Composite Finite Elements and Multi-Grid Part I: Convergence Theory in 1-d

N. Frauböse and S.A. Sauter
Universität Zürich
Institut für Mathematik
Winterthurerstrasse 190
CH-Zürich

May 2001

Abstract

We will study the convergence of multi-grid methods for solving linear systems as they arise from finite element discretisations of elliptic boundary value problems on complicated domains. *Composite finite elements* are employed for the construction of the sequence of coarse-level discretisations, where the minimal dimension of the coarsest linear system is very small, independent of the number and size of geometric details in the domain. The convergence of the corresponding multi-grid method is proved in the framework of geometric multi-grid methods while the emphasis is on the *robustness* with respect to the geometric details in the domain.

Introduction

In this paper, we will present a multi-grid algorithm which is based on a discretisation with so-called composite finite elements. Composite finite elements have been developed for coarse level discretisations of partial differential equations (PDEs) on complicated domains (see [HS97c], [HS96], [HS97a], [HS97b]). They can be considered as a bridge between purely algebraic multi-level methods (see, e.g. [RS87], [Bra95]), where only the given system of linear equations is required as an input, and purely geometric multi-grid methods [Hac85], which are based on a hierarchy of nested finite element meshes of the domain. This new multi-grid method requires as an input

This work was partially supported by the Swiss National Foundation under grant 21-058891.99.

- the underlying boundary value problem,
- the algebraic system of linear equations corresponding to a finite element discretisation, and
- the link between this discretisation and the geometry of the domain, e.g., the meaning of the unknowns.

In [FHSW01], the algorithmic aspects of this new multi-grid method have been presented. Here, its convergence is proved in the framework of geometric multi-grid method, i.e., by employing the so-called *smoothing* and *approximation property.* Various other approaches for applying multi-level methods to elliptic boundary value problems on complicated domains exist in the literature (cf. [BX96], [Xu96], [KY94]). They are formulated in the setting of *multi-level* methods (cf. [Xu92]). For a comparison of these methods, we refer to [BHW98].

As a model problem we consider, as a prototype of an elliptic boundary value problem with smooth coefficients, the operator $Lu = -\Delta u + u$ with mixed boundary conditions . We assume that the discretisation, i.e., the finite element mesh and the system of equations, is *given* while a *hierarchy* of coarse scale discretisations is not at hand. *Composite finite elements* in the context of multi-grid methods have been introduced in [FHSW01] and we focus here on the convergence analysis. We will prove for a one-dimensional model problem that the multi-grid method converges *robustly* with respect to some parameter describing the geometry of the domain and the finite element meshes.

The paper is organised as follows. In Section 1, we will introduce the model problem and the corresponding finite element discretisation. The multi-grid method based on composite finite elements is presented in Section 2, while its convergence is analysed in Section 3.

1 Model Problem & Discretisation

Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal Lipschitz domain with possibly many geometrical details. The boundary of Ω is denoted by $\Gamma := \partial\Omega$. The usual Sobolev space is denoted by $H^k(\Omega)$ and we put $V := H^1(\Omega)$. The weak formulation of the model problem is given by seeking, for given $f \in L^2(\Omega)$, the function $u \in V$ such that

$$a(u, v) = \int_\Omega f v \, \mathrm{d}x \qquad \forall\, v \in V\,, \tag{1}$$

where the bilinear form $a : V \times V \to \mathbb{R}$ is defined by

$$a(u,v) := \int_\Omega \langle \nabla u, \nabla v \rangle + uv \, \mathrm{d}x \, .$$

The discretisation is based on a conforming finite element mesh

$$\mathcal{G} = \{\tau_1, \tau_2, \ldots, \tau_N\}$$

in the sense of Ciarlet [Cia78]. Let S denote the space of continuous, piecewise affine finite elements

$$S := \{u \in V \,|\, \forall \tau \in \mathcal{G} : u|_\tau \text{ is affine}\} \, .$$

The finite element discretisation of problem (1) is given by finding $u_S \in S$ with

$$a(u_S, v) = \int_\Omega fv \, \mathrm{d}x \qquad \forall \, v \in S \, .$$

The set of nodal points - all vertices of triangles in $\mathcal{G}$ - is denoted by Θ and the corresponding local nodal basis by φ_x, $x \in \Theta$. The link between a finite element function and its coefficient vector is given by the prolongation operator

$$P : \mathbb{R}^\Theta \to S, \quad P[\mathbf{u}](x) = \sum_{z \in \Theta} \mathbf{u}(z) \varphi_z(x) \, .$$

Thus, the coefficient vector $\mathbf{u}$ is the solution of

$$\mathbf{A}_S \mathbf{u}_S = \mathbf{f}_S \tag{2}$$

with

$$\left.\begin{array}{lcl} \mathbf{A}_S(x,y) & := & a(\varphi_x, \varphi_y) \\ \mathbf{f}_S(x) & := & \int_\Omega f \varphi_x \, \mathrm{d}x \end{array}\right\} \qquad \forall \, x, y \in \Theta \, .$$

Typically, the linear system (2) is very large and sparse so that multi-grid solvers should be applied for its solution. The efficiency of multi-grid methods strongly relies on a given hierarchy of coarse-level discretisations. However, in the case of very complicated domains such a hierarchy is *not* available in an obvious way. We will apply composite finite elements for constructing appropriate coarse-level systems. In the following, we will prove that the convergence rate of the resulting multi-grid method is independent of the number and sizes of the geometric details. Here, we will restrict to a one-dimensional model problem while further research is devoted to the multi-dimensional case.

2 Multi-Grid Algorithm

A multi-grid algorithm is based on a multi-scale discretisation of the boundary value problem. It combines an iterative solver (called *smoother*) on each discretisation level with a recursive coarse grid correction. The parameter $\ell \in \mathbb{N}$ with $0 \leq \ell \leq L$ describes the discretisation level, where L corresponds to the finest level. The number of levels $L+1$ is not known *a priori*. For ease of notation we rewrite (2) as

$$\mathbf{A}_L \mathbf{u}_L = \mathbf{f}_L \,.$$

Analogously, the finite element space S is denoted, from now on, by S_L and its basis by $\varphi_{L,x}$.

2.1 Abstract Multi-Grid Algorithm

The definition of a sequence of coarse grid finite element spaces $\{S_\ell\}_{0\leq\ell\leq L-1}$ and of appropriate discrete inter-grid transfer operators

$$\mathbf{p}_{\ell,\ell-1} : \mathbb{R}^{\Theta_{\ell-1}} \to \mathbb{R}^{\Theta_\ell}$$

plays the key rôle for a multi-grid algorithm (The explicite definitions will be given in section 2.3.). The transposed operator of $\mathbf{p}_{\ell,\ell-1}$ defines the restriction $\mathbf{r}_{\ell-1,\ell} : \mathbb{R}^{\Theta_\ell} \to \mathbb{R}^{\Theta_{\ell-1}}$. Given the prolongation $\mathbf{p}_{\ell,\ell-1}$, the coarse grid matrices $\mathbf{A}_\ell$, for $\ell < L$, are defined recursively via the Galerkin product

$$\mathbf{A}_\ell := \mathbf{r}_{\ell,\ell+1} \mathbf{A}_{\ell+1} \mathbf{p}_{\ell+1,\ell} \,.$$

In order to define the multi-grid algorithm one has to specify an iterative solver on each level. Here, we restrict to linear solvers of the form

$$\mathbf{u}_\ell^{(i+1)} := \mathbf{u}_\ell^{(i)} - \mathbf{M}_\ell \left(\mathbf{A}_\ell \mathbf{u}_\ell^{(i)} - \mathbf{f}_\ell \right) \,. \tag{3}$$

The ν-fold application of this solver defines the mapping $\mathbf{S}_\ell^{(\nu)} \left(\mathbf{u}_\ell^{(i)}, \mathbf{f}_\ell \right) := \mathbf{u}_\ell^{(i+\nu)}$. The multi-grid algorithm is a recursive procedure which requires as an input parameter $\nu \in \mathbb{N}$ specifying the number of smoothing steps and a parameter $\gamma \in \{1, 2\}$ controlling whether a V- or W-cycle is applied (for details refer to [Hac85]).

For the realisation of the multi-grid algorithm, the definition of a hierarchy of coarse scale discretisations is essential.

2.2 Generation of a Coarse Grid Hierarchy

This subsection is devoted to the construction of a hierarchy of coarse scale grids $(\mathcal{G}_\ell)_{\ell=0}^{L-1}$ from a given fine grid $\mathcal{G}_L$. These grids $\mathcal{G}_\ell$ will be nested for $0 \leq \ell \leq L-1$ but, possibly, do not resolve the geometric details of the domain Ω. However, the construction guarantees that the grid $\mathcal{G}_{L-1}$ has a similar (slightly coarser) distribution of mesh cells as $\mathcal{G}_L$. Let Ω_0 denote a bounding box of Ω and $\mathcal{G}_0 = \{\tau_0, \tau_1, \ldots, \tau_{n_0}\}$ a minimal conforming finite element grid of Ω_0 (in the sense of Ciarlet [Cia78]). Finer meshes $\mathcal{G}_\ell$ are generated by recursively refining the initial mesh $\mathcal{G}_0$. The refinement process is controlled by the distribution of mesh cells in $\mathcal{G}_L$. A mesh cell $\tau \in \mathcal{G}_\ell$ is marked for refinement, if more than three distinct elements of $\mathcal{G}_L$ are contained in τ. Since we apply conforming subdivisions of elements as, e.g., connecting midpoints of edges, either $t \subset \tau$ or $(\operatorname{int} t) \cap (\operatorname{int} \tau) = \emptyset$ holds for any $\tau \in \mathcal{G}_\ell$ and $t \in \mathcal{G}_{\ell+1}$, $\ell + 1 < L$. To avoid hanging nodes the *green-closure algorithm* (cf.[BS81]) is applied. The refinement stops if all mesh cells contain at most three triangles of the given grid. Triangles having no intersection with the domain are removed from these grids. For a detailed description of the algorithm and numerical experiments, we refer to [FHSW01].

The domain covered by the grid $\mathcal{G}_\ell$ is denoted by Ω_ℓ and Θ_ℓ denotes the set of vertices of triangles in $\mathcal{G}_\ell$. The definition of auxiliary grids $(\mathcal{G}_\ell)_{\ell=0}^{L-1}$ implies for all $\ell < L$

$$\Omega_{\ell+1} \subset \Omega_\ell \qquad \text{and} \qquad \Theta_{\ell+1} \subset \overline{\Omega_\ell}\,.$$

Remark 2.1 *The algorithm adapts the auxiliary grids to $\mathcal{G}_L$ even if the given grid is not quasi-uniform.*

Remark 2.2 *The algorithm can be generalised to three-dimensional problems in a straightforward manner.*

2.3 Composite Finite Elements

In this section, the sequence of coarse grid finite element spaces and the appropriate transfer operators for the multi-grid algorithm will be defined. The auxiliary meshes $\mathcal{G}_\ell$ do not necessarily resolve the geometry of Ω for $\ell < L$ and, hence, standard finite elements cannot be applied. Therefore, we will introduce composite finite element spaces based on an iterated prolongation.

Let $\varphi_{\ell,x}$ denote the standard continuous, piecewise affine Lagrange basis on $\mathcal{G}_\ell$. For any grid function $\mathbf{u} \in \mathbb{R}^{\Theta_\ell}$, we associate a finite element function on the *overlapping domain* Ω_ℓ by

$$P_\ell[\mathbf{u}](x) := \sum_{z \in \mathbb{R}^{\Theta_\ell}} \mathbf{u}(z)\varphi_{\ell,z}(x)\,.$$

Due to $\Theta_{\ell+1} \subset \overline{\Omega_\ell}$, the function $P_\ell[\mathbf{u}]$ can be evaluated at the grid points $\Theta_{\ell+1}$ of the finer mesh. In this light, the inter-grid prolongation $\mathbf{p}_{\ell+1,\ell} : \mathbb{R}^{\Theta_\ell} \to \mathbb{R}^{\Theta_{\ell+1}}$ is defined by

$$\mathbf{p}_{\ell+1,\ell}[\mathbf{u}](x) := P_\ell[\mathbf{u}](x)\,, \qquad x \in \Theta_{\ell+1}\,,$$

and the matrix representation is

$$\mathbf{p}_{\ell+1,\ell} \in \mathbb{R}^{\Theta_{\ell+1}\times\Theta_\ell} : \ \mathbf{p}_{\ell+1,\ell}(x,y) = P_\ell[\varphi_{\ell,y}](x)$$

for all $x \in \Theta_{\ell+1}$ and $y \in \Theta_\ell$. The restriction is the transposed of $\mathbf{p}_{\ell+1,\ell}$, i.e.,

$$\mathbf{r}_{\ell,\ell+1} \in \mathbb{R}^{\Theta_\ell\times\Theta_{\ell+1}} : \ \mathbf{r}_{\ell,\ell+1}(x,y) = \mathbf{p}_{\ell+1,\ell}(y,x)\,.$$

Due to the nestedness of the auxiliary grids the composite mapping

$$\mathbf{p}_{L,\ell} := \mathbf{p}_{L,L-1}\mathbf{p}_{L-1,L-2}\cdots\mathbf{p}_{\ell+1,\ell}$$

has the representation

$$\mathbf{p}_{L,\ell} : \mathbb{R}^{\Theta_\ell} \to \mathbb{R}^{\Theta_L}, \qquad \mathbf{p}_{L,\ell}[\mathbf{u}](x) = P_\ell[\mathbf{u}](x) \qquad \forall\, x \in \Theta_L\,.$$

For any coarse grid vector $\mathbf{u} \in \mathbb{R}^{\Theta_\ell}$, the associated *composite finite element function* is given by

$$u = P_{L,\ell}[\mathbf{u}] := P_L\left[\mathbf{p}_{L,\ell}[\mathbf{u}]\right] = \sum_{z\in\mathbb{R}^{\Theta_L}} \left(P_\ell[\mathbf{u}](z)\right)\varphi_{L,z}\,.$$

Since the finest grid resolves the domain, i.e., $\Omega_L = \Omega$, any composite finite element function is a function on the physical domain Ω and contained in the corresponding standard finite element space S. We emphasise that the overlapping domains Ω_ℓ have only auxiliary character.

Definition 2.3 *The space of Composite Finite Elements is the range of* $P_{L,\ell}$

$$S_{L,\ell} := \operatorname{range} P_{L,\ell} := P_{L,\ell}\left(\mathbb{R}^{\Theta_\ell}\right)\,.$$

The definition of composite finite elements by using inter-grid prolongations has geometric and algebraic character. The construction of the auxiliary grids is based on geometric properties while these grids are only used for setting up purely algebraic prolongation operators.

2.4 Multi-Grid Algorithm

Our aim is to solve the given system of linear equations $\mathbf{A}_L\mathbf{u}_L = \mathbf{f}_L$ via a multi-grid algorithm. For $0 \le \ell \le L$, consider the problem

$$\mathbf{A}_\ell\mathbf{u}_\ell = \mathbf{f}_\ell\,.$$

In a first phase, a few steps of an iterative solver (smoothing iteration) of the form (3) are applied yielding an approximation $\tilde{\mathbf{u}}_\ell$ of $\mathbf{u}_\ell$. Denote the error by $\mathbf{e}_\ell := \tilde{\mathbf{u}}_\ell - \mathbf{u}_\ell$ and observe

$$\mathbf{A}_\ell\mathbf{e}_\ell = \mathbf{A}_\ell\tilde{\mathbf{u}}_\ell - \mathbf{f}_\ell =: \mathbf{d}_\ell\,. \tag{4}$$

The principle of multi-grid methods is to exploit the smoothness of the error $\mathbf{e}_\ell$ and solve (4) on a coarser grid. Put $\mathbf{d}_{\ell-1} := \mathbf{r}_{\ell-1,\ell}\mathbf{d}_\ell$ and define $\mathbf{e}_{\ell-1}$ as the solution of

$$\mathbf{A}_{\ell-1}\mathbf{e}_{\ell-1} = \mathbf{d}_{\ell-1}\,. \tag{5}$$

An approximate solution of (4) is given by $\mathbf{p}_{\ell,\ell-1}\mathbf{e}_{\ell-1}$ and is used to correct to approximation $\tilde{\mathbf{u}}_\ell$

$$\tilde{\mathbf{u}}_\ell \mapsto \tilde{\mathbf{u}}_\ell - \mathbf{p}_{\ell,\ell-1}\mathbf{e}_{\ell-1}.$$

A two-grid algorithm is the combination of such a coarse grid correction and the smoothing iteration. The multi-grid method is the recursive application of the two-grid algorithm to the *defect equation* (5) on coarser levels and can be written in the following form (cf. [Hac93]):

```
procedure MGM(u_ℓ, f_ℓ, ν);
if ℓ = 0 then u_0 := A_0^{-1} f_0 else
begin for i := 1 to ν do u_ℓ := S_ℓ(u_ℓ, f_ℓ);
      d_{ℓ-1} := r_{ℓ-1,ℓ}(A_ℓ u_ℓ - f_ℓ);
      e_{ℓ-1}^{(0)} := 0;
      for i := 1 to γ do e_{ℓ-1}^{(i)} := MGM(e_{ℓ-1}^{(i-1)}, d_{ℓ-1}, ν);
      u_ℓ := u_ℓ - p_{ℓ,ℓ-1} e_{ℓ-1}^{(γ)};
end;
```

The choice $\gamma = 1$ results in the so-called V-cycle multi-grid method while $\gamma = 2$ corresponds to the W-cycle (cf. [Hac85]). The two-grid iteration matrix $\mathbf{K}^{TGM}_{\ell,\ell-1}$ is given by

$$\mathbf{K}^{TGM}_{\ell,\ell-1} := \left(\mathbf{A}_\ell^{-1} - \mathbf{p}_{\ell,\ell-1}\mathbf{A}_{\ell-1}^{-1}\mathbf{r}_{\ell-1,\ell}\right)\mathbf{A}_\ell\mathbf{K}_\ell^\nu \tag{6}$$

with the iteration matrix $\mathbf{K}_\ell := \mathbf{I} - \mathbf{M}_\ell\,\mathbf{A}_\ell$ of the smoother.

3 Convergence Analysis

The main objective of this paper is the investigation of the convergence of the multi-grid method based on non-nested grids. The idea is to adapt the general multi-grid convergence theory of [Hac85] to our specific situation and to prove the so-called smoothing and the approximation property. The proof of these properties differs for composite finite element discretisations on overlapping and non-nested grids significantly from the situation considered in [Hac85] for the following reasons.

1. The prolongation which was introduced in Section 2.3 is the canonical finite element prolongation on the levels $0 \dots L-1$. The prolongation from the finest auxiliary level to the given discretisation is defined on the non-nested grids $\mathcal{G}_{L-1}$, $\mathcal{G}_L$ and some of the assumptions in [Hac85] are violated. Hence, for the investigation of the effect of the prolongation on non-nested grids it suffices to consider a *two-grid* method ($\mathcal{G}_{L-1} \to \mathcal{G}_L$).

2. Triangles at the boundary, possibly, have only small overlap with the domain Ω. As a consequence the matrix entries which correspond to nodes lying (essentially) outside the domain have very different scaling compared to those entries corresponding to interior nodes. The stability of the method has to be investigated with respect to these boundary effects.

We recall the basic ingredients of the convergence theory and develop the necessary modifications.

3.1 Convergence of a Two-Grid Algorithm

The two-grid iteration converges if and only if the spectral radius $\varrho(\mathbf{K}_{\ell,\ell-1}^{TGM})$ of the two-grid iteration-matrix is smaller than 1.

Let $\langle \cdot, \cdot \rangle_{\mathcal{U}}$ be a scalar product on $\mathbb{R}^{\Theta_\ell}$ and denote the corresponding norm by

$$\|\mathbf{u}\|_{\mathcal{U}}^2 := \langle \mathbf{u}, \mathbf{u} \rangle_{\mathcal{U}} \qquad \forall \mathbf{u} \in \mathbb{R}^{\Theta_\ell} .$$

Put $\mathcal{U} := \left(\mathbb{R}^{\Theta_\ell}, \|\cdot\|_{\mathcal{U}}\right)$. The spectral radius $\varrho(\mathbf{K}_{\ell,\ell-1}^{TGM})$ is defined as the maximal eigenvalue of $\mathbf{K}_{\ell,\ell-1}^{TGM}$ and can be estimated by

$$\varrho(\mathbf{K}_{\ell,\ell-1}^{TGM}) \leq \|\mathbf{K}_{\ell,\ell-1}^{TGM}(\nu)\|_{\mathcal{U}\leftarrow\mathcal{U}} .$$

The convergence proof is based on a multiplicative splitting of the two-grid iteration matrix

$$\|\mathbf{K}_{\ell,\ell-1}^{TGM}(\nu)\|_{\mathcal{U}\leftarrow\mathcal{U}} = \|\mathbf{A}_\ell^{-1} - \mathbf{p}_{\ell,\ell-1}\mathbf{A}_{\ell-1}^{-1}\mathbf{r}_{\ell-1,\ell}\|_{\mathcal{U}\leftarrow\mathcal{U}} \, \|\mathbf{A}_\ell \mathbf{K}_\ell^\nu\|_{\mathcal{U}\leftarrow\mathcal{U}} .$$

The *approximation property* holds if there is a constant $C_A > 0$ and a number $\alpha \in \mathbb{R}$ such that

$$\|\mathbf{A}_\ell^{-1} - \mathbf{p}_{\ell,\ell-1}\mathbf{A}_{\ell-1}^{-1}\mathbf{r}_{\ell-1,\ell}\|_{\mathcal{U}\leftarrow\mathcal{U}} \leq C_A\, h_\ell^\alpha \qquad \forall\, \ell \geq 1. \tag{7}$$

$\mathbf{K}_\ell^\nu$ is said to possess the *smoothing property* if there exist numbers $0 < \underline{\nu} \leq \overline{\nu}$ and $0 < \beta < 1$ such that

$$\|\mathbf{A}_\ell \mathbf{K}_\ell^\nu\|_{\mathcal{U}\leftarrow\mathcal{U}} \leq \frac{\beta}{C_A h_\ell^\alpha} \qquad \forall\, \underline{\nu} \leq \nu \leq \overline{\nu}\ , \ell \geq 1\,,$$

with α and C_A as in (7).

Note that the definition of the approximation and smoothing property slightly differs from that in [Hac85]. It will be appropriate to prove the convergence of the presented multi-grid method with respect to a certain parameter related to the geometry of Ω.

3.2 The Effect of the Small Overlap of Triangles with the Domain Near the Boundary

In [FHSW01], it was proved that, for quasi-uniform and shape-regular meshes, the multi-grid method converges although the prolongation from the finest auxiliary level to the given discretisation is defined on non-nested grids *provided* the mesh $\mathcal{G}_{L-1}$ has a similar but slightly coarser structure compared to the given mesh. For the details, we refer to [FHSW01].

Here, we will study the effect that, on coarse levels, large triangles might have an arbitrary small overlap with the domain and the degrees of freedom lying (essentially) outside of the domain lead to a very different scaling of the corresponding matrix entries compared to the interior unknowns. We will study the influence of this scaling effect on the convergence rate of a multi-grid algorithm for a one-dimensional model problem.

3.2.1 One-Dimensional Model Problem

Let $\Omega = (0,1)$ and set $V := \{u \in H^1(\Omega) : u(1) = 0\}$. Consider the problem to find, for given $f \in L^2(\Omega)$, a function $u \in V$ so that

$$a(u,v) = \int_\Omega fv \,\mathrm{d}x\,, \quad \forall\, v \in V \quad \text{with} \quad a(u,v) := \int_\Omega u'v' \,\mathrm{d}x\,.$$

Let $(\mathcal{G}_\ell)_{0\leq\ell\leq L}$ denote a sequence of overlapping grids. The grid points are numbered from left to right and we assume

$$x_{\ell,0} < 0 < x_{\ell,1} < x_{\ell,2} < \cdots < x_{\ell,n_\ell} = 1\,.$$

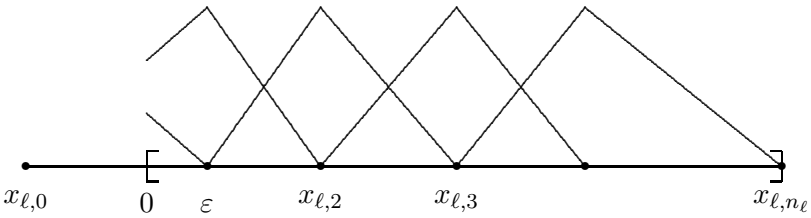

Figure 1: Overlapping grid and corresponding basis functions. The two left-most basis functions are the restrictions of standard basis functions on the overlapping grid to the domain.

The intervals are denoted by $\tau_{\ell,i} := (x_{\ell,i-1}, x_{\ell,i})$ and their lengths by $h_{\ell,i} := x_{\ell,i} - x_{\ell,i-1}$. The maximal step size is $h_\ell := \max_{1\le i\le n_\ell} h_{\ell,i}$. We assume that the grids are quasi-uniform, i.e., there exists a moderate constant C_u such that, for all $0 \le \ell \le L$ and all $1 \le i \le n_\ell$, the estimate $h_\ell \le C_u h_{\ell,i}$ holds. Furthermore, we assume that the grids are generated from an initial grid $\mathcal{G}_0$ by introducing midpoints to the intervals of coarser triangulations. More precisely, for all $0 \le \ell \le L-1$, we assume (cf. Figure 1)

$$x_{\ell+1,2i} = \frac{x_{\ell,i+1} + x_{\ell,i}}{2} \quad \text{and} \quad x_{\ell+1,2i+1} = x_{\ell,i+1} \qquad \forall\, 0 \le i \le 2^{\ell-1}(n_0 - 1)\,.$$

A direct consequence is that $x_{\ell,1} =: \varepsilon$, for all $0 \le \ell \le L$. We shall regard the (small) overlap ε of the first interval $\tau_{\ell,1}$ with the domain as a parameter and investigate the behaviour of the multi-grid convergence as $\varepsilon \to 0$. More precisely, we assume

$$0 < \varepsilon < \alpha h_{\ell+1,1} < 1 \qquad \forall\, 0 \le \ell \le L-1\,, \tag{8}$$

with $0 < \alpha < 1$ fixed, and therefore,

$$0 < \varepsilon < \alpha h_{L,1} < 1\,.$$

As a consequence the overlap of coarse intervals $\tau_{\ell,0}$ with the domain might be arbitrary small.

We use the notation $A \lesssim B$ if there is a constant C depending only on α and C_u such that $A \le CB$. Similarly, $A \gtrsim B$ if $B \lesssim A$ and $A \sim B$ if $A \lesssim B$ and $B \lesssim A$. We will prove that the multi-grid algorithm converges uniformly with respect to L, ℓ, h_ℓ and ε.

For our model problem, the composite finite element space S_ℓ is simply the restriction of the standard finite element space on the overlapping grid $\mathcal{G}_\ell$ to $(0,1)$

$$S_\ell = \left\{ u \in C^0\left(\overline{\Omega}\right) \mid \forall\, \tau \in \mathcal{G}_\ell :\, u|_\tau \text{ is affine} \right\} \cap V\,.$$

The finite element discretisation is given by seeking $u_\ell \in S_\ell$ such that

$$a(u_\ell, v_\ell) = \int_\Omega f v_\ell \,\mathrm{dx}\,, \qquad \forall\, v_\ell \in S_\ell\,.$$

Due to the Dirichlet boundary conditions, the degrees of freedom are associated with the set of grid points $\Theta_\ell := \{x_{\ell,i} : 0 \le i \le n_\ell - 1\}$. As a basis of S_ℓ we choose the usual nodal basis $\{\varphi_{\ell,i}\}_{i=0}^{n_\ell-1}$ (*hat functions*) of the overlapping grid, where the two left-most basis functions are restricted to the domain Ω (cf. Figure 1). Applying the basis representation of u_ℓ and testing with the basis function lead to the system of linear equations

$$\mathbf{A}_\ell \mathbf{u}_\ell = \mathbf{F}_\ell\,,$$

with

$$(\mathbf{A}_\ell)_{i,j} := a\left(\varphi_{\ell,i}, \varphi_{\ell,j}\right)\,, \qquad (\mathbf{F}_\ell)_i := \int_\Omega f\varphi_{\ell,i}\,\mathrm{d}x \qquad \forall\, 0 \le i,j \le n_\ell - 1\,.$$

For simplicity, we restrict to the damped Jacobi iteration with damping factor $\omega \in (0,1)$ as a smoother. Let $\mathbf{D}_\ell$ denote the diagonal part of $\mathbf{A}_\ell$. Then, the iteration matrix is given by $\mathbf{K}_\ell := \mathbf{I} - \omega \mathbf{D}_\ell^{-1}\mathbf{A}_\ell$ and the iteration can be written in the form

$$x^{(i+1)} = x^{(i)} - \omega \mathbf{D}_\ell^{-1}\left(\mathbf{A}_\ell x^{(i)} - \mathbf{F}_\ell\right)\,.$$

3.2.2 The Convergence Rate of a Two-Grid Method

First, we analyse the convergence of a two-grid algorithm with respect to the levels ℓ, $\ell - 1$. As the inter-grid transfer operator we use the canonical finite element prolongation (cf. [Hac85, Sec. 3.6]) $\mathbf{p}_{\ell,\ell-1} : \mathbb{R}^{\Theta_{\ell-1}} \to \mathbb{R}^{\Theta_\ell}$ and, as the restriction $\mathbf{r}_{\ell-1,\ell}$, its transposed. The two-grid iteration matrix with ν (pre-)smoothing steps is given by

$$\mathbf{K}_{\ell,\ell-1}^{TGM} := \left(\mathbf{A}_\ell^{-1} - \mathbf{p}_{\ell,\ell-1}\mathbf{A}_{\ell-1}^{-1}\mathbf{r}_{\ell-1,\ell}\right)\mathbf{A}_\ell\mathbf{K}_\ell^\nu\,.$$

Our aim is to prove the two-grid convergence by investigating the smoothing and the approximation property with respect to appropriate norms. Numerical results show, that the Euklidean norm is *not* appropriate because $\|\mathbf{K}_{\ell,\ell-1}^{TGM}\|_{0,\ell\leftarrow 0,\ell}$ diverges as $\varepsilon \to 0$, while the spectral radius is bounded away from 1 independent of ε. Below, we will prove the robust convergence as $\varepsilon \to 0$ in a weighted norm, where the left-most unknown (lying essentially outside Ω) is damped. For $s \in \mathbb{R}$, let

$$\mathbf{N}_{s,\ell} := h_\ell \,\mathrm{diag}\,[q_\ell^{3-s}, 1, 1, \ldots, 1]\,,$$

with $q_\ell := q_\ell(\varepsilon) := \varepsilon/h_{\ell,1}$. The appropriate scalar product and norm on $\mathbb{R}^{\Theta_\ell}$ are given by

$$\langle \mathbf{u}, \mathbf{v} \rangle_{0,s,\ell} = \mathbf{u}^\intercal \mathbf{N}_{s,\ell} \mathbf{v} \qquad \text{and} \qquad \|\mathbf{u}\|_{0,s,\ell} := \langle \mathbf{u}, \mathbf{u} \rangle_{0,s,\ell}^{1/2} .$$

The indices $0, s$ characterise the correspondence of the norm $\|\cdot\|_{0,s,\ell}$ to a weighted L^2-norm.

Remark 3.1 *It can be shown that the norms $\|P_{L,\ell}(\cdot)\|_{L^2(\Omega)}$ and $\|\cdot\|_{0,0,\ell}$ are equivalent.*

For any vector $\mathbf{u} \in \mathbb{R}^m$, the usual Euklidean norm is denoted by $\|\mathbf{u}\|_0^2 := \sum_{i=1}^m |\mathbf{u}_i|^2$ and the associated matrix norm by $\|\cdot\|_{0\leftarrow 0}$.

Explicit calculations yield that the system matrix $\mathbf{A}_\ell$ has the following representation

$$\mathbf{A}_\ell = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \tilde{\mathbf{A}}_\ell & \\ 0 & & \end{bmatrix} + \frac{q_\ell}{h_{\ell,1}} \left[\begin{array}{ccc|c} 1 & -1 & 0 & \mathbf{0} \\ (K_\ell)_{21} & 0 & (K_\ell)_{23} & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] ,$$

where $\tilde{\mathbf{A}}_\ell$ denotes the finite element discretisation of a reduced Poisson problem

$$\begin{aligned} -u'' &= f \qquad \text{in } (\varepsilon, 1) , \\ u'(\varepsilon) = u(1) &= 0 \end{aligned} \tag{9}$$

on the sub-mesh $\tilde{\mathcal{G}}_\ell := \{\tau \in \mathcal{G}_\ell : \tau \subset (\varepsilon, 1)\}$.

Lemma 3.2 *For all $\gamma \in (0,1)$ there exist $0 < \underline{\omega} < \overline{\omega} < 1$, $1 < \underline{\nu} < \overline{\nu}$, and $\overline{q} > 0$ depending only on α, C_u (cf. (8)) and γ such that*

$$\|\mathbf{A}_\ell \mathbf{K}_\ell^\nu\|_{0,1\leftarrow 0,1} \le \gamma h_\ell \qquad \forall\, q \in]0, \overline{q}] , \quad \nu \in [\underline{\nu}, \overline{\nu}] , \quad \omega \in [\underline{\omega}, \overline{\omega}] .$$

Proof. To simplify the notation we skip the index ℓ in the proof of this lemma.

From the definition of the norm $\|\cdot\|_{0,s}$ it follows

$$\|\mathbf{A}\mathbf{K}^\nu\|_{0,1\leftarrow 0,1} = \|\mathbf{N}_1^{1/2} \mathbf{A}\mathbf{K}^\nu \mathbf{N}_1^{-1/2}\|_{0\leftarrow 0} .$$

The *ansatz* is to use an appropriate multiplicative splitting of this matrix

$$\begin{aligned} \mathbf{N}_1^{1/2} \mathbf{A}\mathbf{K}^\nu \mathbf{N}_1^{-1/2} &= \mathbf{N}_1^{1/2} \mathbf{A} \left(\mathbf{I} - \omega \mathbf{D}^{-1} \mathbf{A}\right)^\nu \mathbf{N}_1^{-1/2} \\ &= \mathbf{N}_1^{1/2} \mathbf{D}^{1/2} \mathbf{X} \left(\mathbf{I} - \omega \mathbf{X}\right)^\nu \mathbf{D}^{1/2} \mathbf{N}_1^{-1/2} \\ &= \mathbf{N}_1^{1/2} \mathbf{D}^{1/2} \mathbf{N}_1^{-k} \mathbf{W} \left(\mathbf{I} - \omega \mathbf{W}\right)^\nu \mathbf{N}_1^{k} \mathbf{D}^{1/2} \mathbf{N}_1^{-1/2} \end{aligned}$$

with $\mathbf{X} := \mathbf{D}^{-1/2}\mathbf{A}\,\mathbf{D}^{-1/2}$ and $\mathbf{W} := \mathbf{N}_1^k\mathbf{X}\,\mathbf{N}_1^{-k}$. Therefore,

$$\begin{aligned}\|\mathbf{A}\mathbf{K}^\nu\|_{0,1\leftarrow 0,1} \le \|\mathbf{N}_1^{1/2}\mathbf{D}^{1/2}\mathbf{N}_1^{-k}\|_{0\leftarrow 0}\cdot\|\mathbf{W}\,(\mathbf{I}-\omega\mathbf{W})^\nu\,\|_{0\leftarrow 0}\\ \cdot\|\mathbf{N}_1^k\mathbf{D}^{1/2}\mathbf{N}_1^{-1/2}\|_{0\leftarrow 0}\,.\end{aligned}$$

Explicit calculations for the first and the last factor yield

$$\|\mathbf{A}\mathbf{K}^\nu\|_{0,1\leftarrow 0,1} \le C_0 h\|\mathbf{W}\,(\mathbf{I}-\omega\mathbf{W})^\nu\,\|_{0\leftarrow 0}\,, \tag{10}$$

for $k = 1/4$ while C_0 is independent of ℓ and q.

It remains to estimate $\|\mathbf{W}\,(\mathbf{I}-\omega\mathbf{W})^\nu\,\|_{0\leftarrow 0}$, with $\mathbf{W} = \mathbf{N}_1^{1/4}\mathbf{X}\mathbf{N}_1^{-1/4}$. We have

$$\mathbf{W} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \tilde{\mathbf{D}}^{-1/2}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-1/2} & \\ 0 & & \end{bmatrix} + \left[\begin{array}{ccc|c} 1 & W_{12} & 0 & \\ W_{21} & 0 & W_{23} & \mathbf{0} \\ 0 & W_{32} & 0 & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right]$$

with $|W_{12}| \sim |W_{23}| \sim |W_{32}| \sim q$ and $|W_{21}| \sim 1$. Let $\mathbf{e}_i$ denote the ith canonical unit vector in $\mathbb{R}^{n-1}$. We apply the splitting

$$\begin{aligned}\mathbf{W} &= \begin{bmatrix} 1 & \mathbf{0} \\ W_{21}\mathbf{e}_1 & \tilde{\mathbf{W}} \end{bmatrix} + \left[\begin{array}{ccc|c} 0 & W_{12} & 0 & \\ 0 & 0 & W_{23} & \mathbf{0} \\ 0 & W_{32} & 0 & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \\ &=: \mathbf{W}_I + \mathbf{W}_{II}\end{aligned} \tag{11}$$

with $\tilde{\mathbf{W}} := \tilde{\mathbf{D}}^{-1/2}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-1/2}$. This leads to

$$\mathbf{W}\,(\mathbf{I}-\omega\mathbf{W})^\nu = \mathbf{W}_I\,(\mathbf{I}-\omega\mathbf{W}_I+\omega\mathbf{W}_{II})^\nu + \mathbf{W}_{II}\,(\mathbf{I}-\omega\mathbf{W})^\nu\,.$$

The first term on the right-hand side above can be rewritten as a sum of terms of the form

$$\mathbf{W}_I\prod_{i=1}^{t}(\mathbf{I}-\omega\mathbf{W}_I)^{\alpha_i}\,(\omega\mathbf{W}_{II})^{\beta_i} \tag{12}$$

with $\alpha_i, \beta_i \in \mathbb{N}_0$ and $|\alpha| + |\beta| = \nu$. Here, $|\mu| := \sum_{i=1}^t \mu_i$. To avoid the identity matrix in between of the factors in (12) we may assume, w.l.o.g.,

$$\alpha_{i+1}, \beta_i > 0 \qquad \forall\, 1 \le i \le t-1\,.$$

There exists only one summand with $|\beta| = 0$, namely

$$\mathbf{W}_I\,(\mathbf{I}-\omega\mathbf{W}_I)^\nu\,.$$

This leads to the estimate

$$\begin{aligned}\|\mathbf{W}\,(\mathbf{I}-\omega\mathbf{W})^{\nu}\|_{0\leftarrow 0} \leq\ & \|\mathbf{W}_I\,(\mathbf{I}-\omega\mathbf{W}_I)^{\nu}\|_{0\leftarrow 0} \\ & + \sum_{j=0}^{\nu-1}\binom{\nu}{j}\,\|\mathbf{I}-\omega\mathbf{W}_I\|_{0\leftarrow 0}^{j}\,\|\omega\mathbf{W}_{II}\|_{0\leftarrow 0}^{\nu-j} \\ & + \|\mathbf{W}_{II}\|_{0\leftarrow 0}\,\|\mathbf{I}-\omega\mathbf{W}\|_{0\leftarrow 0}^{\nu}\end{aligned} \tag{13}$$

and the single factors and sums on the right-hand side in (13) will be considered below.

We begin with the estimate of $\mathbf{W}_I\,(\mathbf{I}-\omega\mathbf{W}_I)^j$ and $(\mathbf{I}-\omega\mathbf{W}_I)^j$. Explicit calculations yield

$$(\mathbf{I}-\omega\mathbf{W}_I)^j = (1-\omega)^j \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{c}^{(j)} & \tilde{\mathbf{B}}^j \end{bmatrix}$$

and

$$\mathbf{W}_I\,(\mathbf{I}-\omega\mathbf{W}_I)^j = (1-\omega)^j \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{b}^{(j)} & \tilde{\mathbf{W}}\tilde{\mathbf{B}}^j \end{bmatrix}, \tag{14}$$

with $\tilde{\mathbf{B}} := (1-\omega)^{-1}\left(\mathbf{I}-\omega\tilde{\mathbf{W}}\right)$ and

$$\mathbf{b}^{(j)} := W_{21}\mathbf{e}_1 - \frac{\omega W_{21}}{1-\omega}\tilde{\mathbf{W}}\sum_{m=0}^{j-1}\tilde{\mathbf{B}}^m\mathbf{e}_1\,, \qquad \mathbf{c}^{(j)} := -\frac{\omega}{1-\omega}W_{21}\sum_{i=0}^{j-1}\tilde{\mathbf{B}}^i\mathbf{e}_1\,.$$

Since the matrix $\tilde{\mathbf{W}}$ is symmetric its spectral radius equals $\|\tilde{\mathbf{W}}\|_{0\leftarrow 0}$. The quasi-uniformity of the reduced grid $\tilde{\mathcal{G}}$ implies that the L^2-norm of $\tilde{\mathbf{W}}$ is bounded by $C_{\tilde{\mathbf{W}}} \sim 1$.

It is well known that $\tilde{\mathbf{A}}$ is positive definite implying that $\tilde{\mathbf{W}} = \tilde{\mathbf{D}}^{-1/2}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-1/2}$ is positive definite as well. The eigenvectors of $\tilde{\mathbf{B}}$ coincide with those of $\tilde{\mathbf{W}}$ and the eigenvalues of $\tilde{\mathbf{B}}$ are of the form $(1-\omega\lambda)/(1-\omega)$ with $\lambda \in \operatorname{spec}\left(\tilde{\mathbf{W}}\right)$. Thus, the L^2-norm of $\tilde{\mathbf{B}}$ is bounded by $(1-\omega)^{-1}$ for all $0 < \omega \leq \bar{\omega} \leq 1/C_{\tilde{\mathbf{W}}}$.

From [Hac85, Lemma 1.3.5], we conclude that for any $0 < \underline{\omega} \leq \bar{\omega}$ and $\omega \in [\underline{\omega}, \bar{\omega}]$,

$$\begin{aligned}\|\tilde{\mathbf{W}}\tilde{\mathbf{B}}^j\|_{0\leftarrow 0} &= \left\|\tilde{\mathbf{W}}\frac{1}{(1-\omega)^j}(\mathbf{I}-\omega\tilde{\mathbf{W}})^j\right\|_{0\leftarrow 0} \\ &\leq \frac{1}{(1-\omega)^j}\max\{f_j(\lambda) = \lambda(1-\omega\lambda)^j\,:\,\lambda\in\operatorname{spec}(\tilde{\mathbf{W}})\}\,.\end{aligned}$$

The maximum above is bounded by

$$\max\{f_j(\lambda)\,:\,\lambda\in\operatorname{spec}(\tilde{\mathbf{W}})\} \leq \frac{1}{\omega\,(1+j)}\,,$$

i.e.,

$$\|\tilde{\mathbf{W}}\tilde{\mathbf{B}}^j\|_{0\leftarrow 0} \leq \frac{1}{\omega\,(1-\omega)^j\,(j+1)}\,.$$

Next, the L^2-norm of the vector $\mathbf{b}^{(j)}$ in (14) is estimated. Let $(\mu_i)_{i=1}^{n-1}$ denote the (orthonormal) eigenvectors of $\tilde{\mathbf{W}}$ and expand $\mathbf{e}_1$ according to

$$\mathbf{e}_1 = \sum_{i=1}^{n-1} \alpha_i \mu_i\,.$$

Thus,

$$\mathbf{b}^{(j)} = W_{21}\mathbf{e}_1 - \frac{\omega W_{21}}{1-\omega} \sum_{i=1}^{n-1} \sum_{m=0}^{j-1} \alpha_i \lambda_i \left(\frac{1-\omega\lambda_i}{1-\omega}\right)^m \mu_i$$

and, for the L^2-norm, one gets

$$\begin{aligned}
\|\mathbf{b}^{(j)}\|_0^2 &\lesssim \|\mathbf{e}_1\|_0^2 + \left\| \frac{\omega}{1-\omega} \sum_{i=1}^{n-1} \sum_{m=0}^{j-1} \frac{\alpha_i}{(1-\omega)^m} |f_m(\lambda_i)| \mu_i \right\|_0^2 \\
&= 1 + \left(\frac{\omega}{1-\omega}\right)^2 \sum_{i=1}^{n-1} \alpha_i^2 \left(\sum_{m=0}^{j-1} \frac{|f_m(\lambda_i)|}{(1-\omega)^m} \right)^2 \\
&\lesssim 1 + \left(\frac{\omega}{1-\omega}\right)^2 \sum_{i=1}^{n-1} \alpha_i^2 \left(\sum_{m=0}^{j-1} \frac{1}{\omega\,(1+m)\,(1-\omega)^m} \right)^2 .
\end{aligned}$$

Straightforward analysis yields

$$\sum_{m=0}^{j-1} (1-\omega)^{j-m} \frac{j+1}{m+1} \leq \frac{1}{\omega} \sum_{m=0}^{j-1} (1-\omega)^{j-m} (1+\omega)^{j-m} \leq \frac{1-\omega^2}{\omega^3}\,.$$

Using this estimate and $\sum \alpha_i^2 = 1$, the norm of $\mathbf{b}^{(j)}$ is bounded by

$$\|\mathbf{b}^{(j)}\|_0 \lesssim \frac{1}{\omega^3\,(j+1)\,(1-\omega)^j}\,,$$

where the constant of equivalence is independent of ω and j. In a similar fashion, one proves

$$\|\mathbf{c}^{(j)}\|_0 \lesssim \frac{1}{(1-\omega)^j}\,.$$

It follows that

$$\left\|\mathbf{W}_I\left(\mathbf{I}-\omega\mathbf{W}_I\right)^j\right\| \lesssim (1-\omega)^j\left(1+\left\|\mathbf{b}^{(j)}\right\|+\left\|\tilde{\mathbf{W}}\tilde{\mathbf{B}}^j\right\|\right) \lesssim \frac{1}{\omega^3\,(j+1)}\,,$$
$$\left\|\left(\mathbf{I}-\omega\mathbf{W}_I\right)^j\right\| \lesssim 1 \qquad \text{and} \qquad \left\|\mathbf{W}_{II}^j\right\| \le (C_{II}q)^j\,. \tag{15}$$

It remains to estimate the norm $\|\mathbf{I}-\omega\mathbf{W}\|_{0\leftarrow 0}^{\nu}$ in (13). We obtain

$$\|\mathbf{W}\|_{0\leftarrow 0} \le \|\mathbf{W}_I\|_{0\leftarrow 0}+\|\mathbf{W}_{II}\|_{0\leftarrow 0} \le C\left(1+\left\|\tilde{\mathbf{W}}\right\|_{0\leftarrow 0}+q\right) =: C_{\mathbf{W}}\,,$$

where $C_{\mathbf{W}}>0$ only depends on α and C_u (cf. (8)). Thus,

$$\overline{\omega} := \min\left\{\frac{2}{C_{\mathbf{W}}}, C_{\tilde{\mathbf{W}}}^{-1}\right\} \tag{16}$$

leads to

$$\|\mathbf{I}-\omega\mathbf{W}\|_{0\leftarrow 0}^{\nu} \le 1\,.$$

Combining this estimate and (15) with (13) results in

$$\begin{aligned}\|\mathbf{W}\left(\mathbf{I}-\omega\mathbf{W}\right)^{\nu}\|_{0\leftarrow 0} &\le C_1\left(\tfrac{1}{\omega^3(\nu+1)}+\textstyle\sum_{j=0}^{\nu-1}\binom{\nu}{j}\left(C_{II}\omega q\right)^{\nu-j}+q\right)\\ &\le C_1\left(\tfrac{1}{\omega^3(\nu+1)}+\left(\left(1+C_{II}\omega q\right)^{\nu}-1\right)+q\right)\,,\end{aligned} \tag{17}$$

where C_1 only depends on α and C_u. Next, we have to determine the ranges of ω, ν and q such that every term in (17) is bounded by $\gamma/3$ for any $0<\gamma<1$.

Fix $0<\underline{\omega}<\overline{\omega}$ and $0<\gamma<1$. Put $\underline{\nu} := 3C_1\left(\underline{\omega}^2\gamma\right)^{-1}-1$ and fix $1\sim\overline{\nu}\ge\underline{\nu}$. Then,

$$\|\mathbf{W}\left(\mathbf{I}-\omega\mathbf{W}\right)^{\nu}\|_{0\leftarrow 0} \le \gamma \qquad \forall\, q\in\,]0,\overline{q}]\,, \quad \nu\in[\underline{\nu},\overline{\nu}]\,, \quad \omega\in[\underline{\omega},\overline{\omega}]\,,$$

with

$$\overline{q} := \min\left\{\frac{\gamma}{3C_1}, \frac{1}{C_{II}\overline{\omega}}\left(\left(1+\frac{\gamma}{3C_1}\right)^{1/\overline{\nu}}-1\right)\right\} \sim 1\,.$$

By replacing γ by γ/C_0 with C_0 as in (10) yields the proof. □

Lemma 3.3 *The approximation property holds with respect to the $\|\cdot\|_{0,1,\ell}$-norm*

$$\|\mathbf{A}_{\ell}^{-1}-\mathbf{p}_{\ell,\ell-1}\mathbf{A}_{\ell-1}^{-1}\mathbf{p}^{\mathbf{T}}{}_{\ell,\ell-1}\|_{0,1,\ell\leftarrow 0,1,\ell} \le C_A h_{\ell}\,,$$

where the constant C_A is independent of ℓ and q_{ℓ}.

Proof. We again apply the splitting

$$\mathbf{A}_\ell = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \tilde{\mathbf{A}}_\ell & \\ 0 & & \end{bmatrix} + \frac{q_\ell}{h_{\ell,1}} \left[\begin{array}{cc|c} 1 & -1 & \\ -1 & 1 & \mathbf{0} \\ \hline & \mathbf{0} & \mathbf{0} \end{array}\right].$$

Let $\mathbf{e}_{\ell,i}$ denote the ith canonical unit vector in $\mathbb{R}^{n_\ell - 1}$. It is easy to check that

$$\mathbf{A}_\ell^{-1} = \begin{bmatrix} \frac{h_{\ell,1}}{q_\ell} + \left(\tilde{\mathbf{A}}_\ell^{-1}\right)_{11} & \left(\tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_{\ell,1}\right)^\intercal \\ \tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_{\ell,1} & \tilde{\mathbf{A}}_\ell^{-1} \end{bmatrix}.$$

Next, we compute the matrix $\mathbf{V}_\ell := \mathbf{A}_\ell^{-1} - \mathbf{p}_{\ell,\ell-1}\mathbf{A}_{\ell-1}^{-1}\,\mathbf{p}^\intercal{}_{\ell,\ell-1}$. We apply

$$\mathbf{p}_{\ell,\ell-1} = \left[\begin{array}{c|ccc} \frac{1}{2} & \frac{1}{2} & 0 & \dots & 0 \\ \hline \mathbf{0} & & \tilde{\mathbf{p}}_{\ell,\ell-1} & & \end{array}\right],$$

with the (canonical) prolongation $\tilde{\mathbf{p}}_{\ell,\ell-1}$ corresponding to the reduced grids $\tilde{\mathcal{G}}_{\ell-1}$, $\tilde{\mathcal{G}}_\ell$. Explicit calculations result in

$$\mathbf{V}_\ell = \begin{bmatrix} \frac{h_{\ell,1}}{q_\ell} + \left(\tilde{\mathbf{A}}_\ell^{-1}\right)_{11} & \left(\tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_{\ell,1}\right)^\intercal \\ \tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_{\ell,1} & \tilde{\mathbf{A}}_\ell^{-1} \end{bmatrix} - \begin{bmatrix} \frac{h_{\ell-1,1}}{q_{\ell-1}} + \left(\tilde{\mathbf{A}}_{\ell-1}^{-1}\right)_{11} & \left(\tilde{\mathbf{A}}_{\ell-1}^{-1}\mathbf{e}_{\ell-1,1}\right)^\intercal \tilde{\mathbf{p}}^\intercal_{\ell,\ell-1} \\ \tilde{\mathbf{p}}_{\ell,\ell-1}\tilde{\mathbf{A}}_{\ell-1}^{-1}\mathbf{e}_{\ell-1,1} & \tilde{\mathbf{p}}_{\ell,\ell-1}\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{p}}^\intercal_{\ell,\ell-1} \end{bmatrix}.$$

Observe that

$$\frac{h_{\ell,1}}{q_\ell} = \frac{h_{\ell,1}^2}{\varepsilon} = \frac{h_{\ell-1,1}^2}{4\varepsilon} = \frac{h_{\ell-1,1}}{4q_{\ell-1}}$$

and, hence, we obtain

$$\mathbf{V}_\ell = \begin{bmatrix} \left(\tilde{\mathbf{A}}_\ell^{-1}\right)_{11} - \left(\tilde{\mathbf{A}}_{\ell-1}^{-1}\right)_{11} & \left(\tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_1\right)^\intercal - \left(\tilde{\mathbf{A}}_{\ell-1}^{-1}\mathbf{e}_{\ell-1,1}\right)^\intercal \tilde{\mathbf{p}}^\intercal_{\ell,\ell-1} \\ \tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_{\ell,1} - \tilde{\mathbf{p}}_{\ell,\ell-1}\tilde{\mathbf{A}}_{\ell-1}^{-1}\mathbf{e}_{\ell-1,1} & \tilde{\mathbf{A}}_\ell^{-1} - \tilde{\mathbf{p}}_{\ell,\ell-1}\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{p}}^\intercal_{\ell,\ell-1} \end{bmatrix}.$$

Put $\tilde{\mathbf{V}}_\ell := \tilde{\mathbf{A}}_\ell^{-1} - \tilde{\mathbf{p}}_{\ell,\ell-1}\tilde{\mathbf{A}}_{\ell-1}^{-1}\tilde{\mathbf{p}}^\intercal_{\ell,\ell-1}$ and note that $\mathbf{e}_{\ell-1,1} = \tilde{\mathbf{p}}^\intercal_{\ell,\ell-1}\mathbf{e}_{\ell,1}$. Thus,

$$\mathbf{V}_\ell = \begin{bmatrix} \left(\tilde{\mathbf{V}}_\ell\right)_{11} & \left(\tilde{\mathbf{V}}_\ell\mathbf{e}_{\ell,1}\right)^\intercal \\ \tilde{\mathbf{V}}_\ell\mathbf{e}_{\ell,1} & \tilde{\mathbf{V}}_\ell \end{bmatrix}.$$

We have to estimate $\|\mathbf{V}_\ell\|_{0,1,\ell\leftarrow 0,1,\ell}$, i.e., the Euklidean norm of

$$\begin{bmatrix} \left(\tilde{\mathbf{V}}_\ell\right)_{11} & q_\ell \left(\tilde{\mathbf{V}}_\ell \mathbf{e}_{\ell,1}\right)^\intercal \\ q_\ell^{-1}\tilde{\mathbf{V}}_\ell \mathbf{e}_{\ell,1} & \tilde{\mathbf{V}}_\ell \end{bmatrix}.$$

Since $\tilde{\mathbf{V}}_\ell$ is independent of q_ℓ we have to prove $\tilde{\mathbf{V}}_\ell \mathbf{e}_{\ell,1} = \mathbf{0}$.

Let us first compute the vector $\tilde{\mathbf{A}}_\ell^{-1}\ \mathbf{e}_{\ell,1}$. Define the 2×2-matrix

$$\mathbf{m} := \left((\varphi_{\ell,i}, \varphi_{\ell,j})_{L^2(x_{\ell,1}, x_{\ell,2})} \right)_{1\le i,j\le 2}$$

and compute explicitly $\beta := \mathbf{m}^{-1}(1,0)^\intercal = 2/h_{\ell,2}(2,-1)^\intercal$. Thus, the function $f \in L^2(\varepsilon, 1)$

$$f(x) := \begin{cases} \beta_1\varphi_{\ell,1} + \beta_2\varphi_{\ell,2} = 2(2x_{\ell,2} - 3x + \varepsilon) & \text{in } (x_{\ell,1}, x_{\ell,2})\,, \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

satisfies $\left(\int_\varepsilon^1 f\varphi_{\ell,i}\,\mathrm{dx}\right)_{i=1}^{n_\ell-1} = \mathbf{e}_{\ell,1}$. Define u as the solution of (9) with f as in (18)

$$u(x) := \int_x^1 \int_\varepsilon^t f(s)\,\mathrm{d}s\,\mathrm{d}t = \begin{cases} \frac{(x-\varepsilon)^2(x-x_{\ell,2}-h_{\ell,2})}{h_{\ell,2}^2} + 1 - \varepsilon & \varepsilon \le x \le x_{\ell,2}\,, \\ 1 - x & x_{\ell,2} \le x \le 1\,. \end{cases} \tag{19}$$

It is well known that the finite element solution for the reduced Poisson problem (9) is the nodal interpolant of the exact solution. Thus,

$$\tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_{\ell,1} = (u(x_{\ell,i}))_{i=1}^{n_\ell-1} = (1 - x_{\ell,i})_{i=1}^{n_\ell-1}\,.$$

Next, we compute $\tilde{\mathbf{p}}_{\ell,\ell-1}\tilde{\mathbf{A}}_{\ell-1}^{-1}\mathbf{e}_{\ell-1,1}$. From (19) we conclude

$$\tilde{\mathbf{A}}_{\ell-1}^{-1}\mathbf{e}_{\ell-1,1} = (1 - x_{\ell,i})_{i=1}^{n_\ell-1}$$

and by applying the definition of the grid points and the canonical finite element prolongation $\tilde{\mathbf{p}}_{\ell,\ell-1}$ we get

$$\tilde{\mathbf{p}}_{\ell,\ell-1}\tilde{\mathbf{A}}_{\ell-1}^{-1}\mathbf{e}_{\ell-1,1} = \tilde{\mathbf{A}}_\ell^{-1}\mathbf{e}_{\ell,1}\,.$$

Thus, $\tilde{\mathbf{V}}_\ell\mathbf{e}_{\ell,1} = \mathbf{0}$ and the assertion follows from the well-known approximation property for the reduced Poisson problem. □

A combination of Lemma 3.2 and 3.3 yields the convergence of the two-grid method with respect to the $\|\cdot\|_{0,1,\ell}$-norm uniform with respect to the overlap $q_\ell \in]0, \overline{q}]$ with $\overline{q} \sim 1$.

3.2.3 Multi-Grid Convergence for grids containing triangles with small overlap

From the convergence of the two-level method, we easily derive the convergence of the W-cycle multi-grid method.

Lemma 3.4 *There exist $0 < \underline{\omega} < \overline{\omega} < 1$, $1 < \underline{\nu} < \overline{\nu}$, and $\overline{q} > 0$ depending only on α and C_u such that*

$$\left\|\mathbf{K}_\ell^{MGM}\right\|_{0,1,\ell\leftarrow 0,1,\ell} \leq 1/2 \qquad \forall\, q \in]0,\overline{q}]\,, \quad \nu \in [\underline{\nu},\overline{\nu}]\,, \quad \omega \in [\underline{\omega},\overline{\omega}]\,,$$

where $\mathbf{K}_\ell^{MGM}$ denotes the W-cycle multi-grid iteration matrix with ν steps of the damped Jacobi method as the smoothing iteration.

For $\overline{q} < q < 1$, we have

$$\left\|\mathbf{K}_\ell^{MGM}\right\|_{0\leftarrow 0} \lesssim (\nu+1)^{-1}$$

for all $\nu \in \mathbb{N}$.

Proof. We adopt the theory and notations of [Hac85, Sec 7.1] and consider first the case of a small overlap $0 < q \leq \overline{q}$. We begin with the estimate of the ν-fold application of the smoothing operator, i.e., $\|\mathbf{K}_\ell^\nu\|_{0,1,\ell\leftarrow 0,1,\ell}$. Similarly as in the proof of the smoothing property in Lemma 3.2 we derive the splitting

$$\mathbf{K}_\ell^\nu = \mathbf{D}_\ell^{-1/2}\mathbf{N}_{1,\ell}^{-1/4}\left(\mathbf{I}_\ell - \omega\mathbf{W}_\ell\right)^\nu \mathbf{N}_{1,\ell}^{1/4}\mathbf{D}_\ell^{1/2}\,,$$

with $\mathbf{W}_\ell := \mathbf{N}_{1,\ell}^{1/4}\mathbf{D}_\ell^{-1/2}\mathbf{A}_\ell\mathbf{D}_\ell^{-1/2}\mathbf{N}_{1,\ell}^{-1/4}$, which leads to

$$\begin{aligned}\|\mathbf{K}_\ell^\nu\|_{0,1,\ell\leftarrow 0,1,\ell} &\leq \left\|\mathbf{N}_{1,\ell}^{1/2}\mathbf{D}_\ell^{-1/2}\mathbf{N}_{1,\ell}^{-1/4}\right\|_{0\leftarrow 0}\|\left(\mathbf{I}_\ell - \omega\mathbf{W}_\ell\right)^\nu\|_{0\leftarrow 0}\left\|\mathbf{N}_{1,\ell}^{1/4}\mathbf{D}_\ell^{1/2}\mathbf{N}_{1,\ell}^{-1/2}\right\|_{0\leftarrow 0}\\ &\lesssim \|\mathbf{I}_\ell - \omega\mathbf{W}_\ell\|_{0\leftarrow 0}^\nu\,.\end{aligned}$$

The assumption $0 < \omega \leq \overline{\omega} \leq 2/C_{\mathbf{W}}$ (cf. 16)) leads to

$$\|\mathbf{K}_\ell^\nu\|_{0,1\leftarrow 0,1} \lesssim 1$$

and the constant C_S in [Hac85, (7.1.1)] satisfies $C_S \sim 1$. In order to estimate the constants $\underline{C}_p$ and $\overline{C}_p$ in [Hac85, (7.1.2)] we have to investigate the eigenvalues of the symmetric matrix

$$\mathbf{N}_{1,\ell-1}^{-1/2}\mathbf{p}_{\ell,\ell-1}^\intercal\mathbf{N}_{1,\ell}\mathbf{p}_{\ell,\ell-1}\mathbf{N}_{1,\ell-1}^{-1/2} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4}q & 0 & \cdots & 0\\ \frac{1}{4}q & \frac{1}{4}q^2+\frac{5}{4} & \frac{1}{4} & \ddots & \vdots\\ 0 & \frac{1}{4} & \frac{3}{2} & \ddots & 0\\ \vdots & \ddots & \ddots & \ddots & \frac{1}{4}\\ 0 & \cdots & 0 & \frac{1}{4} & \frac{3}{2}\end{bmatrix}.$$

By restricting to $\overline{q} \leq 1/2$, Gerschgorin's theorem implies that the eigenvalues of this matrix are bounded from below by $1/8$ and from above by 2. Hence, $\underline{C}_p, \overline{C}_p \sim 1$ and $C^\star := 1 + \underline{C}_p \overline{C}_p \, (1 + C_S) \sim 1$ (cf. [Hac85, (7.1.6)]). Let $\mathbf{K}_\ell^{MGM}$ denote the iteration matrix of the W-cycle multi-grid method and $\mathbf{K}_\ell^{TGM}$ the iteration matrix of the two-grid method (cf. (6)). Then, according to [Hac85, (7.1.5c)] the recursion

$$\left\|\mathbf{K}_\ell^{MGM}\right\|_{0,1,\ell\leftarrow 0,1,\ell} \leq \left\|\mathbf{K}_\ell^{TGM}\right\|_{0,1,\ell\leftarrow 0,1,\ell} + C^\star \left\|\mathbf{K}_{\ell-1}^{MGM}\right\|^2_{0,1,\ell\leftarrow 0,1,\ell}$$

holds. Choose γ in Lemma 3.2 as

$$\gamma := (4C^\star C_A)^{-1}$$

and the bounds $\underline{\nu}, \overline{\nu}$, $\underline{\omega}$, $\overline{\omega}$, $\overline{q} \sim 1$ accordingly. Then, [Hac85, Lemma 7.1.6] implies that

$$\left\|\mathbf{K}_\ell^{MGM}\right\|_{0,1,\ell\leftarrow 0,1,\ell} \leq \frac{1}{2C^\star} \leq 1/2$$

yielding the multi-grid convergence with respect to the $\|\cdot\|_{0,1,\ell}$-norm.

Let us now consider the case $q \in [\overline{q}, 1]$. Since $\overline{q} \sim 1$, the standard multi-grid theory, e.g., [Hac85, Lemma 6.2.1, Prop. 6.2.14, Lemma 6.3.13,] with respect to the $\|\cdot\|_0$-norm can be applied yielding the asserted estimate. The constant in the estimate only depends on $\overline{q} \sim 1$. □

Summarising we have proved for a one-dimensional model problem that the multi-grid convergence is robust with respect to a small overlap of elements with the domain by using a weighted L^2-norm. Numerical examples show that this is not the case with respect to the usual L^2-norm. Our analysis shows that this fact is related to the convergence of unknowns corresponding to grid points lying (essentially) outside the domain. However, this is not considered as a drawback of the proposed method since these unknowns either are not interesting (since they are lying outside of the domain) or can be corrected easily in a post-processing step via extrapolation from the interior.

We emphasise that, although the analysis of the proposed multi-grid method is somewhat technical, the method itself is rather simple and its convergence is robust with respect to some characteristic parameters describing the geometry and finite element meshes. Numerical computations on realistic geometries (cf. [FHSW01]) show that the multi-grid algorithm is efficient also in two space dimensions. Further research activities are directed towards the proof of the convergence in the multi-dimensional case.

References

[BHW98] P. Bastian, W. Hackbusch, and G. Wittum. Additive and Multiplicative Multi-Grid - A Comparison. *Computing*, 60:345–364, 1998.

[Bra95] D. Braess. Towards Algebraic Multigrid for Elliptic Problems of Second Order. *Computing*, 55(4):379–393, 1995.

[BS81] R. Bank and A. Sherman. An Adaptive, Multi-Level Method for Elliptic Boundary Value Problems. *Computing*, 26:91–105, 1981.

[BX96] R.E. Bank and J. Xu. An Algorithm for Coarsening Unstructured Meshes. *Numer. Math.*, 73(1):1–36, 1996.

[Cia78] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North Holland, Amsterdam, 1978.

[FHSW01] D. Feuchter, I. Heppner, S. Sauter, and G. Wittum. Briding the Gap between Geometric and Algebraic Multi-Grid Methods. *Technical report, Universität Zürich*, in preparation, 2001.

[Hac85] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer Verlag, Berlin, 1985.

[Hac93] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations*. Springer Verlag, New York, 1993.

[HS96] W. Hackbusch and S. Sauter. Adaptive Composite Finite Elements for the Solution of PDEs Containing non-uniformly distributed Micro-Scales. *Mat. Model.*, 8(9):31–43, 1996.

[HS97a] W. Hackbusch and S. Sauter. Composite Finite Elements for Problems Containing Small Geometric Details. Part II: Implementation and Numerical Results. *Comput. Visual. Sci.*, 1(1):15–25, 1997.

[HS97b] W. Hackbusch and S. Sauter. Composite Finite Elements for problems with complicated boundary. Part III: Essential boundary conditions. *Technical report, Universität Kiel*, 1997.

[HS97c] W. Hackbusch and S. Sauter. Composite Finite Elements for the approximation of PDEs on Domains with Complicated Micro-Structures. *Numer. Math.*, 75(4):447–472, 1997.

[KY94] R. Kornhuber and H. Yserentant. Multilevel Methods for Elliptic Problems on Domains not Resolved by the Coarse Grid. *Contemporay Mathematics*, 180:49–60, 1994.

[RS87] J.W. Ruge and K. Stüben. Algebraic Multigrid. In S. McCormick, editor, *Multigrid Methods*, pages 73–130, Pennsylvania, 1987. SIAM Philadelphia.

[Xu92] J. Xu. Iterative Methods by Space Decomposition and Subspace Correction. *SIAM Reviews*, 34:581–613, 1992.

[Xu96] J. Xu. The Auxiliary Space Method and Optimal Multigrid Preconditioning Techniques for Unstructured Grids. *Computing*, 56:215–235, 1996.

Department of Mathematics, University of Zürich
http://www.math.unizh.ch/compmath/
e-mail (Frauböse): nafra@amath.unizh.ch
e-mail (Sauter): stas@amath.unizh.ch