Geometric Multigrid for High-Order Regularizations of Early Vision Problems

Stephen L. Keeling[1] and Gundolf Haase[2]

Abstract. The surface estimation problem is used as a model to demonstrate a framework for solving early vision problems by high-order regularization with natural boundary conditions. Because the application of algebraic multigrid is usually constrained by an M-matrix condition which does not hold for discretizations of high-order problems, a geometric multigrid framework is developed for the efficient solution of the associated optimality systems. It is shown that the convergence criteria of [5] are met, and in particular the general elliptic regularity required is proved. Further, the Galerkin formalism is used together with a multi-colored ordering of unknowns to permit vectorization of a symmetric Gauss-Seidel relaxation in image processing systems. The implementation is analyzed computationally and inaccuracies are corrected by lumping and by proper floating point representations. Direct one-dimensional calculations are used to estimate the effect of regularization order, regularization strength, relaxation, and data support on the multigrid reduction factor. A finite difference formulation is ruled out in favor of a finite element formulation. A representative problem from magnetic resonance coil sensitivity estimation is solved using increasingly higher orders of regularization, and the results are compared in terms of accuracy and multigrid convergence.

1 Introduction

Early vision problems are those in which dense three-dimensional information is to be inferred from sparse and generally noisy two-dimensional imaging data [16]. The model problem used here to demonstrate the framework set forth in the present work is the surface estimation problem [11] [19] [20] [22]. However, this framework has also been used for other early vision problems, including optical flow and image registration, and those results will be reported separately. The variational approach to solving such problems involves to minimize the sum of a data residual term and a solution regularization term [16]. The optimality condition determining the mimimizer is a partial differential equation whose discretization can be poorly conditioned depending upon the sparsity of data, the order of regularization, and the formulation of boundary conditions [3] [11] [19] [20] [22]. To reduce ill-conditioning, low-order regularization and low-order non-natural boundary conditions have been used [3] [11] [19] [20] [22]. Also creative conjugate gradient preconditioners [11], hierarchical [19] and wavelet bases [15] [22], and multigrid methods [3] [20] have been developed for such formulations. However, as shown below, low-order regularizations and non-natural boundary conditions have a corrupting effect on the solution to the early vision problem. For instance, in a surface estimation problem such as that reported in [10], the surface to be estimated typically has exponential growth at the boundary. Such boundary growth is naturally frustrated by low-order homogeneous boundary conditions. Thus, in the present work, high-order regularizations with natural boundary conditions are studied. Also, because the application of algebraic multigrid is usually constrained by an M-matrix condition [21] which does not hold for discretizations of high-order problems, a geometric multigrid framework is developed for the efficient solution of the associated optimality systems; see also [12] for related multigrid work.

The present paper can be summarized as follows. In Section 2 the variational formulation is given and its analysis is summarized. The general optimality system is a high-order Neumann problem whose solution regularity is needed in numerical error estimates. Since in the present

[1]Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria; email: stephen.keeling@uni-graz.at; tel: +43–316–380–5156; fax: +43–316–380–9815.

[2]Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria; email: gundolf.haase@uni-graz.at; tel: +43–316–380–5178; fax: +43–316–380–9815.

context the standard assumptions on the smoothness of coefficients and of the domain boundary do not apply, a proof of the required elliptic regularity is given in Section 2. In Section 3 a finite element discretization of the general optimality system of Section 2 is given and its analysis is summarized. In Section 4 a geometric multigrid formulation is presented along with the demonstration that the convergence criteria of [5] are met. Specifically, the approximation property is fulfilled through the Galerkin formalism, and the smoothing property is fulfilled by a symmetric Gauss-Seidel relaxation. Section 5 is concerned with details of computational implementation and with rejected alternative formulations. For instance, lumping of zero-order terms is important in the limit of high data fidelity [10]. Also, a multi-colored ordering of unknowns is used to permit that the relaxation scheme is vectorizable in systems such as MATLAB [14] or particularly in the image processing system IDL [6]. Floating point aspects are analyzed to recover the effective loss of positive definiteness of the discretization of elliptic operators. Then direct one-dimensional calculations are used to estimate the effect of regularization order, regularization strength, relaxation, and data support on the multigrid reduction factor. For a model problem, a finite difference formulation is shown to be inferior in relation to the finite element formulation. Finally, in Section 6 a representative problem from magnetic resonance coil sensitivity estimation [10] is solved using increasingly higher orders of regularization, and the results are compared in terms of accuracy and multigrid convergence.

2 Variational Formulation

The model problem considered in the present work is to estimate a smooth function u satisfying $mu \approx r$ for compactly supported and noisy data m and r. The solution to this problem is defined as the minimizer of the following functional:

$$J(w) = \int_\Omega |mw - r|^2 d\boldsymbol{x} + \mu \sum_{|\alpha|=\nu} \frac{\nu!}{\alpha!} \int_\Omega |D^\alpha w|^2 d\boldsymbol{x} \tag{2.1}$$

where $\Omega = (0,1)^N$. In the imaging examples, $N = 2$ holds, although the spatial dimension is not theoretically limited. Also μ is a regularization parameter controlling the trade-off between the smoothness of the solution and the extent to which the solution matches the given data. The minimizer for (2.1) is well defined according to Theorem 2.1 below; see [7] and [10]. Here and throughout the paper, $L^2(\Omega)$ denotes the Hilbert space of Lebesgue measurable functions which are square integrable in Ω, and $H^\nu(\Omega)$ denotes the Sobolev space of functions with derivatives up to order ν in $L^2(\Omega)$. Also $L^\infty(\Omega)$ denotes the Banach space of essentially bounded Lebesgue measurable functions. The usual inner products, norms and semi-norms on these spaces are denoted by $(\cdot,\cdot)_{H^\nu(\Omega)} = (\cdot,\cdot)_\nu$, $\langle\cdot,\cdot\rangle_{H^\nu(\Omega)} = \langle\cdot,\cdot\rangle_\nu$, $\|\cdot\|_\nu^2 = (\cdot,\cdot)_\nu$, $|\cdot|_\nu^2 = \langle\cdot,\cdot\rangle_\nu$, $\nu \geq 0$ and $\|\cdot\|_{L^\infty(\Omega)} = \|\cdot\|_\infty$. See [1] for further details of these function spaces.

Theorem 2.1 *Let $r \in L^2(\Omega)$ and $m \in L^\infty(\Omega)$ with $m \geq \bar{m} > 0$ on some regular domain $S \subset \Omega$. Then J in (2.1) has a unique minimizer $u \in H^\nu(\Omega)$.*

For a fixed regularization parameter, the continuous dependence of the minimizer on the data is given as follows; see [7] and [10].

Theorem 2.2 *Assume that $\{m_n\} \subset L^\infty(\Omega)$ and $\{r_n\} \subset L^2(\Omega)$ are data for which:*

$$\lim_{n\to\infty} \|m_n - m_0\|_\infty = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \|r_n - r_0\|_0 = 0 \tag{2.2}$$

and $m_n \geq \bar{m} > 0$ for all n on some regular domain $S \subset \Omega$. Also, for each n let $u_n \in H^\nu(\Omega)$ denote the minimizer for J in (2.1) corresponding to data (m_n, r_n). Then:

$$\lim_{n\to\infty} \|u_n - u_0\|_0 = 0. \tag{2.3}$$

Now convergence of minimizers in the limit of increasing data fidelity and vanishing regularization is given as follows; see [7] and [10].

Theorem 2.3 *For data $r^* \in L^2(\Omega)$ and $m^* \in L^\infty(\Omega)$ with $m^* \geq \bar{m} > 0$ on Ω, let $u^* = r^*/m^* \in H^\nu(\Omega)$. Assume that $\{m_n\} \subset L^\infty(\Omega)$ and $\{r_n\} \subset L^2(\Omega)$ are data for which:*

$$\lim_{n\to\infty} \|m_n - m^*\|_\infty = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \|r_n - r^*\|_0 = 0 \tag{2.4}$$

and $m_n \geq \bar{m}$ for all n on some regular domain $S \subset \Omega$. Also, for each n let $u_n \in H^\nu(\Omega)$ denote the minimizer of J in (2.1) corresponding to data (m_n, r_n) and to regularization parameter μ_n. If $\mu_n \to 0$ at a rate for which $\|m_n u^* - r_n\|_0^2/\mu_n$ remains bounded, then:

$$\lim_{n\to\infty} \|u_n - u^*\|_0 = 0. \tag{2.5}$$

Now consider the characterization of a minimizer for J through its optimality system. This system will subsequently be discretized with the finite element method in the next section. For $w \in H^\nu(\Omega)$, a straightforward calculation shows that the variational derivative of J is given by [17]:

$$\frac{1}{2}\frac{\delta J}{\delta w}(w; v) = B(w, v) - F(v) \tag{2.6}$$

where:

$$B(w, v) = (m^2 w, v)_0 + \mu \langle w, v \rangle_\nu \tag{2.7}$$

and:

$$F(v) = (rm, v)_0. \tag{2.8}$$

Let L denote the differential operator associated with B in the sense that $(Lw, v)_0 = B(w, v)$ for all sufficiently smooth w and v, and note that since B is symmetric,

$$B(w, v) = B(v, w), \quad \forall w, v \in H^\nu(\Omega) \tag{2.9}$$

L is self-adjoint:

$$L = \mu(-\Delta)^\nu + m^2 I = L^*, \quad \text{Dom}(L) = H^{2\nu}(\Omega). \tag{2.10}$$

The weak formulation of the necessary optimality condition on the minimizer $u \in H^\nu(\Omega)$ that (2.6) vanish for all $v \in H^\nu(\Omega)$ is:

$$B(u, v) = F(v), \quad \forall v \in H^\nu(\Omega). \tag{2.11}$$

These forms are clearly bounded in the following sense:

$$|B(w, v)| \leq c_1 \|w\|_\nu \|v\|_\nu, \quad c_1 = \|m\|_\infty^2 + \mu \tag{2.12}$$

$$|F(v)| \leq c_2 \|v\|_\nu, \quad c_2 = \|m\|_\infty \|r\|_0. \tag{2.13}$$

Since the kernel of the semi-norm $|\cdot|_\nu$ is a space of polynomials and the support of m has positive measure, it is straightforward to show the following coercivity:

$$|B(w, w)| \geq c_3 \|w\|_\nu^2 \tag{2.14}$$

where c_3 depends upon m and μ. Thus, the Lax-Milgram Lemma gives the following theorem [4].

Theorem 2.4 *Let $r \in L^2(\Omega)$ and $m \in L^\infty(\Omega)$ with $m \geq \bar{m} > 0$ on some regular domain $S \subset \Omega$. Then there exists a unique $u \in H^\nu(\Omega)$ satisfying (2.11) and*

$$\|u\|_\nu \leq c \|r\|_0 \tag{2.15}$$

for a constant c depending upon m and μ.

Under standard assumptions on the regularity of $\partial\Omega$ and of coefficients in B, it is a usual property of a weak solution satisfying (2.11) that it possesses higher regularity than guaranteed by Theorem 2.4. However, the standard regularity assumptions are lacking in the present context, and therefore the following theorem is proved for the present purposes. The techniques used are adapted from [8].

Theorem 2.5 *The solution u to (2.11) satisfies $u \in H^{2\nu}(\Omega)$ and*

$$\|u\|_{2\nu} \leq c\|r\|_0 \tag{2.16}$$

for a constant c depending upon m and μ.

Proof. To make use of tools in [8], the following perturbed bilinear form is defined which has smooth coefficients of all orders:

$$\tilde{B}(w,v) = \mu(w,v)_0 + \mu\langle w,v\rangle_\nu. \tag{2.17}$$

Then a corresponding perturbation of (2.11) is defined:

$$\tilde{F}(v) = (rm + (\mu - m^2)u, v)_0 \tag{2.18}$$

in such a way that u is also the unique solution to:

$$\tilde{B}(u,v) = \tilde{F}(v), \quad \forall v \in H^\nu(\Omega). \tag{2.19}$$

With respect to (2.19), the Corollary to Lemma 15.5 in [8] implies that for any smooth domain $\Omega' \subset\subset \Omega$,

$$\|u\|_{H^{2\nu}(\Omega')} \leq c_1\|rm + (\mu - m^2)u\|_0 \tag{2.20}$$

where c_1 depends upon Ω'. Let Ω' be chosen so that Ω is covered by the union of Ω' and the following cover of $\partial\Omega$. Specifically, let $\partial\Omega$ be covered with a finite set of balls $\{\beta\}$ with the following properties. First there exists a finite set of balls $\{\beta'\}$ which also cover $\partial\Omega$ such that each β' shares its center with some β while satisfying $\beta' \subset\subset \beta$. Additionally, all balls have their centers on $\partial\Omega$ and no ball in the set $\{\beta\}$ contains the center of another ball in that set. Finally, all faces of $\partial\Omega$ of dimension $N-1$, $N-2$, and so on, are similarly covered by the original covering $\{\beta\}$.

Next, let the constants $\{\lambda_k\}_{k=1}^{2\nu}$ be defined as the solution to the system:

$$\sum_{k=1}^{2\nu} \lambda_k(-k)^l = 1, \quad l = k-\nu, \quad k = 1,\ldots,\nu \quad \text{and} \quad l = k, \quad k = (\nu+1),\ldots,2\nu. \tag{2.21}$$

Then on those balls β overlapping $\partial\Omega \cap \{\boldsymbol{x} : x_i = 0\}$ define the reflective extension operators E_i by:

$$E_i u(x_1,\ldots,x_N) = \sum_{k=1}^{2\nu} \lambda_k u(x_1,\ldots,-x_i/k,\ldots,x_N), \quad x_i \leq 0. \tag{2.22}$$

Also for those balls β overlapping $\partial\Omega \cap \{\boldsymbol{x} : x_i = 0\}$ and for a given function $\phi \in C_0^\infty(\beta)$, define the folding operators F_i by:

$$F_i\phi(x_1,\ldots,x_N) = \sum_{k=1}^{2\nu} k\lambda_k \phi(x_1,\ldots,-kx_i,\ldots,x_N). \tag{2.23}$$

Then collect these operators to define the following in the neighborhood of the corner of Ω situated at the origin:

$$Eu = \begin{cases} u, & \boldsymbol{x} > 0 \\ E_i u, & x_i \leq 0, \\ E_iE_j u, & x_i, x_j \leq 0, \\ \vdots & \vdots \\ [\prod_{i=1}^N E_i]u, & \boldsymbol{x} < 0 \end{cases} \qquad F\phi = \begin{cases} (I+F_i)\phi, & \text{in } \beta_i \\ (I+F_i)(I+F_j)\phi, & \text{in } \beta_{ij} \\ \vdots & \vdots \\ [\prod_{i=1}^N (I+F_i)]\phi, & \text{in } \beta_0 \end{cases} \tag{2.24}$$

where for a multi-index $\kappa = (i, j, k, \ldots)$, β_κ denotes a ball overlapping $\partial\Omega \cap \{\boldsymbol{x} : x_i = x_j = x_k = \cdots = 0\}$ and β_0 is centered at the origin. These operators E and F are defined similarly along the rest of $\partial\Omega$. Note that the choice of the constants $\{\lambda_k\}$ implies:

$$\partial_n^l F\phi|_{\partial\Omega} = 0, \quad l = \nu, \ldots, (2\nu - 1) \tag{2.25}$$

and

$$\partial_n^l Eu|_{\partial\Omega^-} = \partial_n^l u|_{\partial\Omega^+}, \quad l = 0, \ldots, (\nu - 1) \tag{2.26}$$

where ∂_n denotes the normal derivative in the outward direction. Further, partial integration of u in (2.19) with v strategically chosen concentrated on the boundary leads to:

$$\partial_n^l Eu|_{\partial\Omega^-} = 0 = \partial_n^l u|_{\partial\Omega^+}, \quad l = \nu, \ldots, (2\nu - 1). \tag{2.27}$$

For a given ball β with $\sigma = \beta \cap \Omega$ and $\phi \in C_0^\infty(\beta)$, a detailed calculation (cf. (17.7) in [8]) shows that with $\tilde{L} = \mu(-\Delta)^\nu + \mu I$,

$$\begin{aligned} (\tilde{L}\phi, Eu)_{L^2(\beta)} = (\tilde{L}F\phi, u)_{L^2(\sigma)} &= \textstyle\sum_{l=0}^{\nu-1} (-1)^{\nu-l} (\partial_n^{2\nu-1-l} F\phi, \partial_n^l u)_{L^2(\partial\sigma)} + \tilde{B}(F\phi, u) \\ &= \tilde{B}(F\phi, u) = (rm + (\mu - m^2)u, \phi). \end{aligned} \tag{2.28}$$

So with the estimate,

$$\begin{aligned} |(\tilde{L}\phi, Eu)_{L^2(\beta)}| &= \tilde{B}(F\phi, u) = (rm + (\mu - m^2)u, \phi) \\ &\leq c_2 \|rm + (\mu - m^2)u\|_0 \|\phi\|_{L^2(\beta)} \end{aligned} \tag{2.29}$$

it follows with Theorem 16.1 of [8] that $Eu \in H^{2\nu}(\beta')$ and with $\sigma' = \beta' \cap \Omega$,

$$\|u\|_{H^{2\nu}(\sigma')} = \|Eu\|_{H^{2\nu}(\sigma')} \leq \|Eu\|_{H^{2\nu}(\beta')} \leq c_3[c_2\|rm + (\mu - m^2)u\|_0 + \|u\|_0]. \tag{2.30}$$

The combination of (2.20) and (2.30) gives:

$$\|u\|_{2\nu} \leq \|u\|_{H^{2\nu}(\Omega')} + \sum_{\sigma' = \beta' \cap \Omega} \|u\|_{H^{2\nu}(\sigma')} \leq (c_1 + c_2 c_3 n)[\|rm + (\mu - m^2)u\|_0 + \|u\|_\nu] \tag{2.31}$$

where n is the number of balls in the set $\{\beta'\}$. Since $r \in L^2(\Omega)$ and $m \in L^\infty(\Omega)$, it follows from (2.15) and (2.31) that $u \in H^{2\nu}(\Omega)$ and that the estimate (2.16) holds. ∎

3 Finite Element Discretization

The boundary value problem of (2.11) is consistently discretized through finite elements as follows. For the following let $C^\nu(\Omega)$, $\Omega \subset \boldsymbol{R}^N$, denote the Banach space of functions whose derivatives up to order ν are continuous in Ω, and recall the definition of $H^\nu(\Omega)$ from Section 2. Let Ω be divided into a grid of 2^{pN} cells, where each cell has unit aspect ratio and width $h = 2^{-p}$. Specifically, with the integer component N-dimensional multi-indices $\kappa = (i, j, k, \ldots)$, $\mathbf{0} = (0, 0, \ldots, 0)$, and $\mathbf{1} = (1, 1, \ldots, 1)$, define the cell corners by $\boldsymbol{x}_{\kappa+\frac{1}{2}} = \kappa h$, $\mathbf{0} \leq \kappa \leq 2^p \cdot \mathbf{1}$, and the cell centroids by $\boldsymbol{x}_\kappa = (\kappa - \frac{1}{2})h$, $\mathbf{1} \leq \kappa \leq 2^p \cdot \mathbf{1}$. Then let $S_h^\nu(\Omega) \subset C^{\nu-1}(\Omega) \cap H^\nu(\Omega)$ denote the N-fold tensor products of B-splines of degree ν defined on the established grid [9]. According to the approximation properties given in Theorems 5.1-2 and 5.2-1 of [2], there is a constant c_1 independent of h such that:

$$\inf_{\chi \in S_h^\nu(\Omega)} \|u - \chi\|_\nu \leq c_1 h^\nu \|u\|_{2\nu}, \quad \forall u \in H^{2\nu}(\Omega). \tag{3.1}$$

According to Theorems 5.1-2 and 5.2-3 of [2], the approximation space also possesses the inverse property, required in Section 4, that there is a c_2 independent of h such that:

$$\|\chi\|_\nu \le c_2 h^{-\nu} \|\chi\|_0, \quad \forall \chi \in S_h^\nu(\Omega). \tag{3.2}$$

Then the finite element approximation to the solution of (2.11) is $u_h \in S_h^\nu(\Omega)$ defined by:

$$B(u_h, \chi) = F(\chi), \quad \forall \chi \in S_h^\nu(\Omega). \tag{3.3}$$

Since $S_h^\nu(\Omega)$ is a closed subspace of $H^\nu(\Omega)$, (2.12), (2.13) and (2.14) hold on $S_h^\nu(\Omega)$ as well, and thus u_h is well defined by (3.3). By Céa's Lemma [4] there is a constant c_3 independent of h such that the solutions to (2.11) and (3.3) satisfy:

$$\|u - u_h\|_\nu \le c_3 \inf_{\chi \in S_h^\nu(\Omega)} \|u - \chi\|_\nu. \tag{3.4}$$

Combining the above inequalities with (2.16) gives the following convergence.

Theorem 3.1 *The respective solutions u and u_h to (2.11) and (3.3) satisfy*

$$\|u - u_h\|_\nu \le c h^\nu \|r\|_0 \tag{3.5}$$

for a constant c independent of h.

4 Geometric Multigrid Formulation

The geometric multigrid formulation follows [5]. The usual strategy is generally to enhance a convergent relaxation scheme by using its initial and rapid smoothing of small scales on finer grids and then to transfer the problem progressively to coarser grids before relaxation is decelerated. The principal ingredients of the strategy include the definition of a smoothing relaxation scheme and the definition of a coarse grid representation of the problem which can be used to provide an improvement or correction on a finer grid. For this, let Ω be divided into a nested sequence of grids ranging from the coarsest at level $l = 0$ to the finest at $l = l_{\max}$. The grid at level l consists of $2^{p_l N}$ cells having unit aspect ratio and width $h_l = 2^{-p_l}$ where $p_{l+1} = p_l + 1$ and p_0 is small enough that a linear system of dimension $2^{p_0 N}$ can easily be solved directly. Thus, the cell widths on adjacent levels satisfy:

$$h_l \le 2h_{l+1}. \tag{4.1}$$

As in Section 3, the cell corners are defined by $\boldsymbol{x}_{\kappa+\frac{1}{2}} = \kappa h_l$, $\mathbf{0} \le \kappa \le 2^{p_l} \cdot \mathbf{1}$, and the cell centroids by $\boldsymbol{x}_\kappa = (\kappa - \frac{1}{2})h_l$, $\mathbf{1} \le \kappa \le 2^{p-l} \cdot \mathbf{1}$. The spline basis for

$$S_l^\nu(\Omega) = S_{h_l}^\nu(\Omega) \tag{4.2}$$

is $(2^{p_l} + \nu)^N$-dimensional as illustrated in Fig. 1 for $N = 1$, $\nu = 1, 2$, $p_l = 1$ and $p_{l+1} = 2$. Note that ghost cells outside Ω are shown in Fig. 1 to emphasize the dimension of $S_l^\nu(\Omega)$; see also the discussion of grids in relation to lumping in Section 5.

In view of (2.12), the Lax-Milgram Lemma [4] guarantees the existence of an operator L_l such that:

$$(L_l \chi, \psi)_0 = B(\chi, \psi), \quad \forall \chi, \psi \in S_l^\nu(\Omega). \tag{4.3}$$

Defining the projection operator $P_l : L^2(\Omega) \to S_l^\nu(\Omega)$ gives:

$$(P_l r m, \chi)_0 = F(\chi), \quad \forall \chi \in S_l^\nu(\Omega). \tag{4.4}$$

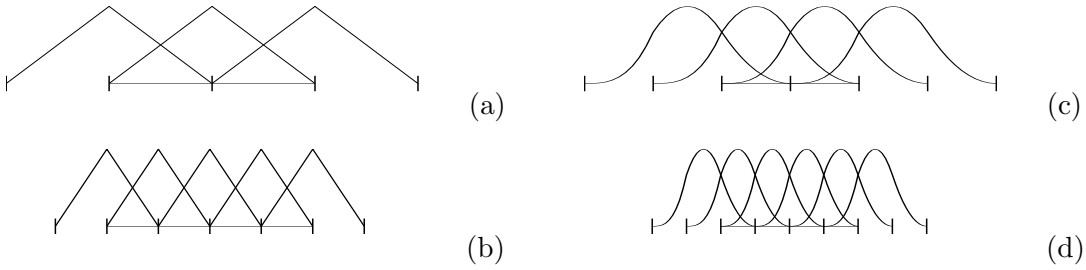

Figure 1: Examples of ($N = 1$) one-dimensional nested grids. (a) $\nu = 1$, $p = 1$, (b) $\nu = 1$, $p = 2$, (c) $\nu = 2$, $p = 1$, and (d) $\nu = 2$, $p = 2$.

Thus, the problem (3.3) is naturally formulated on the subspaces $S_l^\nu(\Omega)$ as follows. For $f_l = P_l rm$, find $u_l \in S_l^\nu(\Omega)$ such that:

$$L_l u_l = f_l, \quad \text{where} \quad L_h u_h = f_h, \quad l = l_{\max}. \tag{4.5}$$

If I_{l-1} denotes the injection operator from $(S_{l-1}^\nu(\Omega), \|\cdot\|_0)$ into $(S_l^\nu(\Omega), \|\cdot\|_0)$, then the variational problems on adjacent levels can be related according to:

$$\begin{aligned} (L_{l-1}\chi, \psi)_0 = B(\chi, \psi) &= B(I_{l-1}\chi, I_{l-1}\psi) = (L_l I_{l-1}\chi, I_{l-1}\psi)_0 \\ &= (I_{l-1}^* L_l I_{l-1}\chi, \psi)_0, \quad \forall \chi, \psi \in S_{l-1}^\nu(\Omega) \end{aligned} \tag{4.6}$$

or:

$$L_{l-1} = I_{l-1}^* L_l I_{l-1}. \tag{4.7}$$

The above relations are given on the operator level but must be implemented in terms of basis coefficients. For this, let $\boldsymbol{R}_l$ denote the $(2^{p_l} + \nu)^N$-dimensional Euclidean space of basis function coefficients U_l for functions $u_l \in S_l^\nu(\Omega)$. Also, for N-dimensional multi-indices $\imath, \jmath$ define the inner product and norm,

$$[X_l, Y_l] = h_l^N \sum_{\imath,\jmath} X_{l,\imath} Y_{l,\jmath}, \quad [X_l] = [X_l, X_l]^{\frac{1}{2}} \tag{4.8}$$

on $\boldsymbol{R}_l$ and let K_l denote the bijective mapping form $(\boldsymbol{R}_l, [\cdot])$ to $(S_l^\nu(\Omega), \|\cdot\|_0)$ so that:

$$u_l = K_l U_l. \tag{4.9}$$

According to Theorems 5.1-2 and 5.2-3 in [2], the norms of coefficients in $\boldsymbol{R}_l$ and of the corresponding functions in $S_l^\nu(\Omega)$ are equivalent:

$$c_1[U] \le \|K_l U\|_0 \le c_2[U], \quad \forall U \in \boldsymbol{R}_l. \tag{4.10}$$

Thus, for a given basis $\{\chi_\imath\}$ of $S_l^\nu(\Omega)$, which can be expressed as $\{K_l \boldsymbol{e}_\imath\}$ for unit vectors $\{\boldsymbol{e}_\imath\} \subset \boldsymbol{R}_l$, the coefficient matrix representations of the operators L_l are given by:

$$A_{l,\imath\jmath} = (L_l \chi_\imath, \chi_\jmath)_0 = (L_l K_l \boldsymbol{e}_\imath, K_l \boldsymbol{e}_j)_0 = [K_l^* L_l K_l \boldsymbol{e}_\imath, \boldsymbol{e}_\jmath] \tag{4.11}$$

or:

$$A_l = K_l^* L_l K_l. \tag{4.12}$$

Also the right side of (4.5) has the coefficient representation:

$$f_l = K_l F_l \tag{4.13}$$

and the problem (4.5) takes the form:

$$A_l U_l = F_l, \quad \text{where} \quad A_h U_h = F_h, \quad l = l_{\max}. \tag{4.14}$$

Now the variational problems on adjacent levels are related on the coefficient space $\boldsymbol{R}_l$ according to:

$$A_{l-1} = K_{l-1}^* L_{l-1} K_{l-1} = K_{l-1}^* I_{l-1}^* L_l I_{l-1} K_{l-1} = R_l^{l-1} K_l^* L_l K_l E_{l-1}^l = R_l^{l-1} A_l E_{l-1}^l \tag{4.15}$$

where $E_{l-1}^l : \boldsymbol{R}_{l-1} \to \boldsymbol{R}_l$ and $R_l^{l-1} : \boldsymbol{R}_l \to \boldsymbol{R}_{l-1}$ are the canonical expansion and restriction operators satisfying:

$$I_{l-1} K_{l-1} = K_l E_{l-1}^l, \quad R_l^{l-1} = (E_{l-1}^l)^* \tag{4.16}$$

and giving:

$$A_{l-1} = R_l^{l-1} A_l E_{l-1}^l. \tag{4.17}$$

In words, the relation $I_{l-1} K_{l-1} = K_l E_{l-1}^l$ means that $E_{l-1}^l U_{l-1}$ produces coefficients $U_l = E_{l-1}^l U_{l-1} \in \boldsymbol{R}_l$ from those $U_{l-1} \in \boldsymbol{R}_{l-1}$ so that the *functions* $K_l U_l \in S_l^\nu(\Omega)$ and $K_{l-1} U_{l-1} \in S_{l-1}^\nu(\Omega)$ are identical.

From (2.14) it follows that the the symmetric matrices A_l of (4.11) are positive definite [4]. Thus, with (4.16) and (4.17) the coarse grid problem and the intergrid transfer operators have been defined. It remains to identify a suitable relaxation scheme and to define the multigrid iteration.

Since the matrices A_l are symmetric, it is natural to use a symmetric relaxation scheme. For this, the symmetric Gauss-Seidel relaxation is used:

$$\begin{aligned} U_l^{k+1} &= \mathcal{S}_l(U_l^k, F_l) &= U_l^k - \omega W_l^{-1}(A_l U_l^k - F_l) \\ &= S_l U_l^k + \omega W_l^{-1} F_l, \quad S_l = I - \omega W_l^{-1} A_l \end{aligned} \tag{4.18}$$

where in practice $\omega = 1$ is used in the present work. Also, W_l is given by:

$$W_l = (\mathcal{D}_l + \mathcal{L}_l)\mathcal{D}_l^{-1}(\mathcal{D}_l + \mathcal{L}_l^{\mathrm{T}}) \tag{4.19}$$

where the diagonal, strictly lower triangular and strictly upper triangular parts of A_l are the respective terms in the sum:

$$A_l = \mathcal{D}_l + \mathcal{L}_l + \mathcal{L}_l^{\mathrm{T}}. \tag{4.20}$$

The following properties of the relaxation scheme are needed below. First, note that:

$$W_l = A_l + \mathcal{L}_l \mathcal{D}_l^{-1} \mathcal{L}_l^{\mathrm{T}}, \quad W_l = W_l^{\mathrm{T}} \tag{4.21}$$

and that the induced matrix norms $\|\cdot\|_{\ell_2}$ and $[\cdot]$ are identical. From the positive definiteness of A_l it follows that:

$$0 < [A_l U, U] \leq [W_l U, U], \quad \forall U \in \boldsymbol{R}_l. \tag{4.22}$$

Also the norm $[W_l]$ can be estimated in terms of the norm $[A_l]$ as follows:

$$[W_l] \leq [A_l] + [\mathcal{L}_l \mathcal{D}_l^{-1} \mathcal{D}_l \mathcal{D}_l^{-1} \mathcal{L}_l^{\mathrm{T}}] \tag{4.23}$$

Since $[\mathcal{L}_l \mathcal{D}_l^{-1}]^2 = [(\mathcal{L}_l \mathcal{D}_l^{-1})^{\mathrm{T}} (\mathcal{L}_l \mathcal{D}_l^{-1})] = [\mathcal{D}_l^{-1} \mathcal{L}_l^{\mathrm{T}}]^2$,

$$[W_l] \leq [A_l] + [\mathcal{D}_l] \|\mathcal{L}_l \mathcal{D}_l^{-1}\|_{\ell_1} \|\mathcal{D}_l^{-1} \mathcal{L}_l^{\mathrm{T}}\|_{\ell_\infty}. \tag{4.24}$$

Since $[\mathcal{D}_l] \leq \max_{ij} |A_{l,ij}| \leq [A_l]$ holds [13] together with $\|\mathcal{L}_l \mathcal{D}_l^{-1}\|_{\ell_1} = \|\mathcal{D}_l^{-1} \mathcal{L}_l^{\mathrm{T}}\|_{\ell_\infty} \leq \|\mathcal{D}_l^{-1} A_l\|_{\ell_\infty}$ it follows that:

$$[W_l] \leq (1 + \|\mathcal{D}_l^{-1} A_l\|_{\ell_\infty})[A_l]. \tag{4.25}$$

Note that A_l may depart from diagonal dominance because of cells near the boundary; nevertheless, $\|\mathcal{D}_l^{-1}A_l\|_{\ell_\infty}$ can be estimated as follows:

$$\begin{aligned} |D_{l,\imath\imath}^{-1}A_{l,\imath\jmath}| &= \left|\frac{\mu\langle\chi_\imath,\chi_\jmath\rangle_\nu + (m^2\chi_\imath,\chi_\jmath)_0}{\mu\langle\chi_\imath,\chi_\imath\rangle_\nu + (m^2\chi_\imath,\chi_\imath)_0}\right| \leq \frac{|\chi_\imath|_\nu|\chi_\jmath|_\nu + \mu^{-1}\|m\|_\infty^2\|\chi_\imath\|_0\|\chi_\jmath\|_0}{|\chi_\imath|_\nu^2} \\ &\leq (1+\mu^{-1}\|m\|_\infty^2)\frac{\|\chi_\imath\|_\nu^2 + \|\chi_\jmath\|_\nu^2}{2|\chi_\imath|_\nu^2} \leq (1+\mu^{-1}\|m\|_\infty^2)\frac{\max_\imath\|\chi_\imath\|_\nu^2}{\min_\imath|\chi_\imath|_\nu^2}. \end{aligned} \tag{4.26}$$

Note that the functions $\{\chi_\imath\}$ are translation invariant and supported in Ω in such a way that the max and min above are computed from a number of cases near $\partial\Omega$ which depend only upon N and ν but not h or l. Thus, since the number of nontrivial elements in any row of A_l is $(2\nu+1)^N$, $\|D_{l,\imath\imath}^{-1}A_{l,\imath\jmath}\|_\infty$ is bounded independently of h and l:

$$[W_l] \leq \theta[A_l]. \tag{4.27}$$

Now that the relaxation scheme has been defined, a symmetric multigrid cycle can be defined as follows:

MGC$(l, U, F; \sigma, \tau)$
  if $l = 0$ then
    $U = A_0^{-1}F$
  else
    for $k = 1, \ldots, \sigma$ set $U = \mathcal{S}_l(U, F)$
    $D = R_l^{l-1}(F - A_lU)$
    $V = 0$
    for $k = 1, \ldots, \tau$ do MGC$(l-1, V, D; \sigma, \tau)$
    $U = U + E_{l-1}^l V$
    for $k = 1, \ldots, \sigma$ set $U = \mathcal{S}_l(U, F)$

where for σ_1 pre-smoothing steps $\hat{\mathcal{S}}_l$ and σ_2 post-smoothing $\mathcal{S}_l$ steps with:

$$\sigma = \sigma_1 = \sigma_2 \tag{4.28}$$

and:

$$\hat{S}_l = S_l \tag{4.29}$$

the multi-grid cycle is a $V(\sigma)$ cycle if $\tau = 1$ and a $W(\sigma)$ cycle if $\tau = 2$. The nested or full multigrid iteration is given by:

FMG$(l_{\max}, U, F; \sigma, \tau)$
  set $F_{l_{\max}} = F$
  for $l = l_{\max}, \ldots, 1$ set $F_{l-1} = R_l^{l-1}F_l$
  set $U = A_0^{-1}F_0$
  for $l = 1, \ldots, l_{\max}$
    $U = E_{l-1}^l U$
    for $k = 1, \ldots, \tau$ do MGC$(l, U, F_l; \sigma, \tau)$

While the FMG iteration is used in practice and followed by further MGC cycles as necessary, the convergence of MGC is considered below. For this, define first the two-grid cycle,

TGC$(l, U, F; \sigma)$
  for $k = 1, \ldots, \sigma$ set $U = S_lU + \omega W_l^{-1}F$
  $D = R_l^{l-1}(F - A_lU)$
  $V = A_{l-1}^{-1}D$
  $U = U + E_{l-1}^l V$
  for $k = 1, \ldots, \sigma$ set $U = S_lU + \omega W_l^{-1}F$

whose iteration matrix is given by:

$$M_l^{\mathrm{TGC}}(\sigma) = S_l^\sigma [I - E_{l-1}^l A_{l-1}^{-1} R_l^{l-1} A_l] S_l^\sigma. \tag{4.30}$$

Similarly, the MGC iteration matrix is given by [5]:

$$M_l^{\mathrm{MGC}}(\sigma,\tau) = M_l^{\mathrm{TGC}}(\sigma) + S_l^\sigma E_{l-1}^l [M_{l-1}^{\mathrm{MGC}}(\sigma,\tau)]^\tau A_{l-1}^{-1} R_l^{l-1} A_l S_l^\sigma. \tag{4.31}$$

The convergence framework in [5] requires to establish an approximation property and a smoothing property. The approximation property is given as follows (cf. Theorem 10.6.17, Remark 10.7.13 and Lemma 10.7.8 in [5]):

Theorem 4.1 *Under the conditions (2.9), (2.10), (2.12), (3.2), (3.5), (4.1), (4.10), (4.11), (4.16), (4.17), (4.21), (4.22), (4.27), (4.28), (4.29), it follows that:*

$$0 \le A_l^{-1} - E_{l-1}^l A_{l-1}^{-1} R_l^{l-1} \le \theta W_l^{-1}. \tag{4.32}$$

The smoothing property is given as follows (cf. Theorem 10.7.11 in [5]):

Theorem 4.2 *Under the assumptions of Theorem 4.1, the smoother S_l, written in the form $X_l = I - A_l^{1/2} S_l A_l^{-1/2} = A_l^{1/2} W_l^{-1} A_l^{1/2}$, satisfies:*

$$0 \le (I - X_l)^\sigma X_l (I - X_l)^\sigma \le \eta(\sigma) I \tag{4.33}$$

where $\eta(\sigma) = \sigma^\sigma/(1+\sigma)^{(1+\sigma)}$ satisfies $\lim_{\sigma\to\infty} \eta(\sigma) = 0$.

Now (4.32) implies

$$\begin{aligned} 0 &\le A_l^{1/2} M_l^{\mathrm{TGC}}(\sigma) A_l^{-1/2} \\ &= [A_l^{1/2} S_l A_l^{-1/2}]^\sigma \{A_l^{1/2} [A_l^{-1} - E_{l-1}^l A_{l-1}^{-1} R_l^{l-1}] A_l^{1/2}\} [A_l^{1/2} S_l A_l^{-1/2}]^\sigma \\ &\le \theta [A_l^{1/2} S_l A_l^{-1/2}]^\sigma [A_l^{1/2} W_l^{-1} A_l^{1/2}] [A_l^{1/2} S_l A_l^{-1/2}]^\sigma \\ &= \theta (I - X_l)^\sigma X_l (I - X_l)^\sigma \quad \le \quad \theta \eta(\sigma) I \end{aligned} \tag{4.34}$$

and thus with (4.33) the two-grid convergence is obtained.

Theorem 4.3 *Under the assumptions of Theorem 4.1, the two grid cycle satisfies:*

$$[A_l^{1/2} M_l^{TGC}(\sigma) A_l^{-1/2}] \le \theta \eta(\sigma). \tag{4.35}$$

The $V(\sigma)$ and $W(\sigma)$ cycles converge according to the following (cf. Theorem 10.7.15 and (10.7.18) in [5]):

Theorem 4.4 *Under the assumptions of Theorem 4.1, the iteration matrices $M_l^V(\sigma) = M_l^{\mathrm{MGC}}(\sigma,1)$ of the $V(\sigma)$ cycle and $M_l^W(\sigma) = M_l^{\mathrm{MGC}}(\sigma,2)$ of the $W(\sigma)$ cycle satisfy:*

$$[A_l^{1/2} M_l^V(\sigma) A_l^{-1/2}] \le \frac{\theta}{\theta+\sigma}, \qquad [A_l^{1/2} M_l^W(\sigma) A_l^{-1/2}] \le \frac{\sqrt{\theta}}{\sqrt{\theta}+\sigma}. \tag{4.36}$$

5 Implementation Aspects and Alternative Formulations

In this section implementation aspects are addressed which are not necessarily deduced from the theoretical foundation above. Also alternative formulations are described together with the grounds for their rejection.

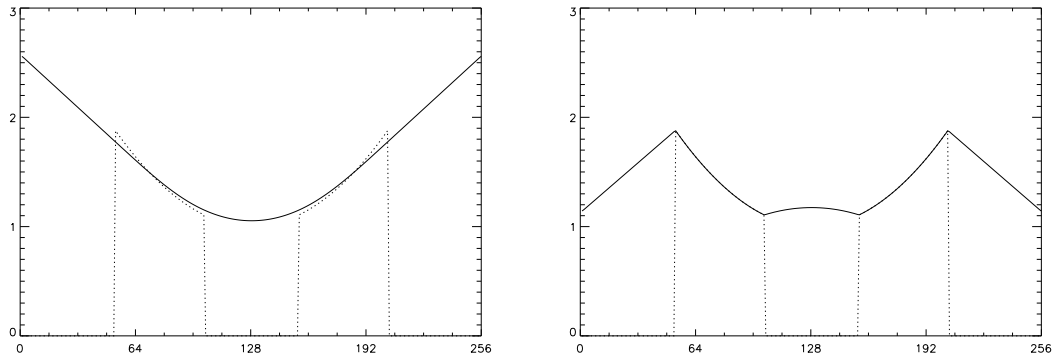

Figure 2: Solutions to (3.3) using the pure finite element method with the case that μ is large shown on the left, and the case that μ is small shown on the right. Here, $N = 1$, $\nu = 2$, r is shown in the dotted curve, and m is the characteristic function for the support of r.

5.1 Lumping Zero-Order Terms

In spite of the convergence guaranteed by Theorem 3.1 as $h \to 0$, the pure finite element method fails to converge as $\mu \to 0$ when the data m are discontinuous. For example let $N = 1$ and $\nu = 2$, and suppose r is the dotted curve shown in Figs. 2a and 2b, while m is the characteristic function for the support of r. Then the solution to (3.3) for large and small μ are shown in Figs. 2a and 2b respectively. This problem is corrected by using a lumping approach in which the Grammian matrix is diagonalized as follows:

$$(m^2\chi_\imath, \chi_\jmath)_0 \to \delta_{\imath\jmath} m_i^2 \|\chi_\imath\|^2 \tag{5.1}$$

where $\{\chi_\imath\}$ is a basis for $S_h^\nu(\Omega)$ and $m_\imath$ denotes m evaluated at the peak of $\chi_\imath$. With this modification, convergence in the limit of increasing data fidelity and vanishing regularization is obtained as explained in [10]. Furthermore, when the lumping in (5.1) is used, the grids, e.g., as illustrated in Fig. 1, can alternatively be taken to include the ghost cells, and the finite element coefficients can be taken as solution values at basis function peaks. As stated in the following theorem and shown in [10], the matrix $A_{l_{\max}} = A_h$ thus obtained is guaranteed to be symmetric and positive definite if the support of m is large enough. Therefore, according to (4.17), all the matrices A_l are symmetric and positive definite when the restriction operators have maximal rank. See [10] for further details.

Theorem 5.1 *Suppose that m is supported at least on a collection of adjacent cells in the finest grid on Ω with centroids $\{\boldsymbol{x}_\kappa : \kappa_1 \le \kappa \le \kappa_1 + \nu\}$. Then the coefficient matrix A_h is symmetric and positive definite.*

To consider the lumping in (5.1) within the theoretical convergence framework, consider the decomposition $L_h = L_h^\mu + L_h^0$ where L_h^μ and L_h^0 are defined as follows:

$$(L_h^\mu \chi_\imath, \chi_\jmath) = \mu\langle \chi_\imath, \chi_\jmath\rangle_\nu, \quad (L_h^0 \chi_\imath, \chi_\jmath) \to \delta_{\imath\jmath} m_{\imath\jmath}^2 \|\chi_\imath\|_0^2 \tag{5.2}$$

where $\{\chi_\imath\}$ is a basis for $S_h^\nu(\Omega)$. Then L_l^μ and L_l^0 are defined through (4.7) for $0 \le l \le l_{\max}$ and $A_l = A_l^\mu + A_l^0$ is similarly decomposed according to (4.17) for $0 \le l \le l_{\max}$. Also, A_l^μ is conveniently computed according to (4.11),

$$A_{l,\imath\jmath}^\mu = \mu\langle \chi_\imath, \chi_\jmath\rangle_\nu \tag{5.3}$$

while A_l^0 is computed iteratively according to (4.17),

$$A_{l-1}^0 = R_l^{l-1} A_l^0 E_{l-1}^l. \tag{5.4}$$

The relations (4.16) and (4.17) hold in any case for A_l, A_l^μ and A_l^0. Then, instead of with (4.26), $D_{l,\imath\imath}^{-1}A_{l,\imath\jmath}$ is estimated as follows:

$$\begin{aligned} |D_{l,\imath\imath}^{-1}A_{l,\imath\jmath}| &= \left|\frac{\mu\langle\chi_\imath,\chi_\jmath\rangle_\nu + \delta_{ij}m_i^2\|\chi_\imath\|_0^2}{\mu\langle\chi_\imath,\chi_\imath\rangle_\nu + m_i^2\|\chi_\imath\|_0^2}\right| \leq \delta_{ij} + (1-\delta_{ij})\frac{|\chi_\imath|_\nu|\chi_\jmath|_\nu}{|\chi_\imath|_\nu^2} \\ &\leq \delta_{ij} + (1-\delta_{ij})\frac{|\chi_\jmath|_\nu}{|\chi_\imath|_\nu} \leq \delta_{ij} + (1-\delta_{ij})\frac{\max_\imath|\chi_\imath|_\nu}{\min_\imath|\chi_\imath|_\nu}. \end{aligned} \tag{5.5}$$

5.2 Multi-Colored Ordering

A multi-colored ordering of unknowns is used in order to facilitate vectorization of the relaxation scheme; see also [13]. Specifically, the diameter of the stencil of the finite element discretization of the 2ν-th order differential operator in (2.10) is $(2\nu+1)$ and the number of weights in the stencils is $(2\nu+1)^N$. Thus, the set of all cells with centroids of the form $\{\boldsymbol{x}_{\kappa+k(\nu+1)e_i}\}$ with $k \in \boldsymbol{Z}$, $(e_i)_j = \delta_{ij}$, and fixed κ, have stencils which do not weight any other cells in the set. Such cells are painted with the same color, ordered sequentially within that color, and then ordered sequentially among the $(\nu+1)^N$ colors. For example, when $\nu = 1$ and $N = 2$, there are 4 colors and the matrices of (4.20) have the form:

$$\mathcal{D}_l = \begin{bmatrix} \mathcal{D}_l^{(1)} & & & \\ & \mathcal{D}_l^{(2)} & & \\ & & \mathcal{D}_l^{(3)} & \\ & & & \mathcal{D}_l^{(4)} \end{bmatrix} \quad \mathcal{L}_l = \begin{bmatrix} & & & \\ \mathcal{L}_l^{(21)} & & & \\ \mathcal{L}_l^{(31)} & \mathcal{L}_l^{(32)} & & \\ \mathcal{L}_l^{(41)} & \mathcal{L}_l^{(42)} & \mathcal{L}_l^{(43)} & \end{bmatrix} \tag{5.6}$$

where each row/column in these forms represents a separate color and the blocks $\mathcal{L}_l^{(\imath\jmath)}$ represent the couplings among colors. With a multi-colored ordering, (4.18) can be implemented by performing a Jacobi iteration on same-colored cells while looping in one direction over the colors:

$$\begin{aligned} &\text{for } c = 1,\ldots,(\nu+1)^N \\ &\qquad U_l^c \leftarrow U_l^c - \omega[\mathcal{D}_l^{-1}(A_lU_l - F_l)]^c \end{aligned} \tag{5.7}$$

and then looping in the opposite direction over the colors:

$$\begin{aligned} &\text{for } c = (\nu+1)^N,\ldots,1 \\ &\qquad U_l^c \leftarrow U_l^c - \omega[\mathcal{D}_l^{-1}(A_lU_l - F_l)]^c. \end{aligned} \tag{5.8}$$

Thus, same-colored cells may be updated simultaneously.

The known stencil diameter can also be used to advantage to vectorize the computation of elements of the coarse grid matrix according to (5.4). Specifically, note that if $V_\kappa = 1$ in the cell $\boldsymbol{x}_\kappa$ and $V_\kappa = 0$ otherwise, then $\tilde{V}_\kappa = A_lV_\kappa$ is non-zero in a neighborhood of $\boldsymbol{x}_\kappa$ of diameter $(2\nu+1)^N$ cells, and the values of $\tilde{V}_\kappa$ give the elements of the c_κth column of A_l where c_κ is the index of cell $\boldsymbol{x}_\kappa$ in the multi-colored ordering. Similarly, define $U_\kappa = 1$ on all the cells $\{\boldsymbol{x}_{\kappa+k(2\nu+1)e_i}\}$ with $k \in \boldsymbol{Z}$, $(e_i)_j = \delta_{ij}$, and fixed κ. Then for the $(2\nu+1)^N$ cells surrounding a given $\boldsymbol{x}_{\kappa+k(2\nu+1)e_i}$, the product $\tilde{U}_\kappa = A_lU_\kappa$ expands the support of U_κ from the single cell $\boldsymbol{x}_{\kappa+k(2\nu+1)e_i}$ to the support of $\tilde{U}_\kappa$ in all $(2\nu+1)^N$ cells in the neighborhood without affecting values in cells of adjacent neighborhoods. Thus, $\tilde{U}_\kappa$ simultaneously gives all columns $\{c_{\kappa+k(2\nu+1)e_i}\}$ of A_l. Since there are $(2\nu+1)^N$ distinct functions U_κ, the computation of the columns of A_l can be vectorized into a loop of length $(2\nu+1)^N$.

5.3 Floating Point Errors

The floating point aspects discussed in this subsection underscore the fact that Theorem 5.1 is only valid in exact arithmetic. Note that the image processing environment IDL is used

for this work and that double precision is used for all floating point variables. According to Theorem 5.1, the Gauss-Seidel iteration (4.18) must converge. When it was found not to converge in initial versions of the code, the positive definiteness of A_l was tested and found to fail because of treatable floating point problems discussed here.

Consider the minimum and maximum magnitudes of stencil values for two-dimensional problems with the increasing orders shown in Table 1. Just like the usual one-dimensional approxi-

ν	$\min\{\lvert A^{\mu}_{l,\imath\jmath}\rvert\}/\mu$	$\max\{\lvert A^{\mu}_{l,\imath\jmath}\rvert\}/\mu$
1	1/3	8/3
2	152/2880	24768/2880
3	135/45360	1353600/45360
4	796672/3715891200	400229580800/3715891200

Table 1: For problems in which the operator of (2.10) is order 2ν, the minimum and maximum magnitudes of stencil values of (5.3) are shown in the respective columns.

mations of higher-order derivatives, the two-dimensional stencils also have signs which alternate with distance from the stencil center. When multiplications with such values are sufficiently inaccurate, the Galerkin multiplication (4.17) can produce a representation of a coarse grid matrix which is so poor as not even to be positive definite. This fact was demonstrated explicitly by finding a vector U in the course of the Gauss-Seidel iteration for which the product $[A_l U, U]$ is negative. This problem was first corrected by using (5.4) only to compute the zero-order terms by the Galerkin multiplication and otherwise to compute the elements of the coarse grid matrix directly according to (5.3). However, a later Gauss-Seidel iteration again revealed a negative product $[A_l U, U]$, and the cause was shown explicitly to be the inaccurate computation of scalar products in matrix-vector multiplications. After all scalar products were reprogramed to be computed with the IDL command `total`, the problem was corrected and convergence of the Gauss-Seidel iteration was obtained.

When for an eighth-order problem ($\nu = 4$) it was once again found possible for the Gauss-Seidel convergence to fail, it remained to determine whether the matrix A_l was being *stored* with sufficient inaccuracy to lose its positive-definiteness. Such was demonstrated as follows to be the case. Let the multiplicative factors in the stencil elements be separated into μ and otherwise integer components according to $A^{\mu}_{l,\imath\jmath} = \mu a_{\imath\jmath}/(dh^{2\nu}2^{N(\nu-1)(l_{\max}-l)})$, where for instance Table 1 gives $d = 3715891200$ for the case $N = 2$ and $\nu = 4$. Then let μ be normalized according to

$$\tilde{\mu} = \frac{\mu}{dh^{2\nu}} \tag{5.9}$$

and let the matrix be correspondingly redefined according to $A^{\tilde{\mu}}_{l,\imath\jmath} = \tilde{\mu} a_{\imath\jmath}/2^{N(\nu-1)(l_{\max}-l)}$. Once $\tilde{\mu}$ was chosen, for instance as a power of 2, so that $A^{\tilde{\mu}}_{l,\imath\jmath}$ can be stored exactly, then convergence of the Gauss-Seidel iteration was obtained for the eighth-order problem. Thus, problems of all orders were so programed.

Because of such floating point considerations, it was also considered to perform the following pre- and post-processing steps before and after relaxation on every level. In the pre-processing step, the current problem can be solved directly on the polynomial kernel of A^{μ}_l. Then the current problem can be relaxed, in effect, on the orthogonal complement of the kernel. In the post-processing step, the components of the direct solve can be combined with the results of relaxation. In spite of the apparent potential accuracy advantage in these steps, they were found to have no real effect on the solution.

5.4 Effect of Parameters on Neumann and Dirichlet Formulations

To estimate the effect of order 2ν, regularization μ, relaxation ω, and data support on the multigrid reduction factor $[A_l^{1/2} M_l^{\mathrm{MGC}} A_l^{-1/2}]$, one-dimensional model problems are used.

Specifically, for $N = 1$ let the data m be the characteristic function as defined for Fig. 2 when the data are compactly support. Otherwise let $m = 1$ hold globally. Also assume that $2^p = 256$, $l_{\max} = 4$, $\sigma = 1$ and $\tau = 1$ hold. Then using MATLAB the norm $[A_{l_{\max}}^{1/2} M_{l_{\max}}^{V(1)} A_{l_{\max}}^{-1/2}]$ is plotted in Fig. 3 as a function of 2ν and $\log_{10}(\mu/h^{2\nu})$ for fixed $\omega = 1$. The norm is shown in the left and right columns, respectively, for homogeneous Neumann ($d^k u/dx^k = 0$, $k = \nu, \ldots, 2\nu - 1$, $\partial\Omega$) and homogeneous Dirichlet ($d^k u/dx^k = 0$, $k = 0, \ldots, \nu - 1$, $\partial\Omega$) boundary conditions. The norm is shown in the top and bottom rows for data m having compact and full support, respectively.

While the reduction factor shown in Fig. 3 is smaller for Dirichlet rather than for Neumann boundary conditions when $\nu = 1$ holds, the trend reverses as the order increases. Observe further in Fig. 3 that, for vanishingly small μ, the graphs reflect the fact that the coefficient matrix of (3.3) becomes ever more indefinite when the data are compactly supported or else ever more nearly the identity when the data are fully supported. The practical range within which μ should be chosen is also suggested by Fig. 3. On the other hand, as discussed in [18], μ should be chosen according to data fidelity by matching the magnitude of the two terms in (2.1) when the first term is essentially the data variance.

The transition in the reduction factor with respect to regularization is seen in Fig. 3 to be in the range $0 \leq \log_{10}(\mu/h^{2\nu}) \leq 5$. Therefore, the norm $[A_{l_{\max}}^{1/2} M_{l_{\max}}^{V(1)} A_{l_{\max}}^{-1/2}]$ is plotted in Fig. 4 as a function of 2ν and ω for fixed $\mu = h^{2\nu} \times 10^5$. Again, the norm is shown in the left and right columns, respectively, for Neumann and Dirichlet boundary conditions, and it is shown in the top and bottom rows for data m having compact and full support, respectively. Note that as the order increases, the optimal ω shifts from 1 to values nearer to 2. For the examples presented in Section 6, $\omega = 1$ is used.

While Figs. 3 and 4 suggest numerical advantages of the Neumann boundary conditions for higher-order problems, the following two-dimensional model problem shows that the Dirichlet formulation fails to deliver a natural solution to the early vision problem. Specifically, for $\nu = 1$, let the data in (2.11) be given by $m = \chi_S$ where $S = (\frac{1}{4}, \frac{3}{4}) \times (\frac{1}{4}, \frac{3}{4})$ and $r = m \cdot u^*$ where $u^* = \exp(x - 1)\exp(y - 1)$ is shown on the right in Fig. 5. The images on the left and in the middle are approximate solutions to (3.3) obtained respectively by using Dirichlet and Neumann boundary conditions in an FMG-$V(1)$ iteration with additional MGC-$V(1)$ iterations performed until a relative error criterion of $[U_h^k - U_h^{k-1}] < 10^{-7}[U_h^k]$ is met. All input parameters are equal in the two calculations: $l_{\max} = 5$, $2^p = 256$, $\nu = 1$, $N = 2$, $\sigma = 1$, $\tau = 1$, $\omega = 1$, and $\tilde{\mu} = 2^{-1}$. The respective errors of the finite difference and finite element solutions are quantified in Table 2. Note that the Neumann estimate of u^* is conspicuously more natural, although a

	$[A_h U_h - F_h]$	$\|\chi_m(u - U_h)\|_{\ell_\infty}$	$\|u - U_h\|_{\ell_\infty}$
Dirichlet:	6.7×10^{-13}	0.043	1.0
Neumann:	1.3×10^{-13}	0.016	0.54

Table 2: For the results shown in Fig. 5, the equation residuals are shown here in the $[\cdot]$-norm in the first column, the solution errors on the data support ($\chi_m = (m \neq 0)$) are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the second column, and the globally support solution errors are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the third column.

still higher order provides an improvement as seen in Section 6.

5.5 Finite Difference Formulation

In a first multigrid implementation it was considered to compute the operators of (4.17) independently of finite elements as follows. For a given ν, let the expansion operator be determined as a transformation from cell centers in a coarse grid to cell centers in a fine grid. This transformation is determined by first extrapolating on the coarse grid to cell centers of ν layers of ghost cells outside Ω. The resulting values on the larger set of cell centers are then transformed into values on the cell centers of the fine grid in such a way that the transformation

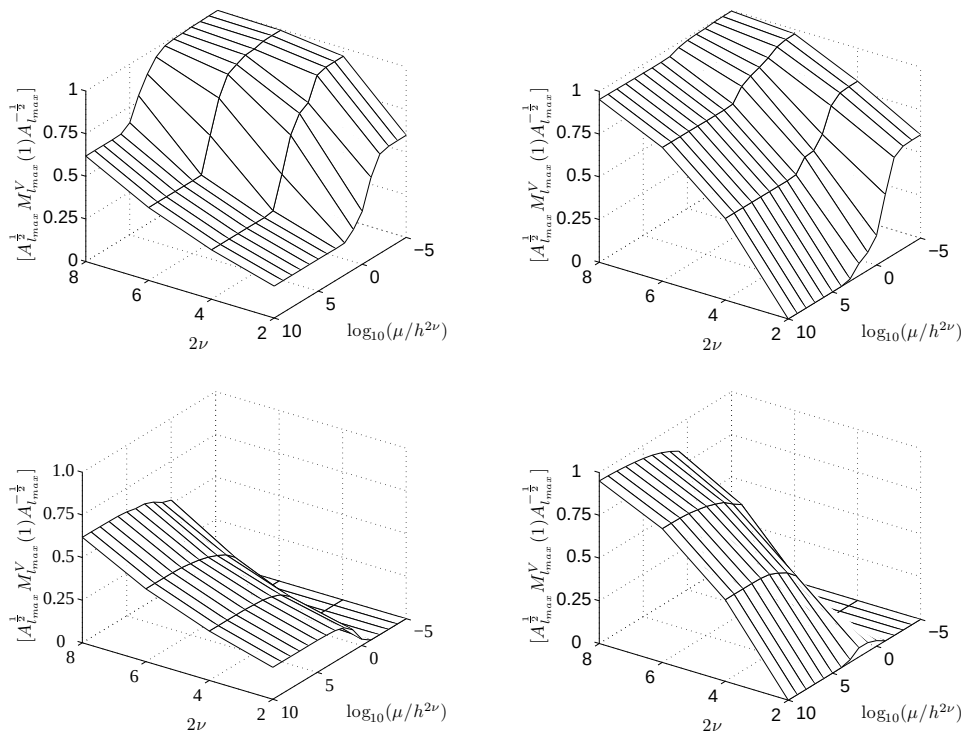

Figure 3: Multigrid reduction factor $[A_{l_{\max}}^{1/2} M_{l_{\max}}^{V(1)} A_{l_{\max}}^{-1/2}]$ as a function of order 2ν and regularization $\log_{10}(\mu/h^{2\nu})$ for fixed $\omega = 1$. The norm is shown in the left and right columns, respectively, for Neumann and Dirichlet boundary conditions, and it is shown in the top and bottom rows for data m having compact and full support, respectively.

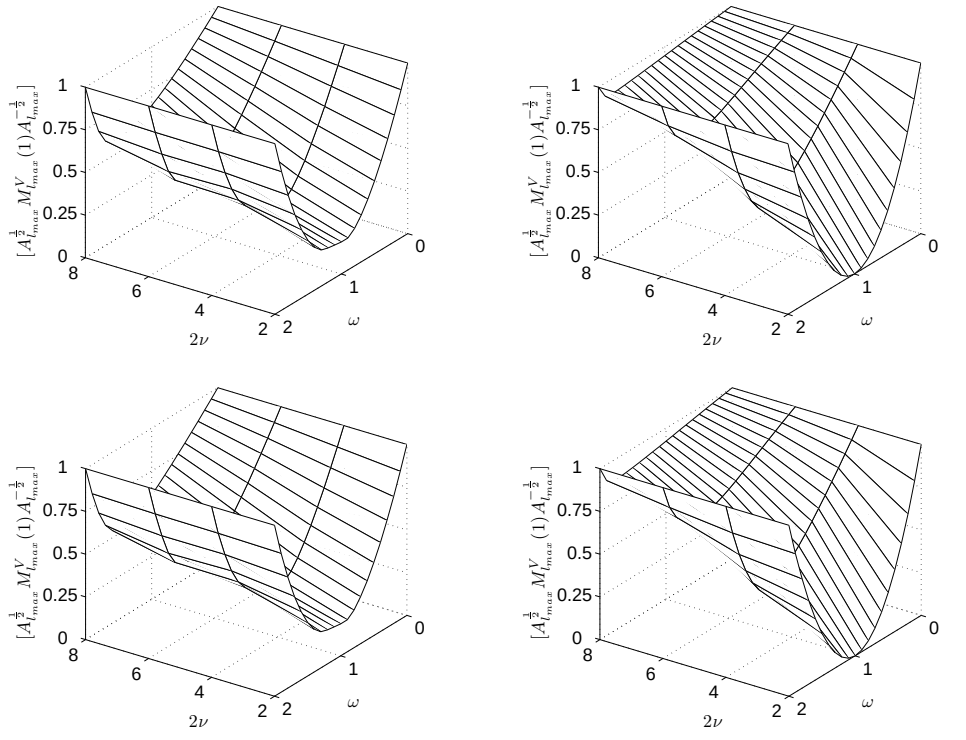

Figure 4: Multigrid reduction factor $[A_{l_{\max}}^{1/2} M_{l_{\max}}^{V(1)} A_{l_{\max}}^{-1/2}]$ as a function of order 2ν and relaxation ω for fixed $\mu = h^{2\nu} \times 10^5$. The norm is shown in the left and right columns respectively for Neumann and Dirichlet boundary conditions, and it is shown in the top and bottom rows for data m having compact and full support, respectively.

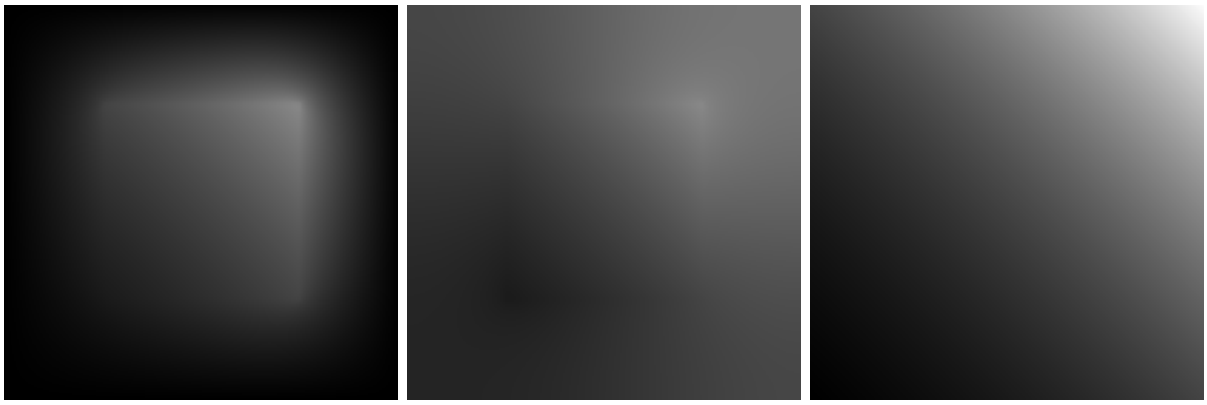

Figure 5: On the left (a) and in the middle (b) are approximate solutions to (3.3) for $\nu = 1$ obtained respectively by using Dirichlet and Neumann boundary conditions. The data m and r satisfy $r = m \cdot u^*$ where $u^* = \exp(x-1)\exp(y-1)$ is shown (c) on the right. Note that the data support is clearly visible in the two left images but less so in (b) than in (a).

is exact for polynomials of degree ν. While one might consider to construct the restriction operator similarly, the canonical transpose of (4.16) is much more accurate and of course maintains symmetry. With this finite difference approach to the intergrid transfer operators, the discrete coarse grid operator is defined by the Galerkin multiplication (4.17). In spite of the apparent naturalness of this construction, the following example shows that it is inadequate.

For $\nu = 2$ let the data in (2.11) be given by $m = \chi_S$ where $S = (\frac{1}{4}, \frac{3}{4}) \times (\frac{1}{4}, \frac{3}{4})$ and $r = m \cdot u^*$ where $u^* = (x - \frac{1}{2})^3 (y - \frac{1}{2})^3$. In order to test numerical accuracy, let $U^*_{l_{\max}}$ be the discretization of u^*, and in (4.14) let $F_{l_{\max}} = A_{l_{\max}} U^*_{l_{\max}}$. Then the image on the right in Fig. 6 is the exact

Figure 6: On the left (a) and in the middle (b) are approximate solutions to (2.11) for $\nu = 2$ using respectively the finite difference and the finite element formulation in an FMG-$V(1)$ iteration with all input parameters equal. The exact solution is shown on the right (c).

solution $U^*_{l_{\max}}$. The images on the left and in the middle are approximate solutions to (2.11) obtained respectively by using the finite difference and finite element formulations in an FMG-$V(1)$ iteration with all input parameters equal: $l_{\max} = 5$, $2^p = 256$, $\nu = 2$, $N = 2$, $\sigma = 2$, $\tau = 1$, $\omega = 1$, and $\tilde{\mu} = 2^{-35}$. The respective errors of the finite difference and finite element solutions are quantified in Table 3. Note that the accuracy of the finite element solution is conspicuously higher.

6 Computational Examples

In this section a model surface estimation problem is solved using increasingly higher orders of regularization, and the results are compared in terms of accuracy and multigrid convergence.

	$[A_h U_h - F_h]$	$\|\chi_m(u - U_h)\|_{\ell_\infty}$	$\|u - U_h\|_{\ell_\infty}$
finite differences:	8.2×10^{-16}	5.7×10^{-12}	0.34
finite elements:	3.3×10^{-18}	1.1×10^{-13}	5.1×10^{-5}

Table 3: For the results shown in Fig. 6, the equation residuals are shown here in the $[\cdot]$-norm in the first column, the solution errors on the data support ($\chi_m = (m \neq 0)$) are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the second column, and the globally support solution errors are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the third column.

The chosen model problem is representative of magnetic resonance coil sensitivity estimation [10]. Specifically, the surface $u^* = \exp(x-1)\exp(y-1)$, shown in Fig. 5c, satisfies $m \cdot u^* = r$ for data m and r which are shown in Fig. 7, and u^* is to be estimated by u which minimizes

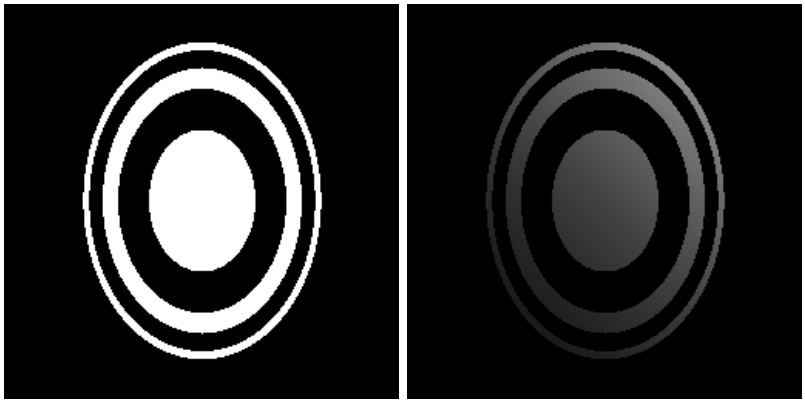

Figure 7: The model data m and r in (2.1) are shown respectively on the left and right. These data satisfy $mu^* = r$ for u^* shown in Fig. 5c.

(2.1). The estimate u is computed first for noise free data and then for noise corrupted data.

For the noise free data shown in Figs. 7a and 7b, approximate solutions U to (4.14) are shown in Fig. 8 for orders $\nu = 1$ to $\nu = 4$. In each case, U is obtained using an FMG-$V(1)$ iteration with additional MGC-$V(1)$ iterations performed until a relative error criterion of $[U_h^k - U_h^{k-1}] < 10^{-7}[U_h^k]$ is met. All the following input parameters are equal in these calculations: $l_{\max} = 5$, $2^p = 256$, $N = 2$, $\sigma = 1$, $\tau = 1$, and $\omega = 1$. Furthermore, following Fig. 3, μ is chosen for each ν so that $\mu h^{-2\nu} \approx 1$. Specifically, $\tilde{\mu} = 2^{-1}$ for $\nu = 1$, $\tilde{\mu} = 2^{-12}$ for $\nu = 2$, $\tilde{\mu} = 2^{-16}$ for $\nu = 3$, and $\tilde{\mu} = 2^{-32}$ for $\nu = 4$. Note the following aspects of the results shown in Fig. 8. In Fig. 8a the data support is clearly visible but the other estimates provide a more natural extrapolation of the data. In the lower left of Fig. 8b the values of the extrapolation are negative and are wrapped to larger intensity values in the interval $[0, 1]$. In Figs. 8c and Figs. 8d the extrapolations are less flat and brighter in their brightest points. Such results are to be expected on the basis of the natural boundary conditions which support polynomial behavior of degree $\nu - 1$ at the boundary.

The respective errors of these cases are quantified in Table 4. On the basis of Theorems 2.3 and 3.1, the solutions U can be expected to converge to u^* only as the data support fills Ω and μ vanishes at a certain rate. Nevertheless, the higher-order solutions clearly provide a more accurate extrapolation. The lower-order solutions are even less accurate on the data support, apparently since their lower-order boundary conditions divert the estimate away from a suitable growth rate. The higher-order solutions are clearly more accurate globally.

Now let the data shown in Figs. 7a and 7b be corrupted by 10% uniformly distributed noise. Then the approximate solutions U to (4.14) are shown in Fig. 9 for orders $\nu = 1$ to $\nu = 4$. As with Fig. 8, U is obtained in each case here using an FMG-$V(1)$ iteration with additional MGC-$V(1)$ iterations performed until a relative error criterion is met. The relative error criterion is $[U_h^k - U_h^{k-1}] < 10^{-7}[U_h^k]$ and it is met after 2 MGC iterations for $\nu = 1$, 3 MGC iterations for

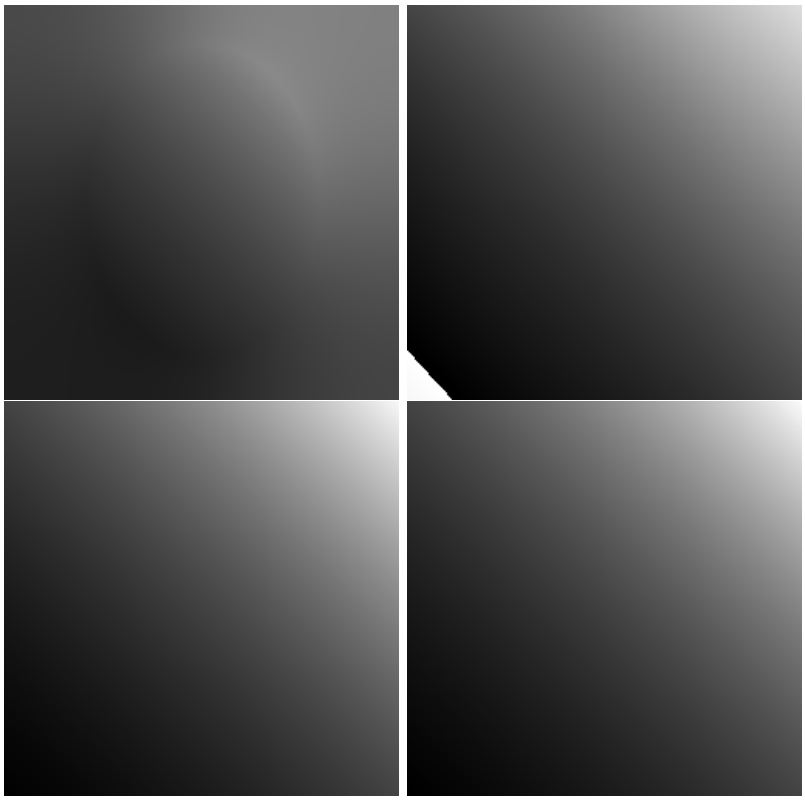

Figure 8: For the data shown in Figs. 7a and 7b, approximate solutions to (3.3) are shown here for (a) $\nu = 1$ at top left, (b) $\nu = 2$ at top right, (c) $\nu = 3$ at bottom left, and (d) $\nu = 4$ at bottom right. The data support is clearly visible in (a). Note in the lower left of (b) that the values of the extrapolation are negative and are wrapped to larger intensity values in the interval $[0, 1]$. Otherwise note that these results are progressively less flat or progressively brighter in their brightest points.

ν	$[A_h U_h - F_h]$	$\|\chi_m(u - U_h)\|_{\ell_\infty}$	$\|u - U_h\|_{\ell_\infty}$
1:	1.5×10^{-13}	5.4×10^{-3}	0.50
2:	9.1×10^{-12}	2.2×10^{-5}	0.12
3:	2.2×10^{-13}	3.2×10^{-8}	0.018
4:	2.8×10^{-15}	1.5×10^{-10}	0.0015

Table 4: For the results shown in Fig. 8, the equation residuals are shown here in the $[\cdot]$-norm in the first column, the solution errors on the data support ($\chi_m = (m \neq 0)$) are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the second column, and the globally support solution errors are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the third column.

$\nu = 2$, and 12 MGC iterations for $\nu = 3$. However, for the case $\nu = 4$ the convergence criterion is not met even after 44 MGC iterations when the norm $\|A_h^{1/2} e_h^k\|_{\ell_2}$, $e_h^k = U_h^k - U_h^{k-1}$, ceases to decrease. Nevertheless, as shown in Table 5, the equation residual for the case $\nu = 4$ is comparable to the other cases, although the trend of decreasing residuals evident in Table 4 is violated in Table 5 by $\nu = 4$. All the following input parameters are equal in these calculations: $l_{\max} = 5$, $2^p = 256$, $N = 2$, $\sigma = 1$, $\tau = 1$, and $\omega = 1$. Furthermore, μ is chosen for each ν so that $\mu h^{-2\nu} \approx 10^5$. Specifically, $\tilde{\mu} = 2^{16}$ for $\nu = 1$, $\tilde{\mu} = 2^5$ for $\nu = 2$, $\tilde{\mu} = 2^1$ for $\nu = 3$, and $\tilde{\mu} = 2^{-15}$ for $\nu = 4$. Note the following aspects of the results shown in Fig. 9. The data support is no longer clearly visible in Fig. 9a as it is in Fig. 8a, but the estimate for the case $\nu = 1$ is relatively flat exhibiting no particular brightness in the upper right corner as in other cases. In the lower left of Figs. 9b and 9c the values of the extrapolation are negative and are wrapped to larger intensity values in the interval $[0, 1]$. In contrast to Fig. 9a, Figs. 9b and

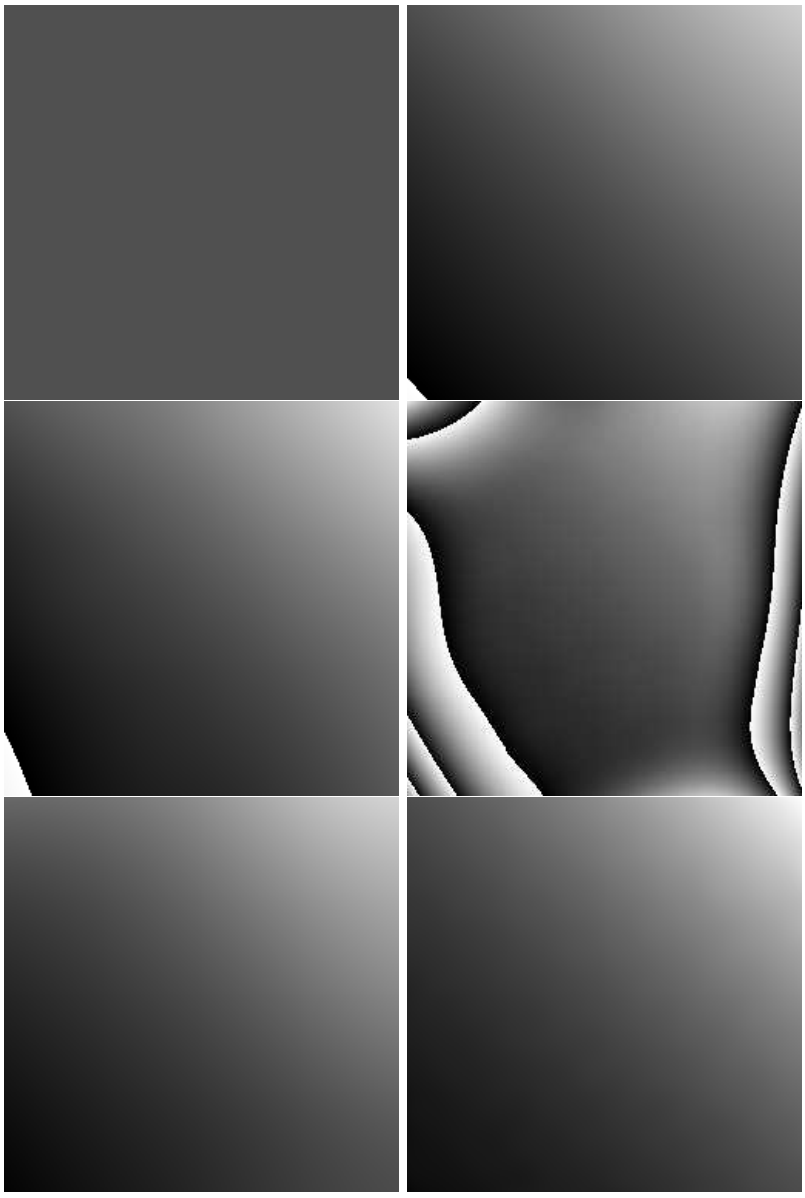

Figure 9: For the data shown in Figs. 7c and 7d, approximate solutions to (3.3) are shown here for (a) $\nu = 1$ at top left, (b) $\nu = 2$ at top right, (c) $\nu = 3$ at middle left, and (d) $\nu = 4$ at middle right, as well as (e) $\nu = 4$ with higher smoothing than in (d) and (f) $\nu = 4$ with full data support. Observe that in the lower left of (b) and (c) the values of the extrapolation are negative and are wrapped to larger intensity values in the interval $[0, 1]$. The values of the extrapolation in image (d) are wrapped through the interval $[0, 1]$ several times. Note that the images (a-c) are progressively less flat or progressively brighter in their brightest points, as is image (e) flatter than image (f).

9c are progressively less flat or progressively brighter in their brightest points. The values of the extrapolation in Fig. 9d are wrapped through the interval $[0, 1]$ several times. In spite of this obvious inaccuracy of extrapolation, it can be seen in Table 5 for the case $\nu = 4$ that the estimate is nevertheless accurate on the data support.

To achieve an improved extrapolation for $\nu = 4$, the estimates shown in Figs. 9e and 9f are obtained as follows. The image in Fig. 9e is obtained with the higher smoothing $\tilde{\mu} = 2^{-2}$ instead of 2^{-15} as used in Fig. 9d, and the image in Fig. 9f is obtained with the fully supported data given by a 10% uniformly distributed noise corruption of the data $m = 1$ and $r = u^*$. As with other images in Figs. 8 and 9, the numerical estimate U is obtained in each case here using an FMG-$V(1)$ iteration with additional MGC-$V(1)$ iterations performed until a relative error criterion is met. The relative error criterion $[U_h^k - U_h^{k-1}] < 10^{-7}[U_h^k]$ is met after 8

MGC iterations for Fig. 9e and after 6 MGC iterations for Fig. 9f. Thus, Fig. 9e shows that the inaccuracy of extrapolation apparent in Fig. 9d can be ameliorated with sufficiently high regularization. Nevertheless, the higher smoothing can be seen to be flattening by comparing the brightness of the brightest points in Figs. 9e and Figs. 9f. The difference between Fig. 9d and Fig. 9f clearly demonstrates the challenging effect of compactly supported data.

The respective errors of the cases demonstrated in Fig. 9 are quantified in Table 5. As in

conditions	ν	$[A_hU_h - F_h]$	$\|\chi_m(u - U_h)\|_{\ell_\infty}$	$\|u - U_h\|_{\ell_\infty}$
	1	1.9×10^{-8}	0.23	0.68
	2	9.2×10^{-9}	0.041	0.18
	3	2.8×10^{-11}	0.045	0.11
	4	7.3×10^{-9}	0.052	2.4
higher smoothing:	4	1.1×10^{-5}	0.044	0.16
full data support:	4	3.7×10^{-7}	0.052	0.052

Table 5: For the results shown in Fig. 9, the equation residuals are shown here in the $[\cdot]$-norm in the first column, the solution errors on the data support ($\chi_m = (m \neq 0)$) are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the second column, and the globally supported solution errors are shown in the $\|\cdot\|_{\ell_\infty}$-norm in the third column.

Table 8, it can be seen in Table 9 that the accuracy on the data support is the poorest for the case $\nu = 1$. Nevertheless, the error on the data support increases from $\nu = 2$. On the other hand, the global errors decrease up to $\nu = 3$. The global error jumps upward for the case $\nu = 4$, and while the global error is reduced considerably by higher smoothing, the minimal global error for compactly supported data is still achieved by $\nu = 3$. Of course for the case of fully supported data, the two solution errors reported in Table 5 are the same.

Graphical representations of the convergence of MGC-$V(1)$ and MGC-$W(1)$ iterations for $\nu = 1$ to $\nu = 4$ are shown in Fig. 10 by plotting on a $\log_{10}$-scale the differences $e_k = U_h^k - U_h^{k-1}$ as functions $\|e_h^k\|_{\ell_2}$ and $\|A_h^{1/2}e_h^k\|_{\ell_2}$ of iteration number k. In each case, noise corrupted data are used while the underlying exact surface is u^* shown in Fig. 5c. The iteration is carried out until the relative error criterion $[U_h^k - U_h^{k-1}] < 10^{-7}[U_h^k]$ is met, except for the case of $\nu = 4$ and compactly supported data in which after 52 MGC-$V(1)$ iterations and 11 MGC-$W(1)$ iterations the convergence criterion is not met and the norm $\|A_h^{1/2}e_h^k\|_{\ell_2}$ ceases to decrease. Therefore, the case of $\nu = 4$ and fully supported data is also shown in Fig. 10. This case is represented by the steeper dotted curves, while the other curves are convergence histories for compactly supported data. The fully supported data are a 10% uniformly distributed noise corruption of the data $m = 1$ and $r = u^*$, while the compactly supported data are a 10% uniformly distributed noise corruption of the data shown in Figs. 7c and 7d. Note in Fig. 10 that the norm $\|A_h^{1/2}e_h^k\|_{\ell_2}$ is reliably more monotone. For the results shown in Fig. 10, reduction factors $\rho^{(k)} = \|A_h^{1/2}e_h^k\|_{\ell_2}/\|A_h^{1/2}e_h^{k-1}\|_{\ell_2}$ and average CPU-times per MGC-iteration (programed in IDL on an Intel Pentium M 212MHz laptop with 1MB Cache) are shown in Table 6. Note that the MGC-$W(1)$ iteration converges more rapidly than the MGC-$V(1)$ iteration, but at higher CPU costs, although the absolute CPU-times can be expected to be smaller when programed in C++. Convergence can be seen to be frustrated by compactly supported data. Notice the fluctuation in the norm $\|e_h^k\|_{\ell_2}$ in relation to the monotonicity of the norm $\|A_h^{1/2}e_h^k\|_{\ell_2}$, particularly for the case of $\nu = 4$ with compactly supported data but also for the case of the MGC-$W(1)$ iteration with $\nu = 3$. The compactly supported data apparently leave considerable freedom for the approximate solution to fluctuate outside the data support although this effect is masked by the considerably less weight given to the data support complement in the norm $\|A_h^{1/2}e_h^k\|_{\ell_2}$. On the basis of these results, the MGC iterations can be expected to solve (3.3) accurately when the data are fully supported, but for compactly supported data a reliably accurate solution to

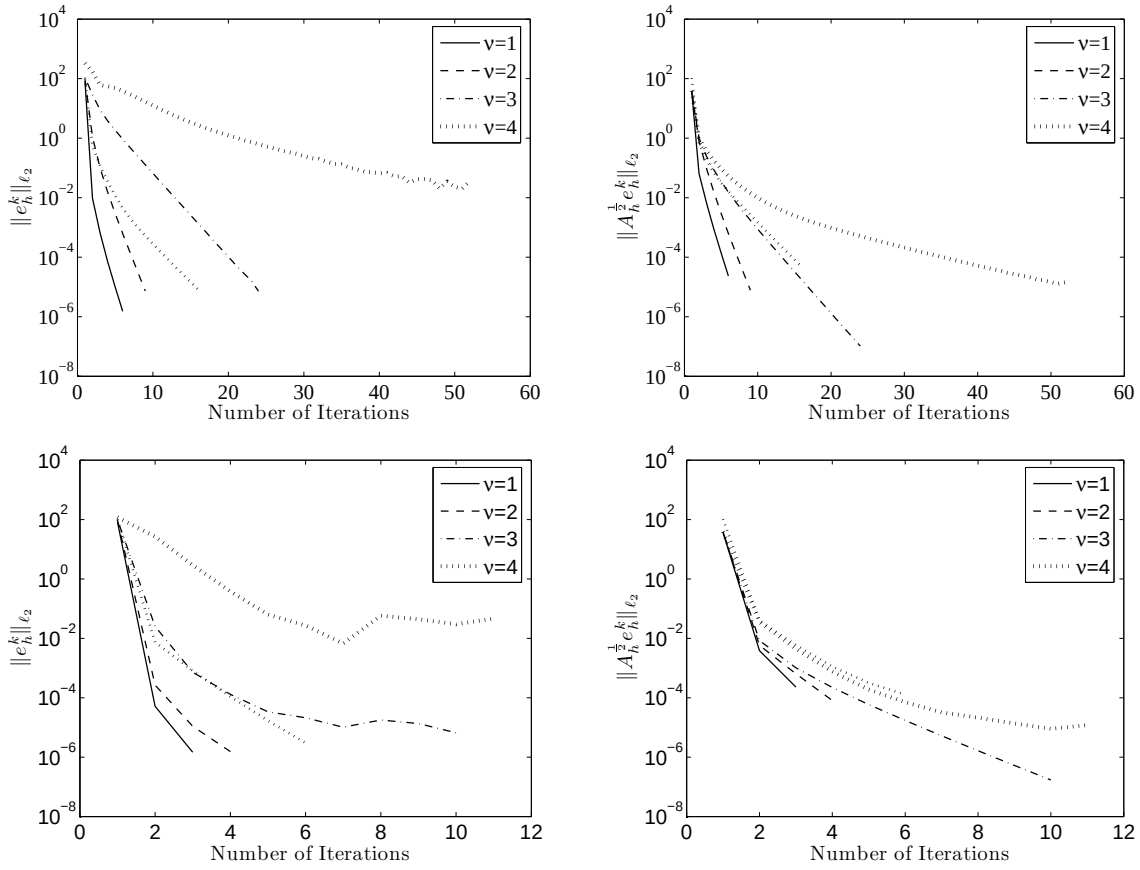

Figure 10: For $\nu = 1$ to $\nu = 4$, convergence histories of $e_h^k = U_h^k - U_h^{k-1}$ are shown here for (a) MGC-$V(1)$ iterations in the norm $\|e_h^k\|_{\ell_2}$ in the upper left, (b) MGC-$V(1)$ iterations in the norm $\|A_h^{1/2} e_h^k\|_{\ell_2}$ in the upper right, (c) MGC-$W(1)$ iterations in the norm $\|e_h^k\|_{\ell_2}$ in the lower left, and (d) MGC-$W(1)$ iterations in the norm $\|A_h^{1/2} e_h^k\|_{\ell_2}$ in the lower right. The steeper dotted curves are the convergence histories for $\nu = 4$ and full data support, while the other curves are convergence histories for compactly supported data.

(3.3) has been obtained here for at most $\nu = 3$.

References

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] J. Aubin, *Approximation of Elliptic Boundary-Value Problems*, Krieger, Huntington, New York, 1980.

[3] D. Choi and J.R. Kender, *Solving the Depth Interpolation Problem on a Parallel Architecture with a Multigrid Approach*, Proc. IEEE Computer Vision and Pattern Recognition Conf., CS Press, Los Alamitos, CA, pp. 189–194, 1988.

[4] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.

[5] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer, 1993.

[6] `http://www.rsinc.com/idl/index.cfm`.

[7] H.W. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

ν:	1	2	3	4	4
support:	compact	compact	compact	compact	full
reduction factor for MGC-$V(1)$:	$0.17_{\rho^{(6)}}$	$0.24_{\rho^{(8)}}$	$0.54_{\rho^{(22)}}$	$0.88_{\rho^{(50)}}$	$0.57_{\rho^{(15)}}$
reduction factor for MGC-$W(1)$:	$0.06_{\rho^{(2)}}$	$0.12_{\rho^{(3)}}$	$0.32_{\rho^{(9)}}$	$0.66_{\rho^{(9)}}$	$0.40_{\rho^{(5)}}$
CPU-time per MGC-$V(1)$ iteration:	1.1 sec	4.6 sec	13 sec	31 sec	31 sec
CPU-time per MGC-$W(1)$ iteration:	1.7 sec	6.5 sec	18 sec	42 sec	42 sec

Table 6: For the results shown in Fig. 10 reduction factors $\rho^{(k)} = \|A_h^{1/2} e_h^k\|_{\ell_2} / \|A_h^{1/2} e_h^{k-1}\|_{\ell_2}$ and average CPU-times per MGC-iteration (programed in IDL on an Intel Pentium M 212MHz laptop with 1MB Cache) are shown here.

[8] A. Friedman, *Partial Differential Equations*, Holt, Rinehart and Winston, Inc., New York, 1969.

[9] K. Höllig, *Finite Element Methods with B-Splines*, Frontiers in Applied Mathematics 26, SIAM, 2003.

[10] S.L. Keeling and R. Bammer, *A Variational Approach to Magnetic Resonance Coil Sensitivity Estimation*, Appl. Math. Comp., Vol. 158, pp. 359–388, 2004.

[11] S.H. Lai and B.C. Vemuri, *Physically-based Adaptive Preconditioning for Early Vision*, IEEE Trans. Pattern Anal. Machine Intell., Vol. 19, No. 6, pp. 594–607, 1997.

[12] U. Langer, *Iterative Solution of Some Schemes of the Finite Element Method for Elliptic Equations of Order* $2n$*,* $n \geq 1$, U.S.S.R. Comput. Maths. Math. Phys., Vol. 23, No. 4, pp. 69–76, 1983.

[13] G. Meurant, *Computer Solution of Large Linear Systems*, Elsevier, New York, 1999.

[14] `http://www.mathworks.com`.

[15] A.P. Pentland, *Interpolation Using Wavelet Bases*, IEEE Trans. Pattern Anal. Machine Intell., Vol. 16, No. 4, pp. 410–414, 1994.

[16] T. Poggio, V. Torre, and C. Koch, *Computational Vision and Regularization Theory*, Nature, Vol. 317, pp 314-319, 1985.

[17] K. Rektorys, *Variational Methods in Mathematics, Science and Engineering*, Reidel, Dordrecht, Boston, 1977.

[18] L.I. Rudin, S. Osher, and E. Fatemi, *Total Variation Based Noise Removal Algorithms*, Physica D, Vol. 60, pp. 259–269, 1992.

[19] R. Szeliski, *Fast Surface Interpolation Using Hierarchical Basis Functions*, IEEE Trans. Pattern Anal. Machine Intell., Vol. 12, No. 6, pp. 513–528, 1990.

[20] D. Terzopoulos, *The Computation of Visible Surface Representations*, IEEE Trans. Pattern Anal. Machine Intell., Vol. 10, No. 4, pp. 417–437, 1988.

[21] U. Trottenberg, C. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, San diego, San Francisco, New York, Boston, London, Sydney, Tokyo, 2001.

[22] M.-H. Yaou and W.-T. Chang, *Fast Surface Interpolation Using Multiresolution Wavelet Transform*, IEEE Trans. Pattern Anal. Machine Intell., Vol. 16, No. 7, pp. 673–688, 1994.