite-volume multigrid, Report NM-R9421, CWI, Amsterdam (1994).
[19] W.A. Mulder, A new multigrid approach to convection problems, *J. Comput. Phys.* 83 (1989) 303–323.
[20] W.A. Mulder, A high-resolution Euler solver based on multigrid, semi-coarsening, and defect correction, *J. Comput. Phys.* 100 (1992) 91–104.
[21] N.H. Naik and J. Van Rosendale, The improved robustness of multigrid elliptic solvers based on multiple semi-coarsened grids, *SIAM J. Numer. Anal.* 30 (1993) 215–229.
[22] U. Rüde, Multilevel, extrapolation, and sparse grid methods, in: P.W. Hemker and P. Wesseling, eds., *Multigrid Methods IV*, International Series of Numerical Mathematics 116 (Birkhäuser, Basel, 1994) 281–294.
[23] L. Schwartz and D. Huet, *Méthodes Mathématiques pour les Sciences Physiques* (Hermann, Paris, 1969).
[24] P. Wesseling, A robust and efficient multigrid method, in: W. Hackbusch and U. Trottenberg, eds., *Multigrid Methods*, Lecture Notes in Mathematics 960 (Springer, Berlin, 1982) 614–630.
[25] P. Wesseling, *An Introduction to Multigrid Methods* (Wiley, Chichester, 1991).
[26] C. Zenger, Sparse grids, in: W. Hackbusch, ed., *Parallel Algorithms for Partial Differential Equations*, Notes on Numerical Fluid Mechanics 31 (Vieweg, Braunschweig, 1991) 241–251.