Mathematical Models and Methods in Applied Sciences

Multigrid solvers for isogeometric discretizations of the second biharmonic problem

Jarle Sogn
Department of Mathematics,
University of Oslo,
Postboks 1053 Blindern, 0316 Oslo, Norway
jarlesog@math.uio.no

Stefan Takacs
Institute of Computational Mathematics,
Johannes Kepler University Linz,
Altenberger Str. 69, 4040 Linz, Austria
stefan.takacs@numa.uni-linz.ac.at

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

We develop a multigrid solver for the second biharmonic problem in the context of Isogeometric Analysis (IgA), where we also allow a zero-order term. In a previous paper, the authors have developed an analysis for the first biharmonic problem based on Hackbusch's framework. This analysis can only be extended to the second biharmonic problem if one assumes uniform grids. In this paper, we prove a multigrid convergence estimate using Bramble's framework for multigrid analysis without regularity assumptions. We show that the bound for the convergence rate is independent of the scaling of the zero-order term and the spline degree. It only depends linearly on the number of levels, thus logarithmically on the grid size. Numerical experiments are provided which illustrate the convergence theory and the efficiency of the proposed multigrid approaches.

Keywords: Biharmonic problem; Isogeometric Analysis; Multigrid methods.

AMS Subject Classification: 31A30, 65D07, 65N55

1. Introduction

We consider multigrid methods for biharmonic problems discretized by Isogeometric Analysis (IgA). In particular, we consider the following model problem: Given a bounded domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, with Lipschitz boundary $\partial\Omega$, a parameter $\beta \geq 0$ and sufficiently smooth functions f, g_1, and g_2, find a function u such that

$$\begin{aligned} \beta u + \Delta^2 u &= f \quad \text{in} \quad \Omega, \\ u &= g_1 \quad \text{on} \quad \partial\Omega, \\ \Delta u &= g_2 \quad \text{on} \quad \partial\Omega \end{aligned} \tag{1.1}$$

holds in a variational sense. For $\beta = 0$, this problem is known as the *second biharmonic problem*, which is of interest for plate theory (cf. Ref. 7) and Stokes streamline equations (cf. Ref. 10). Problems with $\beta > 0$ are of particular interest in the context of optimal control problems, where the constraint is a second order elliptic operator. The optimality systems associated to these optimal control problems can be preconditioned robustly using preconditioners that rely on solving (1.1), see Refs. 21, 30, 1, 22. The problem (1.1) is obtained when considering the full observation; if one considers an optimal control problem with limited observation, one would obtain a similar problem, where the mass term βu is multiplied with the characteristic function for the observation domain.

We derive a standard variational formulation of the model problem, which lives in the Sobolev space $H^2(\Omega)$. For the discretization, we use Isogeometric Analysis (IgA) since it easily allows for H^2-conforming discretizations. Particularly, we consider a discretization based on tensor product B-splines of some degree $p > 1$ and maximum smoothness, i.e., $p - 1$ times continuously differentiable. For the derivation of the multigrid solver, we set up a hierarchy of grids as obtained by uniform refinement. Since we keep spline degree and spline smoothness fixed, we obtain nested spaces.

Concerning the choice of the smoother, there are many possibilities. We are interested in a smoother that yields a p-robust multigrid method. The first p-robust multigrid solvers for isogeometric analysis were based on the boundary corrected mass smoother (see Ref. 16) and the subspace corrected mass smoother (see Ref. 15). Both have been formulated for the Poisson problem. Since the subspace corrected mass smoother is more flexible and has proven itself more efficient in practice, we restrict ourselves to that smoother. The multigrid solvers with subspace corrected mass smoother have been extended to the first biharmonic problem in Ref. 29 and to the second and third biharmonic problem in the thesis Ref. 28. The convergence estimates are shown using the standard splitting of the analysis into approximation property and smoothing property, as proposed by Hackbusch, cf. Ref. 13.

The theory in all of these papers requires that the grids are uniform since they have been based on the p-robust approximation error estimates from Ref. 32, which are valid only in this case. Since then, newer p-robust approximation error estimates, see Refs. 26, 25, have been proposed, which do not require uniform grids. Using these new estimates, it is straightforward to relax this assumption and to show analogous results for the Poisson problem as well as the first biharmonic problem for quasi uniform grids. However, this is not straightforward for the second biharmonic problem, since the proof requires a certain commutativity property (cf. Lemma 9.2 in Ref. 28), which is only valid in case of uniform grids.

In this paper, we go another way. We base the analysis on the framework introduced by Bramble et al., cf. Refs. 3, 2. This allows us to drop the requirement that the grids are uniform. While this analysis could also be performed for other kinds of boundary conditions, like the first biharmonic problem, we restrict ourselves to the second biharmonic problem since it has previously turned out to be the more challenging one. For this setting, we prove a multigrid convergence estimate (Corollary 5.1) which is robust with respect to the spline degree p and which only depends logarithmically on the grid size h.

Moreover, that convergence result is robust in the parameter $\beta \geq 0$. This analysis is motivated by the mentioned optimal control problem. Such parameter-robust multigrid solvers are also known for the Poisson problem, see Ref. 23 for an analysis based on Hackbusch's framework. There, the authors also provide a regularity result for the corresponding partial differential equation (PDE), which is based on standard results for the Poisson problem. In our case, we do not need to do that since Bramble's analysis is not based on any regularity assumptions.

In the numerical experiments, one can observe that the convergence of a multigrid solver with subspace corrected mass smoother degrades if the geometry gets distorted. While this is also true for the Poisson problem, this dependence is significantly amplified for the biharmonic problem. The reason for the geometry dependence of the convergence rates is that the subspace corrected mass smoother is based on the tensor product structure of the spline space. This tensor product structure is distorted by the geometry mapping. So, the contributions of the geometry function are ignored when setting up the smoother. We aim to overcome this problem by considering a hybrid smoother that combines the proposed smoother with Gauss-Seidel sweeps, see also Refs. 29, 28.

Alternative smoothers based on overlapping multiplicative Schwarz techniques have been considered in Refs. 8, 22. Both approaches give good numerical results for the biharmonic problem. However, there is no rigorous, p-robust convergence theory available for these methods. It is worth mentioning that, as an alternative for solving biharmonic problems on the primal form, various kinds of mixed or non-conforming formulations have been developed, cf. Refs. 4, 34, 14, 24, 6.

The remainder of the paper is organized as follows. We introduce IgA, the biharmonic model problem in its variational form and its discretization in Section 2. In Section 3, the multigrid method is introduced and we state sufficient conditions for its convergence. We develop the approximation error estimates needed for the convergence estimates in Section 4. The choice of the smoother, the smoothing properties and the resulting multigrid convergence results are addressed in Section 5. Finally, we provide numerical results in Section 6.

2. Model problem and its discretization

2.1. *The biharmonic model problem*

Following the usual design principles of IgA, we assume that the computational domain $\Omega \subset \mathbb{R}^d$ has a Lipschitz boundary $\partial\Omega$ and that it is parameterized by a bijective geometry function

$$\mathbf{G} : \widehat{\Omega} = (0,1)^d \to \Omega = \mathbf{G}(\widehat{\Omega}),$$

with first and second weak derivatives which are almost everywhere uniformly bounded:

$$\|\nabla^r \mathbf{G}\|_{L^\infty(\widehat{\Omega})} \leq c_1 \quad \text{and} \quad \|\nabla^r(\mathbf{G}^{-1})\|_{L^\infty(\Omega)} \leq c_2, \quad \text{for} \quad r = 1, 2, \tag{2.1}$$

for some constants c_1 and c_2. Note that this condition is satisfied if $\mathbf{G}$ is a C^1-continuous and bijective B-spline, where the inverse of the Jacobian is uniformly bounded.

After homogenization, the variational formulation of the model problem (1.1) reads as follows. Given $f \in L^2(\Omega)$ and $\beta \in \mathbb{R}$ with $\beta \geq 0$, find $u \in V := H^2(\Omega) \cap H^1_0(\Omega)$ such that

$$\beta(u, v)_{L^2(\Omega)} + (\Delta u, \Delta v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall\, v \in V. \tag{2.2}$$

Here and in what follows, $L^2(\Omega)$ and $H^r(\Omega)$ denote the standard Lebesgue and Sobolev spaces with standard inner products $(\cdot,\cdot)_{L^2(\Omega)}$, $(\cdot,\cdot)_{H^r(\Omega)}$ and norms $\|\cdot\|_{L^2(\Omega)}$, $\|\cdot\|_{H^r(\Omega)}$. $H^1_0(\Omega)$ is the standard subspace of $H^1(\Omega)$ containing the functions with vanishing trace. On V, we define the bilinear form $(\cdot,\cdot)_{\mathcal{B}}$ via

$$(u, v)_{\mathcal{B}} := (\Delta u, \Delta v)_{L^2(\Omega)} \quad \forall\, u, v \in V,$$

which is an inner product since we have the Poincaré like inequality

$$\|u\|_{H^2(\Omega)} \leq c_\Omega \|\Delta u\|_{L^2(\Omega)} = c_\Omega \|u\|_{\mathcal{B}} \quad \forall\, u \in V, \tag{2.3}$$

where c_Ω is a constant that depends only on the shape of Ω, cf. Ref. 22.

Using the substitution rule for integration and the chain rule for differentiation, (2.2) can be expressed in terms of integrals on the parameter domain $\widehat{\Omega}$. In IgA, this is usually done in order to simplify the evaluation of the integrals using quadrature rules. Besides these inner products, there are also standard inner products for the parameter domain, like $(\cdot,\cdot)_{L^2(\widehat{\Omega})}$ and $(\cdot,\cdot)_{\widehat{\mathcal{B}}}$, where the latter is given by

$$(\widehat{u}, \widehat{v})_{\widehat{\mathcal{B}}} := (\Delta \widehat{u}, \Delta \widehat{v})_{L^2(\widehat{\Omega})} \quad \forall\, \widehat{u}, \widehat{v} \in \widehat{V} := H^2(\widehat{\Omega}) \cap H^1_0(\widehat{\Omega}).$$

Also for the parameter domain $\widehat{\Omega}$, the result (2.3) holds. So, we know

$$\|\widehat{u}\|_{H^2(\widehat{\Omega})} \leq c_{\widehat{\Omega}} \|\Delta \widehat{u}\|_{L^2(\widehat{\Omega})} = c_{\widehat{\Omega}} \|\widehat{u}\|_{\widehat{\mathcal{B}}} \quad \forall\, \widehat{u} \in \widehat{V}.$$

Here and in what follows, differential operators applied to functions defined on the parameter domain, like $\widehat{u}$, refer to the coordinates on the parameter domain. We know (cf. Ref. 28) that there exist constants $\underline{c}_M$, $\overline{c}_M$, $\underline{c}_B$ and $\overline{c}_B$ only depending on the constants c_1, c_2 and the shape of Ω such that

$$\begin{aligned} \underline{c}_M\, (u, u)_{L^2(\Omega)} &\leq (\widehat{u}, \widehat{u})_{L^2(\widehat{\Omega})} \leq \overline{c}_M\, (u, u)_{L^2(\Omega)} \quad \text{and} \\ \underline{c}_B\, (u, u)_{\mathcal{B}} &\leq (\widehat{u}, \widehat{u})_{\widehat{\mathcal{B}}} \leq \overline{c}_B\, (u, u)_{\mathcal{B}} \end{aligned} \tag{2.4}$$

for all $u \in V$ with $\widehat{u} = u \circ \boldsymbol{G} \in \widehat{V}$. We define a simplified bilinear form $(\cdot,\cdot)_{\overline{\mathcal{B}}}$ as the inner product obtained by removing the cross terms from the inner product $(\cdot,\cdot)_{\widehat{\mathcal{B}}}$, that is,

$$(\widehat{u},\widehat{v})_{\bar{\mathcal{B}}} := \sum_{k=1}^{d} \left(\partial_{x_k}^2 \widehat{u}, \partial_{x_k}^2 \widehat{v}\right)_{L^2(\widehat{\Omega})} \quad \forall\, \widehat{u},\widehat{v} \in \widehat{V}.$$

Here and in what follows, $\partial_x := \frac{\partial}{\partial x}$ and $\partial_{xy} := \partial_x \partial_y$ and $\partial_x^r := \frac{\partial^r}{\partial x^r}$ denote partial derivatives. The original bilinear form and the simplified bilinear form are spectrally equivalent, which implies that also the simplified bilinear form is an inner product.

Lemma 2.1. *The inner products $(\cdot,\cdot)_{\widehat{\mathcal{B}}}$ and $(\cdot,\cdot)_{\bar{\mathcal{B}}}$ are spectrally equivalent, that is,*

$$(\widehat{u},\widehat{u})_{\bar{\mathcal{B}}} \le (\widehat{u},\widehat{u})_{\widehat{\mathcal{B}}} \le d\,(\widehat{u},\widehat{u})_{\bar{\mathcal{B}}} \quad \forall\, \widehat{u} \in \widehat{V} = H^2(\widehat{\Omega}) \cap H_0^1(\widehat{\Omega}).$$

Proof. From the results of Refs. 11 and 12, it follows that $\|\Delta \widehat{u}\|_{L^2(\widehat{\Omega})} = \|\nabla^2 \widehat{v}\|_{L^2(\widehat{\Omega})}$ for $\widehat{u},\widehat{v} \in \widehat{V}$, for a detailed proof, see, e.g., Lemma 3.3 in Ref. 30. Using this, we obtain

$$\begin{aligned}\|\widehat{u}\|_{\widehat{\mathcal{B}}}^2 &= \|\Delta \widehat{u}\|_{L^2(\widehat{\Omega})}^2 = \|\nabla^2 \widehat{u}\|_{L^2(\widehat{\Omega})}^2 \\ &= \underbrace{\sum_{k=1}^{d} \|\partial_{x_k}^2 \widehat{u}\|_{L^2(\widehat{\Omega})}^2}_{=\|\widehat{u}\|_{\bar{\mathcal{B}}}^2} + \underbrace{\sum_{k=1}^{d} \sum_{l \in \{1,\dots,d\}\setminus\{k\}} \|\partial_{x_k x_l} \widehat{u}\|_{L^2(\widehat{\Omega})}^2}_{\ge 0},\end{aligned}$$

which shows the first side of the inequality. Using the Cauchy-Schwarz inequality and $ab \le \frac{1}{2}(a^2+b^2)$, we obtain

$$\|\widehat{u}\|_{\widehat{\mathcal{B}}}^2 = \sum_{k=1}^{d}\sum_{l=1}^{d} \left(\partial_{x_k}^2 \widehat{u}, \partial_{x_l}^2 \widehat{u}\right)_{L^2(\widehat{\Omega})} \le \frac{1}{2} \sum_{k=1}^{d}\sum_{l=1}^{d} \left(\left\|\partial_{x_k}^2 \widehat{u}\right\|_{L^2(\widehat{\Omega})}^2 + \left\|\partial_{x_l}^2 \widehat{u}\right\|_{L^2(\widehat{\Omega})}^2 \right) = d \|\widehat{u}\|_{\bar{\mathcal{B}}}^2,$$

which shows second side of the inequality. □

Remark 2.1. A analogous result holds for the domain Ω, which satisfies condition (2.1). In this case, the constants also depend on the shape of Ω.

2.2. *Discretization*

We consider a discretization using tensor product B-splines in the context of IgA. We start by defining these splines on the parameter domain $\widehat{\Omega}$. Let $C^k(0,1)$ denote the space of all continuous functions mapping $(0,1) \to \mathbb{R}$ that are k times continuously differentiable and let $\mathcal{P}_p$ be the space of polynomials of degree at most p. For any sequence of grid points $\boldsymbol{\tau} := (\tau_0, \dots, \tau_{N+1})$ with

$$0 = \tau_0 < \tau_1 < \cdots < \tau_N < \tau_{N+1} = 1,$$

we define the space $S_{p,\boldsymbol{\tau}}$ of splines of degree p with maximum smoothness by

$$S_{p,\boldsymbol{\tau}} := \left\{ v \in C^{p-1}(0,1) : v|_{(\tau_j,\tau_{j+1})} \in \mathcal{P}_p,\ j = 0,1,\dots,N \right\}.$$

The size of the largest and the smallest interval are denoted by

$$h_{\boldsymbol{\tau}} := \max_{j=0,\dots,N} (\tau_{j+1} - \tau_j) \quad \text{and} \quad h_{\boldsymbol{\tau},\min} := \min_{j=0,\dots,N} (\tau_{j+1} - \tau_j),$$

respectively. For the parameter domain, we define a spline space by tensorization, which we transfer to the physical domain using the pull-back principle, thus we define for given sequences of grid points $\boldsymbol{\tau}_{\ell,1}, \dots, \boldsymbol{\tau}_{\ell,d}$ the spaces

$$\widehat{V}_\ell := \left(\bigotimes_{i=1}^d S_{p,\boldsymbol{\tau}_{\ell,i}} \right) \cap H^1_0(\widehat{\Omega}) \subset \widehat{V} \quad \text{and} \quad V_\ell := \{ f \circ \mathbf{G}^{-1} : f \in \widehat{V}_\ell \} \subset V.$$

Here and in what follows, the tensor product space $\bigotimes_{i=1}^d S_{p,\boldsymbol{\tau}_{\ell,i}}$ is the space of all linear combinations of functions of the form $v(x_1,\dots,x_d) = v_1(x_1)\cdots v_d(x_d)$ with $v_i \in S_{p,\boldsymbol{\tau}_{\ell,i}}$. The spline degree p could be different for each of the spacial directions. For notational convenience, we restrict ourselves to a uniform choice of the degree.

The corresponding minimum and maximum grid size are denoted by

$$h_\ell := \max_{i=1,\dots,d} h_{\boldsymbol{\tau}_{\ell,i}} \quad \text{and} \quad h_{\ell,\min} := \min_{i=1,\dots,d} h_{\boldsymbol{\tau}_{\ell,i},\min}.$$

For the multigrid methods we, set up a sequence nested spline spaces

$$V_0 \subset V_1 \subset \cdots \subset V_L \subset V \quad \text{with} \quad h_0 > h_1 > \cdots > h_L > 0$$

based on a sequence of nested grids.

We assume that all grids are quasi uniform, that is, there is a constant c_q such that

$$h_\ell \le c_q\, h_{\ell,\min} \quad \text{for} \quad \ell = 0, 1, \dots, L. \tag{2.5}$$

We also assume that the ratio of the grid sizes of any two consecutive grids is bounded, that is, there is a constant c_r such that

$$h_{\ell-1} \le c_r\, h_\ell \quad \text{for} \quad \ell = 1, \dots, L. \tag{2.6}$$

If the grids are obtained by uniform refinements of the coarsest grid, then this condition is naturally satisfied with $c_r = 2$.

By applying a Galerkin discretization, we obtain the following discrete problem: Find $u_\ell \in V_\ell$ such that

$$\beta(u_\ell, v_\ell)_{L^2(\Omega)} + (u_\ell, v_\ell)_{\mathcal{B}} = (f, v_\ell)_{L^2(\Omega)} \quad \forall\, v_\ell \in V_\ell. \tag{2.7}$$

By fixing a basis for the space V_ℓ, we can rewrite (2.7) in matrix-vector notation as

$$(\beta \mathcal{M}_\ell + \mathcal{B}_\ell) \underline{u}_\ell = \underline{f}_\ell, \tag{2.8}$$

where $\mathcal{B}_\ell$ is the biharmonic stiffness matrix, $\mathcal{M}_\ell$ is the mass matrix, $\underline{u}_\ell$ is the vector representation of the corresponding function u_ℓ with respect to the chosen basis and the vector $\underline{f}_\ell$ is obtained by testing the right-hand side functional $(f, \cdot)_{L^2(\Omega)}$ with the basis functions.

Notation 2.1. Throughout this paper, c is a generic positive constant that is independent of h and p, but may depend on d, the constants c_1, c_2, c_q, and c_r and the shape of Ω.

For any two square matrices $A, B \in \mathbb{R}^{n\times n}$, $A \leq B$ means that

$$\underline{x}^T A \underline{x} \leq \underline{x}^T B \underline{x} \quad \forall \underline{x} \in \mathbb{R}^n.$$

3. The multigrid solver

In this section, we present an abstract multigrid method and give a convergence theorem that is based on the analysis by Bramble et al., see Theorem 1 in Ref. 3.

3.1. *The multigrid framework*

Let us assume that we have nested spaces $V_0 \subset V_1 \subset \cdots \subset V_L \subset V$. Let $I_{\ell-1}^{\ell}$ be the matrix representation of the canonical embedding from $V_{\ell-1}$ into V_ℓ and let the restriction matrix $I_\ell^{\ell-1}$ be its transpose, this is $I_\ell^{\ell-1} := (I_{\ell-1}^{\ell})^T$.

On each grid level, $\ell = 0, \ldots, L$, we have a linear system

$$\mathcal{A}_\ell \, \underline{u}_\ell = \underline{f}_\ell,$$

which is obtained by discretizing a symmetric, bounded and coercive bilinear form $a(\cdot,\cdot)$ in the space V_ℓ using the Galerkin principle. The matrix induces a norm via $\|\underline{u}_\ell\|_{\mathcal{A}_\ell} := (\mathcal{A}_\ell \underline{u}_\ell, \underline{u}_\ell)^{1/2} = \|\mathcal{A}_\ell^{1/2} \underline{u}_\ell\|$. Here and in what follows, $(\cdot,\cdot)$ and $\|\cdot\|$ are the Euclidean scalar product and norm, respectively. In the continuous setting, the matrix can be represented by an operator

$$\mathcal{A} : V \to V' \quad \text{with} \quad \mathcal{A}u = a(u,\cdot).$$

We have $\|u_\ell\|_{\mathcal{A}} = \|\underline{u}_\ell\|_{\mathcal{A}_\ell}$ for all $u_\ell \in V_\ell$ with coefficient representation $\underline{u}_\ell$.

For the analysis, we can additionally choose symmetric positive definite matrices X_ℓ for all grid levels $\ell = 0, 1, \ldots, L$, which induce norms via $\|\underline{u}_\ell\|_{X_\ell} = (X_\ell \underline{u}_\ell, \underline{u}_\ell)^{1/2} = \|X_\ell^{1/2} \underline{u}_\ell\|$. The norm $\|u_\ell\|_{X_\ell}$ of a function $u_\ell \in V_\ell$ is interpreted as $\|\underline{u}_\ell\|_{X_\ell}$, where $\underline{u}_\ell$ is the coefficient representation of u_ℓ.

For the abstract framework, we assume to have a symmetric and positive definite matrix $\tau_\ell L_\ell^{-1}$ for every grid level $\ell = 1, \ldots, L$, representing the smoother.

Later, for the model problem, the bilinear form $a(\cdot,\cdot)$, the matrices $\mathcal{A}_\ell$, $\ell = 0, \ldots, L$ and our choice of X_ℓ will be

$$a(u,v) = \beta(u,v)_{L^2(\Omega)} + (u,v)_{\mathcal{B}}, \quad \mathcal{A}_\ell = \beta \mathcal{M}_\ell + \mathcal{B}_\ell \quad \text{and} \quad X_\ell = (\beta + h_\ell^{-4})\mathcal{M}_\ell + \mathcal{B}_\ell.$$

As smoothers, we will choose a subspace corrected mass smoother, a symmetric Gauss-Seidel smoother and a hybrid smoother in Section 5. Based on these choices, the overall algorithm reads as follows.

Algorithm 3.1. One multigrid cycle, applied to some iterate $\underline{u}_\ell^{(0)}$ and a right-hand side $\underline{f}_\ell$ consists of the following steps:

- Apply ν_ℓ pre-smoothing steps, i.e., compute
$$\underline{u}_\ell^{(i)} = \underline{u}_\ell^{(i-1)} + \tau_\ell L_\ell^{-1}(\underline{f}_\ell - \mathcal{A}_\ell \underline{u}_\ell^{(i-1)}) \quad \text{for} \quad i = 1, \ldots, \nu_\ell. \tag{3.1}$$
- Apply recursive coarse-grid correction, i.e., apply the following steps. Compute the residual and restrict it to the next coarser grid level:
$$\underline{r}_{\ell-1} = I_\ell^{\ell-1}(\underline{f}_\ell - \mathcal{A}_\ell \underline{u}_\ell^{(\nu_\ell)}).$$
If $\ell - 1 = 0$, compute the update $\underline{q}_0 := A_0^{-1}\underline{r}_0$ using a direct solver. Otherwise, compute the update $\underline{q}_{\ell-1}$ by applying the algorithm r ($r \in \mathbb{N} := \{1, 2, \ldots\}$) times recursively to the right-hand side $\underline{r}_{\ell-1}$ and a zero vector as initial guess. Then set
$$\underline{u}_\ell^{(\nu_\ell+1)} = \underline{u}_\ell^{(\nu_\ell)} + I_{\ell-1}^\ell \underline{q}_{\ell-1}.$$
- Apply ν_ℓ post-smoothing steps, i.e., compute $\underline{u}_\ell^{(i)}$ using (3.1) for $i = \nu_\ell + 2, \ldots, 2\nu_\ell + 1$ to obtain the next iterate $\underline{u}_\ell^{(2\nu_\ell+1)}$.

This abstract algorithm coincides with the algorithm presented in Ref. 3. Since each multigrid cycle is linear, its application can be expressed by the matrix B_ℓ^s, which is recursively given by $B_0^s := \mathcal{A}_0^{-1}$ and
$$B_\ell^s := \big(I - (I - \tau_\ell L_\ell^{-1}\mathcal{A}_\ell)^{\nu_\ell}(I - I_{\ell-1}^\ell B_{\ell-1}^s I_\ell^{\ell-1}\mathcal{A}_\ell)^r (I - \tau_\ell L_\ell^{-1}\mathcal{A}_\ell)^{\nu_\ell}\big)\mathcal{A}_\ell^{-1}, \quad \ell = 1, \ldots, L.$$
The iteration matrix corresponding to one multigrid cycle is given by
$$I - B_\ell^s \mathcal{A}_\ell = (I - \tau_\ell L_\ell^{-1}\mathcal{A}_\ell)^{\nu_\ell}(I - I_{\ell-1}^\ell B_{\ell-1}^s I_\ell^{\ell-1}\mathcal{A}_\ell)^r (I - \tau_\ell L_\ell^{-1}\mathcal{A}_\ell)^{\nu_\ell}, \quad \ell = 1, \ldots, L.$$

Remark 3.1. The coarse-grid correction is realized by applying r iterations of the algorithm on the next coarser grid level. Thus, $r = 1$ corresponds to the V-cycle and $r = 2$ corresponds to the W-cycle.

3.2. *Abstract convergence framework*

The assumptions used to show convergence can be split into two groups: *approximation properties* and *smoother properties.*

Theorem 3.1. *Let λ_ℓ be the largest eigenvalue of $X_\ell^{-1}\mathcal{A}_\ell$. Assume that the following estimates hold:*

- Approximation properties. *There are constants C_1 and C_2, independent of ℓ, and linear operators $Q_\ell : V_L \to V_\ell$ for $\ell = 0, 1, \ldots, L$ with $Q_L = I$ such that*
$$\|(Q_\ell - Q_{\ell-1})u_L\|_{X_\ell}^2 \le C_1 \lambda_\ell^{-1}(u_L, u_L)_{\mathcal{A}} \qquad \textit{for} \quad \ell = 1, \ldots, L, \tag{3.2}$$
$$(Q_\ell u_L, Q_\ell u_L)_{\mathcal{A}} \le C_2 (u_L, u_L)_{\mathcal{A}} \qquad \textit{for} \quad \ell = 0, \ldots, L-1, \tag{3.3}$$
for all $u_L \in V_L$.

- Smother properties. *We assume there exist a constant C_S independent of ℓ such that*

$$\frac{\|\underline{u}_\ell\|_{X_\ell}^2}{\lambda_\ell} \leq C_S(\tau_\ell L_\ell^{-1} X_\ell \underline{u}_\ell, \underline{u}_\ell)_{X_\ell} \quad \forall\, \underline{u}_\ell \in \mathbb{R}^{\dim V_\ell} \tag{3.4}$$

  and

$$(\tau_\ell L_\ell^{-1} \mathcal{A}_\ell \underline{u}_\ell, \underline{u}_\ell)_{\mathcal{A}_\ell} \leq (\underline{u}_\ell, \underline{u}_\ell)_{\mathcal{A}_\ell} \quad \forall\, \underline{u}_\ell \in \mathbb{R}^{\dim V_\ell} \tag{3.5}$$

  holds for $\ell = 1, \dots, L$.

Then, the estimate

$$((I - B_L^s \mathcal{A}_L)\underline{u}_L, \underline{u}_L)_{\mathcal{A}_L} \leq \left(1 - \frac{1}{CL}\right)(\underline{u}_L, \underline{u}_L)_{\mathcal{A}_L},$$

holds for all $\underline{u}_L \in \mathbb{R}^{\dim V_L}$, where $C = [1 + C_2^{1/2} + (C_S C_1)^{1/2}]^2$.

For a proof, see Theorem 1 in Ref. 3.

Remark 3.2. Condition (3.4) is only required for functions u_ℓ in the range of $Q_\ell - Q_{\ell-1}$. However, since we do not exploit this, we have stated the stronger condition.

Now, we provide conditions that guarantee (3.4) and (3.5), which fit our needs better than the original conditions.

Lemma 3.1. *If there exists a constant C_S, independent of ℓ, which satisfies*

$$(\mathcal{A}_\ell \underline{u}_\ell, \underline{u}_\ell) \leq \frac{1}{\tau_\ell}(L_\ell \underline{u}_\ell, \underline{u}_\ell) \leq \lambda_\ell C_S (X_\ell \underline{u}_\ell, \underline{u}_\ell) \quad \forall\, \underline{u}_\ell \in \mathbb{R}^{\dim V_\ell} \tag{3.6}$$

for each $\ell = 1, \dots, L$. Then, the assumptions (3.4) and (3.5) hold true for the same C_S.

Proof. We start by showing that the first inequality implies (3.5), i.e., the smoothing operator $I - \tau_\ell L_\ell^{-1} \mathcal{A}_\ell$ is nonnegative in the scalar product induced by $\mathcal{A}_\ell$, i.e., $(\mathcal{A}_\ell (I - \tau_\ell L_\ell^{-1} \mathcal{A}_\ell)\underline{u}_\ell, \underline{u}_\ell) \geq 0$. Let $\underline{w}_\ell \in \mathbb{R}^{\dim V_\ell}$ be an arbitrary vector. Using the Cauchy-Schwarz inequality and the first inequality in (3.6), we obtain

$$\begin{aligned}
\tau_\ell (L_\ell^{-1} \underline{w}_\ell, \underline{w}_\ell) &= \tau_\ell (\mathcal{A}_\ell^{1/2} L_\ell^{-1} \underline{w}_\ell, \mathcal{A}_\ell^{-1/2} \underline{w}_\ell) \\
&\leq \tau_\ell (\mathcal{A}_\ell L_\ell^{-1} \underline{w}_\ell, L_\ell^{-1} \underline{w}_\ell)^{1/2} (\mathcal{A}_\ell^{-1} \underline{w}_\ell, \underline{w}_\ell)^{1/2} \\
&\leq \tau_\ell^{1/2} (L_\ell^{-1} \underline{w}_\ell, \underline{w}_\ell)^{1/2} (\mathcal{A}_\ell^{-1} \underline{w}_\ell, \underline{w}_\ell)^{1/2}
\end{aligned}$$

It follows that

$$\tau_\ell (L_\ell^{-1} \underline{w}_\ell, \underline{w}_\ell) \leq (\mathcal{A}_\ell^{-1} \underline{w}_\ell, \underline{w}_\ell) \quad \forall\, \underline{w}_\ell \in \mathbb{R}^{\dim V_\ell}.$$

By substituting $\underline{w}_\ell$ with $\mathcal{A}_\ell \underline{u}_\ell$, we get (3.5). Next, we use the Cauchy-Schwarz inequality and the second inequality in (3.6) to show (3.4). Let $\underline{w}_\ell \in \mathbb{R}^{\dim V_\ell}$, we have

$$\begin{aligned}(X_\ell^{-1}\underline{w}_\ell, \underline{w}_\ell) &= (L_\ell^{1/2} X_\ell^{-1}\underline{w}_\ell, L_\ell^{-1/2}\underline{w}_\ell) \le (L_\ell X_\ell^{-1}\underline{w}_\ell, X_\ell^{-1}\underline{w}_\ell)^{1/2}(L_\ell^{-1}\underline{w}_\ell, \underline{w}_\ell)^{1/2}\\ &\le \tau_\ell^{1/2}\lambda_\ell^{1/2}C_S^{1/2}(X_\ell^{-1}\underline{w}_\ell, \underline{w}_\ell)^{1/2}(L_\ell^{-1}\underline{w}_\ell, \underline{w}_\ell)^{1/2}.\end{aligned}$$

By squaring the inequality, we get

$$(X_\ell^{-1}\underline{w}_\ell, \underline{w}_\ell) = \tau_\ell \lambda_\ell C_S (L_\ell^{-1}\underline{w}_\ell, \underline{w}_\ell) \quad \forall\, \underline{w}_\ell \in \mathbb{R}^{\dim V_\ell}.$$

By substituting $\underline{w}_\ell$ with $X_\ell \underline{u}_\ell$, we get (3.4). □

4. Approximation error estimates

In this section, we prove some approximation error estimates and provide a projector which will be used to prove (3.2) and (3.3).

4.1. *Error and stability estimates for the univariate case*

We start by introducing a periodic spline space. For any given sequence of grid points $\boldsymbol{\tau} = (0, \tau_1, \dots, \tau_N, 1)$, we define

$$\boldsymbol{\tau}^{per} := (-1, -\tau_N, \cdots, -\tau_1, 0, \tau_1, \cdots, \tau_N, 1).$$

For each $p \in \mathbb{N}$, we define the periodic spline space

$$S_{p,\boldsymbol{\tau}}^{per} := \left\{ v \in S_{p,\boldsymbol{\tau}^{per}} \,:\, \partial^l v\,(-1) = \partial^l v\,(1) \quad \forall\, l \in \mathbb{N}_0 \text{ with } l < p \right\}$$

and a spline space with vanishing even derivatives on the boundary

$$S_{p,\boldsymbol{\tau}}^{0} := \left\{ v \in S_{p,\boldsymbol{\tau}} \,:\, \partial^{2l} v\,(0) = \partial^{2l} v\,(1) = 0 \quad \forall\, l \in \mathbb{N}_0 \text{ with } 2l < p \right\}. \tag{4.1}$$

We also define the periodic Sobolev space

$$H^q_{per}(-1,1) := \left\{ v \in H^q(-1,1) \,:\, \partial^l v\,(-1) = \partial^l v\,(1)\,, \quad \forall\, l \in \mathbb{N}_0 \text{ with } l < q \right\}$$

for each $q \in \mathbb{N}$. Let $\Pi^{per}_{p,\boldsymbol{\tau}} : H^2_{per}(-1,1) \to S^{per}_{p,\boldsymbol{\tau}}$ be the H^2-orthogonal projector satisfying

$$\begin{aligned}\left(\partial^2 \Pi^{per}_{p,\boldsymbol{\tau}} u, \partial^2 v\right)_{L^2(-1,1)} &= \left(\partial^2 u, \partial^2 v\right)_{L^2(-1,1)} \quad \forall\, v \in S^{per}_{p,\boldsymbol{\tau}},\\ \left(\Pi^{per}_{p,\boldsymbol{\tau}} u, 1\right)_{L^2(-1,1)} &= (u,1)_{L^2(-1,1)}\,.\end{aligned} \tag{4.2}$$

We use the following approximation error estimate for spline spaces which does not require uniform knot spans.

Theorem 4.1. *For any $p \ge 3$, we have*

$$\|\partial^2(u - \Pi^{per}_{p,\boldsymbol{\tau}} u)\|_{L^2(-1,1)} \le \frac{h_{\boldsymbol{\tau}}^2}{\pi^2} \|\partial^4 u\|_{L^2(-1,1)} \quad \forall\, u \in H^4_{per}(-1,1).$$

For a proof, see Theorem 4 in Ref. 25.

Using the H^2–H^4 result above and an Aubin-Nitsche duality trick, we obtain the following L^2–H^2 result.

Theorem 4.2. *For any $p \geq 3$, we have*

$$\|u - \Pi^{per}_{p,\boldsymbol{\tau}} u\|_{L^2(-1,1)} \leq \frac{h^2_{\boldsymbol{\tau}}}{\pi^2} \|\partial^2 u\|_{L^2(-1,1)} \quad \forall\, u \in H^2_{per}(-1,1).$$

Proof. Let $u \in H^2_{per}(-1,1)$ be arbitrary but fixed. Let $w \in H^4(-1,1) \cap H^3_{per}(-1,1)$ be such that $\partial^4 w = u - \Pi^{per}_{p,\boldsymbol{\tau}} u$. Note that (4.2) gives $0 = (u - \Pi^{per}_{p,\boldsymbol{\tau}} u, 1)_{L^2(-1,1)} = (\partial^4 w, 1)_{L^2(-1,1)} = \partial^3 w(1) - \partial^3 w(-1)$. So, we know that $w \in H^4_{per}(-1,1)$.

Using integration by parts (which does not introduce boundary terms since $u - \Pi^{per}_{p,\boldsymbol{\tau}} u \in H^2_{per}(-1,1)$ and $w \in H^4_{per}(-1,1)$) and using Theorem 4.1, we obtain

$$\begin{aligned}
\|u - \Pi^{per}_{p,\boldsymbol{\tau}} u\|^2_{L^2} &= \frac{(u - \Pi^{per}_{p,\boldsymbol{\tau}} u, u - \Pi^{per}_{p,\boldsymbol{\tau}} u)_{L^2}}{\|u - \Pi^{per}_{p,\boldsymbol{\tau}} u\|_{L^2}} = \frac{(u - \Pi^{per}_{p,\boldsymbol{\tau}} u, \partial^4 w)_{L^2}}{\|\partial^4 w\|_{L^2}} \\
&= \frac{(\partial^2(u - \Pi^{per}_{p,\boldsymbol{\tau}} u), \partial^2 w)_{L^2}}{\|\partial^4 w\|_{L^2}} \leq \frac{h^2_{\boldsymbol{\tau}}}{\pi^2} \frac{(\partial^2(u - \Pi^{per}_{p,\boldsymbol{\tau}} u), \partial^2 w)_{L^2}}{\|\partial^2(w - \Pi^{per}_{p,\boldsymbol{\tau}} w)\|_{L^2}}.
\end{aligned}$$

From the definition of $\Pi^{per}_{p,\boldsymbol{\tau}}$, see (4.2), we have $(\partial^2(u - \Pi^{per}_{p,\boldsymbol{\tau}} u), \partial^2 \Pi^{per}_{p,\boldsymbol{\tau}} w)_{L^2} = 0$. This, together with the Cauchy-Schwarz inequality and the H^2-stability of $\Pi^{per}_{p,\boldsymbol{\tau}}$, gives

$$\begin{aligned}
\|u - \Pi^{per}_{p,\boldsymbol{\tau}} u\|^2_{L^2} &\leq \frac{h^2_{\boldsymbol{\tau}}}{\pi^2} \frac{(\partial^2(u - \Pi^{per}_{p,\boldsymbol{\tau}} u), \partial^2(w - \Pi^{per}_{p,\boldsymbol{\tau}} w))_{L^2}}{\|\partial^2(w - \Pi^{per}_{p,\boldsymbol{\tau}} w)\|_{L^2}} \\
&\leq \frac{h^2_{\boldsymbol{\tau}}}{\pi^2} \|\partial^2(u - \Pi^{per}_{p,\boldsymbol{\tau}} u)\|^2_{L^2} \leq \frac{h^2_{\boldsymbol{\tau}}}{\pi^2} \|\partial^2 u\|^2_{L^2},
\end{aligned}$$

which completes the proof. □

Let $\Pi^0_{p,\boldsymbol{\tau}} : H^2(0,1) \cap H^1_0(0,1) \to S^0_{p,\boldsymbol{\tau}}$ be the H^2-orthogonal projector satisfying

$$\left(\partial^2 \Pi^0_{p,\boldsymbol{\tau}} u, \partial^2 v\right)_{L^2(0,1)} = \left(\partial^2 u, \partial^2 v\right)_{L^2(0,1)} \quad \forall\, v \in S^0_{p,\boldsymbol{\tau}}.$$

Theorem 4.3. *For any $p \geq 3$, we have*

$$\|u - \Pi^0_{p,\boldsymbol{\tau}} u\|_{L^2(0,1)} \leq \frac{h^2_{\boldsymbol{\tau}}}{\pi^2} \|\partial^2 u\|_{L^2(0,1)} \quad \forall\, u \in H^2(0,1) \cap H^1_0(0,1).$$

Proof. Let $u \in H^2(0,1) \cap H^1_0(0,1)$ be arbitrary but fixed. Define w on $(-1,1)$ to be

$$w(x) := \operatorname{sign}(x)\, u(|x|).$$

Observe that we obtain $w \in H^2_{per}(-1,1)$. From Theorem 4.2, we have

$$\|(I - \Pi^{per}_{p,\boldsymbol{\tau}}) w\|_{L^2(-1,1)} \leq \frac{h^2_{\boldsymbol{\tau}}}{\pi^2} \|\partial^2 w\|_{L^2(-1,1)}.$$

Observe that $\|\partial^2 w\|_{L^2(-1,1)} = 2^{1/2} \|\partial^2 u\|_{L^2(0,1)}$. Define $w_{\boldsymbol{\tau}} := \Pi^{per}_{p,\boldsymbol{\tau}} w$ and let $u_{\boldsymbol{\tau}}$ be the restriction of $w_{\boldsymbol{\tau}}$ to $(0,1)$. Observe that $w_{\boldsymbol{\tau}}$ is anti-symmetric, which implies that $u_{\boldsymbol{\tau}} \in S^0_{p,\boldsymbol{\tau}}$. It follows that $\|w - w_{\boldsymbol{\tau}}\|_{L^2(-1,1)} = 2^{1/2}\|u - u_{\boldsymbol{\tau}}\|_{L^2(0,1)}$. Using this, we obtain

$$\|u - u_{\boldsymbol{\tau}}\|_{L^2(0,1)} \le \frac{h_{\boldsymbol{\tau}}^2}{\pi^2}\|\partial^2 u\|_{L^2(0,1)}.$$

It remains to show that $u_{\boldsymbol{\tau}}$ coincides with $\Pi^{per}_{p,\boldsymbol{\tau}} u$, i.e., to show that $u - u_{\boldsymbol{\tau}}$ is H^2-orthogonal to $S^0_{p,\boldsymbol{\tau}}$. By definition, this means that we have to show

$$(\partial^2(u - u_{\boldsymbol{\tau}}), \partial^2 \tilde{u}_{\boldsymbol{\tau}})_{L^2(0,1)} = 0 \quad \forall\, \tilde{u}_{\boldsymbol{\tau}} \in S^0_{p,\boldsymbol{\tau}}.$$

Let $\tilde{w}_{\boldsymbol{\tau}} \in S^{per}_{p,\boldsymbol{\tau}}$ be $\tilde{w}_{\boldsymbol{\tau}} := \mathrm{sign}(x)\, \tilde{u}_{\boldsymbol{\tau}}(|x|)$ and observe that $2(\partial^2(u - u_{\boldsymbol{\tau}}), \partial^2 \tilde{u}_{\boldsymbol{\tau}})_{L^2(0,1)} = (\partial^2(w - w_{\boldsymbol{\tau}}), \partial^2 \tilde{w}_{\boldsymbol{\tau}})_{L^2(0,1)}$, since u, $u_{\boldsymbol{\tau}}$ and $\tilde{u}_{\boldsymbol{\tau}}$ are restrictions of w, $w_{\boldsymbol{\tau}}$ and $\tilde{w}_{\boldsymbol{\tau}}$, respectively. Furthermore, $(\partial^2(w - w_{\boldsymbol{\tau}}), \partial^2 \tilde{w}_{\boldsymbol{\tau}})_{L^2(-1,1)} = 0$ by construction, since $w_{\boldsymbol{\tau}} := \Pi^{per}_{p,\boldsymbol{\tau}} w$, which completes the proof. □

Let $Q^0_{p,\boldsymbol{\tau}} : H^2(0,1) \cap H^1_0(0,1) \to S^0_{p,\boldsymbol{\tau}}$ be the L^2-orthogonal projector satisfying

$$\left(Q^0_{p,\boldsymbol{\tau}} u, v\right)_{L^2(0,1)} = (u, v)_{L^2(0,1)} \quad \forall\, v \in S^0_{p,\boldsymbol{\tau}}.$$

Since the L^2-orthogonal projector minimizes the error in the L^2-norm, Theorem 4.3 immediately implies the following statement.

Theorem 4.4. *For any $p \ge 3$, we have*

$$\|u - Q^0_{p,\boldsymbol{\tau}} u\|_{L^2(0,1)} \le \frac{h_{\boldsymbol{\tau}}^2}{\pi^2}\|\partial^2 u\|_{L^2(0,1)} \quad \forall\, u \in H^2(0,1) \cap H^1_0(0,1).$$

Next, we show the stability of $Q^0_{p,\boldsymbol{\tau}}$ with respect to the H^2-seminorm. Such a proof is possible since the space $S^0_{p,\boldsymbol{\tau}}$ satisfies the following p-robust inverse inequality, while the space $S_{p,\boldsymbol{\tau}} \cap H^1_0(0,1)$ does not satisfy such an inverse inequality, cf. Ref. 32.

Theorem 4.5. *Let $p \in \mathbb{N}$ with $p \ge 2$. We have*

$$\|\partial^2 u_{\boldsymbol{\tau}}\|_{L^2(0,1)} \le 12 h^{-2}_{\boldsymbol{\tau},\min}\|u_{\boldsymbol{\tau}}\|_{L^2(0,1)} \quad \forall u_{\boldsymbol{\tau}} \in S^0_{p,\boldsymbol{\tau}}.$$

A proof can be found in Theorem 12 in Ref. 29. It is done be induction, starting from $p = 2$. For the induction step, one uses integration by parts. The boundary terms vanish due to the boundary conditions baked into the definition of $S^0_{p,\boldsymbol{\tau}}$.

Theorem 4.6. *Let $p \in \mathbb{N}$ with $p \ge 3$. Then there exists a constant $c > 0$ such that*

$$\|\partial^2(Q^0_{p,\boldsymbol{\tau}} u)\|^2_{L^2(0,1)} \le c\|\partial^2 u\|^2_{L^2(0,1)} \quad \forall u \in H^2(0,1) \cap H^1_0(0,1).$$

Proof. The proof is analogous to that of Theorem 14 in Ref. 29, however it is given here for completeness. Using the triangle inequality and the inverse inequality in Theorem 4.5, we obtain

$$\begin{aligned}
\|\partial^2 Q^0_{p,\boldsymbol{\tau}} u\|^2_{L^2} &\le 2\|\partial^2 \Pi^0_{p,\boldsymbol{\tau}} u\|^2_{L^2} + 2\|\partial^2(Q^0_{p,\boldsymbol{\tau}} u - \Pi^0_{p,\boldsymbol{\tau}} u)\|^2_{L^2} \\
&\le 2\|\partial^2 \Pi^0_{p,\boldsymbol{\tau}} u\|^2_{L^2} + c h^{-2}_{\boldsymbol{\tau},\min}\|Q^0_{p,\boldsymbol{\tau}} u - \Pi^0_{p,\boldsymbol{\tau}} u\|^2_{L^2} \\
&\le 2\|\partial^2 \Pi^0_{p,\boldsymbol{\tau}} u\|^2_{L^2} + c h^{-2}_{\boldsymbol{\tau},\min}\|u - \Pi^0_{p,\boldsymbol{\tau}} u\|^2_{L^2} + c h^{-2}_{\boldsymbol{\tau},\min}\|u - Q^0_{p,\boldsymbol{\tau}} u\|^2_{L^2}.
\end{aligned}$$

The Theorems 4.3 and 4.4 and assumption (2.5) give the desired result. □

4.2. *Proof of the approximation properties*

In this subsection, we consider the discretization framework from Section 2. We choose

$$X_\ell := \mathcal{B}_\ell + (\beta + h_\ell^{-4})\mathcal{M}_\ell,$$

which corresponds to the norm $\|\cdot\|_{X_\ell}$ that satisfies

$$\|u\|_{X_\ell}^2 = \|u\|_{\mathcal{B}}^2 + (\beta + h_\ell^{-4})\|u\|_{L^2(\Omega)}^2 \quad \forall\, u \in V.$$

Now, we give a bound for the eigenvalues of $X_\ell^{-1}\mathcal{A}_\ell$.

Lemma 4.1. *Let λ_ℓ with $\ell \ge 1$ be the largest eigenvalue of $X_\ell^{-1}\mathcal{A}_\ell$. For $p \ge 3$, we have $\lambda_\ell \in (\frac{1}{1+c}, 1)$ for some positive constant c.*

We give a proof of this Lemma below.

Next, we prove (3.2) and (3.3). This requires that we choose the projectors $\mathbf{Q}_{p,\ell}^0$, which have to map into the space V_ℓ. We first define a projector that maps from $\widehat{V}$ into $\widehat{V}_\ell$ by tensorization of the univariate projectors:

$$\widehat{\mathbf{Q}}_{p,\ell}^0 := Q_{p,\boldsymbol{\tau}_{\ell,1}}^0 \otimes \cdots \otimes Q_{p,\boldsymbol{\tau}_{\ell,d}}^0,$$

where the tensor product is to be understood as in Section 3.2 in Ref. 31. The next two theorems follow from Theorems 4.4 and 4.6 by standard arguments.

Theorem 4.7. *Let $p \in \mathbb{N}$ with $p \ge 3$. Then there exists a constant c such that*

$$\|(I - \widehat{\mathbf{Q}}_{p,\ell}^0)\widehat{u}\|_{L^2(\widehat{\Omega})} \le c h_\ell^2 \|\widehat{u}\|_{\widehat{\mathcal{B}}} \quad \forall\, \widehat{u} \in H^2(\widehat{\Omega}) \cap H_0^1(\widehat{\Omega}).$$

Proof. The result follows from the definition of $\widehat{\mathbf{Q}}_{p,\ell}^0$, the L^2-stability of the L^2-projectors, triangle inequality and Theorem 4.4. □

Now, we can give a proof of Lemma 4.1.

Proof. (of Lemma 4.1). Since $\mathcal{M}_\ell$ is symmetric positive definite and $h_\ell^{-4} > 0$, we have $\mathcal{A}_\ell < X_\ell$, which implies $\lambda_\ell < 1$.

For the lower bound, we use $V_{\ell-1} \subsetneqq V_\ell$, which implies that there is some nonzero $w_\ell \in V_\ell$ that is L_2-orthogonal to $V_{\ell-1}$, that is $(w_\ell, u_{\ell-1})_{L^2(\Omega)} = 0$ for all $u_{\ell-1} \in V_{\ell-1}$. By combining Theorem 4.7 and (2.4), we obtain

$$\|w_\ell\|_{L^2(\Omega)} = \sup_{u_{\ell-1} \in V_{\ell-1}} \|w_\ell - u_{\ell-1}\|_{L^2(\Omega)} \le c\, h_{\ell-1}^2 \|w_\ell\|_{\mathcal{B}}.$$

In matrix-vector notation, this reads as

$$\underline{w}_\ell^T \mathcal{M}_\ell \underline{w}_\ell \le c\, h_{\ell-1}^4\, \underline{w}_\ell^T \mathcal{B}_\ell \underline{w}_\ell.$$

Using (2.6), we know that there is a constant $c > 0$ such that

$$\underline{w}_\ell^\top X_\ell \underline{w}_\ell = \underline{w}_\ell^\top \mathcal{A}_\ell \underline{w}_\ell + h_\ell^{-4} \underline{w}_\ell^\top \mathcal{M}_\ell \underline{w}_\ell < (1+c)\underline{w}_\ell^\top \mathcal{A}_\ell \underline{w}_\ell,$$

which shows $\lambda_\ell > 1/(1+c)$. □

Theorem 4.8. *Let $p \in \mathbb{N}$ with $p \geq 3$. Then there exists a constant $c > 0$ such that*

$$\|\widehat{\boldsymbol{Q}}^0_{p,\ell}\widehat{u}\|^2_{\widehat{\mathcal{B}}} \leq c\|\widehat{u}\|^2_{\widehat{\mathcal{B}}} \quad \forall \widehat{u} \in H^2(\widehat{\Omega}) \cap H^1_0(\widehat{\Omega}).$$

Proof. The proof is based on the definition that $\|w\|^2_{\widehat{\mathcal{B}}} = \sum_{k=1}^d \|\partial^2_{x_k} w\|_{L^2(\widehat{\Omega})}$ and the boundedness of the L^2-orthogonal projectors $Q^0_{p,\boldsymbol{\tau}_{\ell,k}}$ in the L^2-norm and in the H^2-norm (Theorem 4.6). □

The projectors $\mathbf{Q}^0_{p,\ell}$ are now defined via the pull-back principle, such that

$$\mathbf{Q}^0_{p,\ell} u := (\widehat{\mathbf{Q}}^0_{p,\ell}(u \circ \boldsymbol{G})) \circ \boldsymbol{G}^{-1} \quad \forall\, u \in V. \tag{4.3}$$

Note that, by construction, $\mathbf{Q}^0_{p,\ell}$ maps into a subspace of V_ℓ, where all even outer normal derivatives on the boundary vanish.

Theorem 4.9. *Let $d \in \mathbb{N}$ and $p \in \mathbb{N}$ with $p \geq 3$. For each level $\ell = 0, 1, \ldots, L-1$, let $\mathbf{Q}^0_{p,\ell} : H^2(\Omega) \cap H^1_0(\Omega) \to V_\ell$ be the projectors defined in (4.3). There exist constants C_1 and C_2 such that*

$$\|(\mathbf{Q}^0_{p,\ell} - \mathbf{Q}^0_{p,\ell-1})u_L\|^2_{X_\ell} \leq C_1 \lambda_\ell^{-1}(u_L, u_L)_{\mathcal{A}} \qquad \textit{for} \quad \ell = 1, \ldots, L, \tag{4.4}$$

$$(\mathbf{Q}^0_{p,\ell}\, u_L, \mathbf{Q}^0_{p,\ell}\, u_L)_{\mathcal{A}} \leq C_2 (u_L, u_L)_{\mathcal{A}} \qquad \textit{for} \quad \ell = 0, \ldots, L-1, \tag{4.5}$$

for all $u_L \in V_L$.

Proof. Let $u_L \in V_L$ be arbitrary but fixed and let $\widehat{u}_L := u_L \circ \boldsymbol{G} \in \widehat{V}_L$. Using (2.4), Lemma 2.1 and Theorem 4.8 and the L^2-stability of $\widehat{\mathbf{Q}}^0_{p,\ell}$, we obtain

$$\begin{aligned}(\mathbf{Q}^0_{p,\ell}\, u_L, \mathbf{Q}^0_{p,\ell}\, u_L)_{\mathcal{A}} &\leq c(\widehat{\mathbf{Q}}^0_{p,\ell}\, \widehat{u}_L, \widehat{\mathbf{Q}}^0_{p,\ell}\, \widehat{u}_L)_{\widehat{\mathcal{A}}} = c\beta\|\widehat{\mathbf{Q}}^0_{p,\ell}\widehat{u}_L\|^2_{L^2(\widehat{\Omega})} + c\|\widehat{\mathbf{Q}}^0_{p,\ell}\widehat{u}_L\|^2_{\widehat{\mathcal{B}}} \\ &\leq c\beta\|\widehat{u}_L\|^2_{L^2(\widehat{\Omega})} + c\|\widehat{u}_L\|^2_{\widehat{\mathcal{B}}} \leq c(\widehat{u}_L, \widehat{u}_L)_{\widehat{\mathcal{A}}} \leq C_2(u_L, u_L)_{\mathcal{A}},\end{aligned}$$

which shows (4.5). Next we prove the auxiliary result

$$\|(I - \mathbf{Q}^0_{p,\ell-1})u_L\|^2_{X_\ell} \leq c\lambda_\ell^{-1}(u_L, u_L)_{\mathcal{A}} \quad \text{for} \quad \ell = 1, \ldots, L. \tag{4.6}$$

Using (2.4), (2.1), Theorem 4.8, Theorem 4.7 and the L^2-stability of $\mathbf{Q}^0_{p,\ell-1}$, we get

$$\begin{aligned}\|(I - \mathbf{Q}^0_{p,\ell-1})u_L\|^2_{X_\ell} &= \|(I - \mathbf{Q}^0_{p,\ell-1})u_L\|^2_{\mathcal{B}} + (\beta + h_\ell^{-4})\|(I - \mathbf{Q}^0_{p,\ell-1})u_L\|^2_{L^2(\Omega)} \\ &\leq c\|(I - \widehat{\mathbf{Q}}^0_{p,\ell-1})\widehat{u}_L\|^2_{\widehat{\mathcal{B}}} + c(\beta + h_\ell^{-4})\|(I - \widehat{\mathbf{Q}}^0_{p,\ell-1})\widehat{u}_L\|^2_{L^2(\widehat{\Omega})} \\ &\leq c\|\widehat{u}_L\|^2_{\widehat{\mathcal{B}}} + ch_\ell^{-4}h_{\ell-1}^4\|\widehat{u}_L\|^2_{\widehat{\mathcal{B}}} + c\beta\|\widehat{u}_L\|^2_{L^2(\widehat{\Omega})} \\ &\leq c(1 + h_\ell^{-4}h_{\ell-1}^4)\|u_L\|^2_{\mathcal{B}} + c\beta\|u_L\|^2_{L^2(\Omega)}.\end{aligned}$$

We use assumption (2.6) and Lemma 4.1 to get (4.6). To complete the proof, we use the fact that $\mathbf{Q}^0_{p,\ell-1}\mathbf{Q}^0_{p,\ell} = \mathbf{Q}^0_{p,\ell-1}$, (4.6) and (4.5), to obtain

$$\begin{aligned}\|(\mathbf{Q}^0_{p,\ell} - \mathbf{Q}^0_{p,\ell-1})u_L\|^2_{X_\ell} &= \|(I - \mathbf{Q}^0_{p,\ell-1})\mathbf{Q}^0_{p,\ell}u_L\|^2_{X_\ell} \leq c\lambda_\ell^{-1}(\mathbf{Q}^0_{p,\ell}u_L, \mathbf{Q}^0_{p,\ell}u_L)_{\mathcal{A}} \\ &\leq C_1\lambda_\ell^{-1}(u_L, u_L)_{\mathcal{A}}.\end{aligned}$$

This shows (4.4) and finishes the proof. □

Remark 4.1. In Lemma 9.2 in Ref. 28, a similar result to Theorem 4.7 is shown. There, the $\mathcal{B}_\ell$-orthogonal projector is considered. That proof only holds true for uniform grids. By using an L^2-orthogonal projector, we avoid these difficulties. Since the convergence theory by Hackbusch, cf. Ref. 13, requires the error estimates for the $\mathcal{B}_\ell$-orthogonal projector, this motivated us to use the convergence theory by Bramble, cf. Ref. 2, where this is not the case.

5. The smoothers and the overall convergence results

5.1. *Subspace corrected mass smoother*

We consider the subspace corrected mass smoother, which was originally proposed in Ref. 15 for a second order problem and was one of the first smoothers to produce a multigrid method for IgA which is robust in both the grid size and the spline degree. In Refs. 29, 28, this smoother was extended to biharmonic problems. The smoother is based on the inverse inequality in Theorem 4.5, which is independent of the spline degree.

First, we introduce a splitting for the one-dimensional case as follows:

$$S_{p,\boldsymbol{\tau}} \cap H_0^1(0,1) = S^0_{p,\boldsymbol{\tau}} \oplus S^1_{p,\boldsymbol{\tau}},$$

where $S^0_{p,\boldsymbol{\tau}}$ is as defined in (4.1) and $S^1_{p,\boldsymbol{\tau}}$ is its L^2-orthogonal complement in $S_{p,\boldsymbol{\tau}} \cap H_0^1(0,1)$. For each of these spaces, we define the corresponding L^2-orthogonal projection

$$\begin{aligned} Q^0_{p,\boldsymbol{\tau}} &: H^2(0,1) \cap H_0^1(0,1) \to S^0_{p,\boldsymbol{\tau}}, \\ Q^1_{p,\boldsymbol{\tau}} &: H^2(0,1) \cap H_0^1(0,1) \to S^1_{p,\boldsymbol{\tau}}. \end{aligned}$$

The next step, is to extend the splitting to the multivariate case. Let $\alpha := (\alpha_1, \ldots, \alpha_d) \in \{0,1\}^d$ be a multiindex. The tensor product B-spline space $\widehat{V}_\ell = S_{p,\boldsymbol{\tau}_\ell} \cap H_0^1(\widehat{\Omega})$ with $\boldsymbol{\tau}_\ell = (\boldsymbol{\tau}_{\ell,1}, \ldots, \boldsymbol{\tau}_{\ell,d})$ is split into the direct sum of 2^d subspaces

$$\widehat{V}_\ell = \bigoplus_{\alpha \in \{0,1\}^d} S^\alpha_{p,\boldsymbol{\tau}_\ell} \quad \text{where} \quad S^\alpha_{p,\boldsymbol{\tau}_\ell} = S^{\alpha_1}_{p,\boldsymbol{\tau}_{\ell,1}} \otimes \cdots \otimes S^{\alpha_d}_{p,\boldsymbol{\tau}_{\ell,d}}. \tag{5.1}$$

Again, we define L^2-orthogonal projectors

$$\widehat{\mathbf{Q}}^\alpha_{p,\boldsymbol{\tau}_\ell} := Q^{\alpha_1}_{p,\boldsymbol{\tau}_{\ell,1}} \otimes \cdots \otimes Q^{\alpha_d}_{p,\boldsymbol{\tau}_{\ell,d}} : \widehat{V} \to S^\alpha_{p,\boldsymbol{\tau}_\ell}.$$

The projector $\widehat{\mathbf{Q}}^0_{p,\boldsymbol{\tau}_\ell}$ from Section 4.2 is consistent with this definition, for the choice $\alpha = 0$. Since the splitting is L^2-orthogonal, we obviously have the following result.

$$\widehat{u}_\ell = \sum_{\alpha \in \{0,1\}^d} \widehat{\mathbf{Q}}^\alpha_{p,\boldsymbol{\tau}_\ell} \widehat{u}_\ell \quad \text{and} \quad \|\widehat{u}_\ell\|^2_{L^2(\widehat{\Omega})} = \sum_{\alpha \in \{0,1\}^d} \|\widehat{\mathbf{Q}}^\alpha_{p,\boldsymbol{\tau}_\ell} \widehat{u}_\ell\|^2_{L^2(\widehat{\Omega})} \quad \forall \widehat{u}_\ell \in \widehat{V}_\ell. \tag{5.2}$$

The next theorem shows that the splitting is also stable in H^2.

Theorem 5.1. *Let $p \in \mathbb{N}$ with $p \geq 3$. Then there exists a constant $c > 0$ such that*

$$c^{-1}\|\widehat{u}_\ell\|^2_{\bar{\mathcal{B}}} \leq \sum_{\alpha\in\{0,1\}^d} \|\widehat{\mathbf{Q}}^\alpha_{p,\boldsymbol{\tau}_\ell}\widehat{u}_\ell\|^2_{\bar{\mathcal{B}}} \leq c\|\widehat{u}_\ell\|^2_{\bar{\mathcal{B}}} \quad \forall \widehat{u}_\ell \in \widehat{V}_\ell.$$

Proof. Theorem 4.6 states the stability of $Q^0_{p,\boldsymbol{\tau}_\ell}$ in the H^2-seminorm. The stability of $Q^1_{p,\boldsymbol{\tau}_\ell}$ in the H^2-seminorm follows using the triangle inequality. The stability of these statements in the L^2-norm is obvious. From these observations, the right inequality follows by arguments that are completely analogous to those of the proof of Theorem 4.8.

The left inequality follows from (5.2) and the triangle inequality. □

For notational convenience, we restrict the setup of the smoother to the two-dimensional case. We write the splitting (5.1) as

$$\widehat{V}_\ell = S^{00}_{p,\boldsymbol{\tau}_\ell} \oplus S^{01}_{p,\boldsymbol{\tau}_\ell} \oplus S^{10}_{p,\boldsymbol{\tau}_\ell} \oplus S^{11}_{p,\boldsymbol{\tau}_\ell}, \quad \text{where} \quad S^{\alpha_1,\alpha_2}_{p,\boldsymbol{\tau}_\ell} = S^{\alpha_1}_{p,\boldsymbol{\tau}_{\ell,1}} \otimes S^{\alpha_2}_{p,\boldsymbol{\tau}_{\ell,2}}.$$

Following the ideas of Refs. 15, 29, we construct local smoothers L_α for any of the spaces $V_{\ell,\alpha} := S^\alpha_{p,\boldsymbol{\tau}_\ell}$. These local contributions are chosen such that they satisfy the corresponding local condition

$$\bar{\mathcal{B}}_{\ell,\alpha} + \beta\widehat{\mathcal{M}}_{\ell,\alpha} \leq L_{\ell,\alpha} \leq c(\bar{\mathcal{B}}_{\ell,\alpha} + (\beta + h^{-4})\widehat{\mathcal{M}}_{\ell,\alpha}),$$

where

$$\bar{\mathcal{B}}_{\ell,\alpha} := \mathbf{P}^T_{\ell,\alpha}\bar{\mathcal{B}}_\ell\mathbf{P}_{\ell,\alpha} \qquad \text{and} \qquad \widehat{\mathcal{M}}_{\ell,\alpha} := \mathbf{P}^T_{\ell,\alpha}\widehat{\mathcal{M}}_\ell\mathbf{P}_{\ell,\alpha}$$

and $\mathbf{P}_{\ell,\alpha}$ is the matrix representation of the canonical embedding $V_{\ell,\alpha} \to V_\ell$. The canonical embedding has tensor product structure, i.e., $P_{\ell,\alpha_1} \otimes \cdots \otimes P_{\ell,\alpha_d}$, where the P_{ℓ,α_i} are the matrix representations of the corresponding univariate embeddings. In the two-dimensional case, $\bar{\mathcal{B}}_\ell$ and $\widehat{\mathcal{M}}_\ell$ have the representation

$$\bar{\mathcal{B}}_\ell = B \otimes M + M \otimes B \quad \text{and} \quad \widehat{\mathcal{M}}_\ell = M \otimes M,$$

where B and M are the corresponding univariate stiffness and mass matrices (not necessarily equal for both spacial directions). For notational convenience, we do not indicate the spacial direction and the grid level for these matrices. Restricting $\bar{\mathcal{B}}_\ell$ to the subspace $V_{\ell,(\alpha_1,\alpha_2)}$ gives

$$\bar{\mathcal{B}}_{\ell,(\alpha_1,\alpha_2)} = B_{\alpha_1} \otimes M_{\alpha_2} + M_{\alpha_1} \otimes B_{\alpha_2},$$

where $B_{\alpha_i} = P^T_{\ell,\alpha_i}BP_{\ell,\alpha_i}$ and $M_{\alpha_i} = P^T_{\ell,\alpha_i}MP_{\ell,\alpha_i}$. We define

$$\bar{\mathcal{A}}_\ell := \bar{\mathcal{B}}_\ell + \beta\widehat{\mathcal{M}}_\ell \quad \text{and} \quad \bar{\mathcal{A}}_{\alpha_1,\alpha_2} := \bar{\mathcal{B}}_{\alpha_1,\alpha_2} + \beta\widehat{\mathcal{M}}_{\alpha_1,\alpha_2}.$$

The inverse inequality for $S^0_{p,\boldsymbol{\tau}_{\ell,i}}$ (Theorem 4.5), allows us to estimate

$$B_0 \leq \sigma M_0,$$

where $\sigma = \sigma_0 h_{\ell,\min}^{-4}$ and $\sigma_0 = 144$. Using this, we define the smoothers L_{α_1,α_2} and obtain estimates for them as follows:

$$\begin{aligned}
\bar{\mathcal{A}}_{00} &\le (2\sigma+\beta) M_0 \otimes M_0 && =: L_{00} \le c(\bar{\mathcal{A}}_{00} + h^{-4}\widehat{\mathcal{M}}_{00}),\\
\bar{\mathcal{A}}_{01} &\le M_0 \otimes ((\sigma+\beta) M_1 + B_1) && =: L_{01} \le c(\bar{\mathcal{A}}_{01} + h^{-4}\widehat{\mathcal{M}}_{01}),\\
\bar{\mathcal{A}}_{10} &\le (B_1 + (\sigma+\beta) M_1) \otimes M_0 && =: L_{10} \le c(\bar{\mathcal{A}}_{10} + h^{-4}\widehat{\mathcal{M}}_{10}),\\
\bar{\mathcal{A}}_{11} &= B_1 \otimes M_1 + M_1 \otimes B_1 + \beta M_1 \otimes M_1 && =: L_{11} \le c(\bar{\mathcal{A}}_{11} + h^{-4}\widehat{\mathcal{M}}_{11}).
\end{aligned}$$

The extension to three and more dimensions is completely straightforward (cf. Ref. 15). For each of the subspaces $V_{\ell,\alpha}$, we have defined a symmetric and positive definite smoother L_α. The overall smoother is given by

$$L_\ell := \sum_{\alpha\in\{0,1\}^d} (\mathbf{Q}^{D,\alpha})^T L_\alpha \mathbf{Q}^{D,\alpha},$$

where $\mathbf{Q}^{D,\alpha} = \widehat{\mathcal{M}}_\alpha^{-1} \mathbf{P}_{\ell,\alpha}^T \widehat{\mathcal{M}}_\ell$ is the matrix representation of the L^2-projection from V_ℓ to $V_{\ell,\alpha}$. Completely analogous to Section 5.2 in Ref. 15, we obtain

$$L_\ell^{-1} = \sum_{\alpha\in\{0,1\}^d} \mathbf{P}_{\ell,\alpha} L_\alpha^{-1} \mathbf{P}_{\ell,\alpha}^T.$$

Theorem 5.2. *Let $d \in \mathbb{N}$ and $p \in \mathbb{N}$ with $p \ge 3$. The subspace corrected mass smoother L_ℓ, satisfies* (3.6)*, i.e.,*

$$(\mathcal{A}_\ell \underline{u}_\ell, \underline{u}_\ell) \le \frac{1}{\tau_\ell}(L_\ell \underline{u}_\ell, \underline{u}_\ell) \le C_S\, \lambda_\ell\, ((\mathcal{A}_\ell + h^{-4}\mathcal{M}_\ell)\underline{u}_\ell, \underline{u}_\ell) \quad \forall\, \underline{u}_\ell \in \mathbb{R}^{\dim V_\ell}$$

for all $\tau \in (0,\tau_0)$, where $\tau_0 > 0$ is some constant.

Proof. The inequalities

$$(\bar{\mathcal{A}}_\ell \underline{u}_\ell, \underline{u}_\ell) \le (L_\ell \underline{u}_\ell, \underline{u}_\ell) \le c((\bar{\mathcal{A}}_\ell + h^{-4}\widehat{\mathcal{M}}_\ell)\underline{u}_\ell, \underline{u}_\ell)$$

were shown in Theorem 17 in Ref. 29 for $\beta = 0$. Note that no part of that proof requires uniform grids. So, the proof can be used almost verbatim also in the context of this paper. Using (5.2), the extension to $\beta > 0$ is straightforward. Using this and Lemma 2.1, we get

$$(\widehat{\mathcal{A}}_\ell \underline{u}_\ell, \underline{u}_\ell) \le d(\bar{\mathcal{A}}_\ell \underline{u}_\ell, \underline{u}_\ell) \le \frac{d}{\tau_\ell}(L_\ell \underline{u}_\ell, \underline{u}_\ell) \le c((\widehat{\mathcal{A}}_\ell + h^{-4}\widehat{\mathcal{M}}_\ell)\underline{u}_\ell, \underline{u}_\ell)$$

for some constant $c > 0$. Using (2.4), we obtain

$$(\mathcal{A}_\ell \underline{u}_\ell, \underline{u}_\ell) \le \frac{c_1}{\tau_\ell}(L_\ell \underline{u}_\ell, \underline{u}_\ell) \le c_2((\mathcal{A}_\ell + h^{-4}\mathcal{M}_\ell)\underline{u}_\ell, \underline{u}_\ell)$$

for some constants $c_1, c_2 > 0$, which finishes the proof since λ_ℓ is bounded from below by a constant (Lemma 4.1). $\square$

Corollary 5.1. *Suppose that we solve the linear system* (2.8) *using a multigrid solver as outlined in Section 3 and using the subspace corrected mass smoother as* *outlined in Section 5, then the convergence of the multigrid solver is described by the relation*

$$((I - B_L^s \mathcal{A}_L)\underline{u}_L, \underline{u}_L)_{\mathcal{A}_L} \leq \left(1 - \frac{1}{CL}\right) (\underline{u}_L, \underline{u}_L)_{\mathcal{A}_L}\,, \tag{5.3}$$

where the constant C is independent of the grid sizes h_ℓ, the number of levels L, the spline degree p and the choice of the scaling parameter β. It may depend on d, the constants c_1, c_2, c_q, and c_r and the shape of Ω, cf. Notation 2.1.

Proof. We use Theorem 3.1, whose assumptions are shown by Theorem 4.9 and the combination of Lemma 3.1 and Theorem 5.2. □

Remark 5.1. The operator L_ℓ^{-1} can be applied efficiently because all of the local contributions L_{00}, L_{01} and L_{10} can be inverted efficiently because they are tensor products. For example, we have $L_{00}^{-1} = \frac{1}{2\sigma+\beta}(M_0^{-1} \otimes I)(I \otimes M_0^{-1})$, where both $M_0^{-1} \otimes I$ and $I \otimes M_0^{-1}$ can be realized by applying direct solvers for the univariate mass matrix to several right-hand sides. The operator L_{11} is the sum of two tensor products. So, it has to be inverted as a whole. However, the dimension of the corresponding space is so small that the corresponding computational costs are negligible. More details on how to realize the smoother computationally efficient are given in Section 5 in Ref. 15. There, it is outlined where an efficient realization of the subspace corrected mass smoother is also possible in case of more than two dimensions.

5.2. *Symmetric Gauss-Seidel smoother and a hybrid smoother*

The second smoother we consider is a symmetric Gauss-Seidel smoother consisting of one forward sweep and one backward sweep. It can be shown that this smoother satisfies Condition (3.6), where the constant C_S depends on the spline degree, see Ref. 29. This means that also the overall convergence result (5.3) holds true, where again C depends on the spline degree. The symmetric Gauss-Seidel smoother works well for domains with nontrivial geometry transformations, but degenerated for large spline degrees (cf. Refs. 9, 17).

Since the symmetric Gauss-Seidel smoother works well for nontrivial geometry transformations and the subspace corrected mass smoother is robust with respect to the spline degree, we combine these smoothers into a hybrid smoother, which was first introduced in Ref. 29. This hybrid smoother consists of one forward Gauss-Seidel sweep, followed by one step of the subspace corrected mass smoother, finally followed by one backward Gauss-Seidel sweep.

6. Numerical experiments

In this section, we present the results of numerical experiments performed with the proposed algorithm. As computational domains, we first consider the unit square, then we consider the nontrivial geometries displayed in Figures 1 (two-dimensional domain) and 2 (three-dimensional domain). We consider the problem

$$\begin{aligned} \beta u + \Delta^2 u &= f \quad \text{in} \quad \Omega, \\ u &= g_1 \quad \text{on} \quad \partial\Omega, \\ \Delta u &= g_2 \quad \text{on} \quad \partial\Omega, \end{aligned}$$

where

$$f(x) = (\beta + d^2\pi^4) \prod_{k=1}^{d} \sin(\pi x_k), \quad g_1(x) = \prod_{k=1}^{d} \sin(\pi x_k), \quad g_2(x) = -d\pi^2 \prod_{k=1}^{d} \sin(\pi x_k).$$

The discretization space on the parameter domain is the space of tensor-product B-splines. On the coarsest level ($\ell = 0$), we choose

$$\boldsymbol{\tau}_{0,i} = (0,\, 1/3,\, 1/2,\, 4/5,\, 1), \tag{6.1}$$

for all spatial directions $i = 1, \dots, d$. The discretization on level ℓ is obtained by preforming ℓ uniform h-refinement steps. The spline spaces have maximum continuity and spline degree p. We solve the resulting system using the preconditioned conjugate gradient (PCG) with a V-cycle multigrid method with 1 pre and 1 post smoothing step, as preconditioner. A random initial guess is used and the stopping criterion is

$$\|\underline{r}_L^{(k)}\| \le 10^{-8} \|\underline{r}_L^{(0)}\|,$$

where $\underline{r}_L^{(k)} := \underline{f}_L - \mathcal{A}_L \underline{x}_L^{(k)}$ is the residual at step k and $\|\cdot\|$ denotes the Euclidean norm. All numerical experiments are implemented using the G+Smo library, see Ref. 20.

6.1. *Numerical experiments on parameter domain*

We start with the unit square as the domain, that is, $\Omega = (0,1)^2$. Note that $g_1(x) = g_2(x) = 0$ for this domain. For now, we consider the symmetric Gauss-Seidel smoother and the subspace corrected mass smoother. For both smoothers, we choose $\tau = 1$. The iteration counts are displayed in Table 1 for $\beta = 1$, and in Table 2 for $\beta = 10^7$.

From the tables, we see that the symmetric Gauss-Seidel smoother preforms well for small spline degrees, but degenerates for larger spline degrees. These results are not surprising since it is known that standard smoothers do not work well for large spline degrees (cf. Refs. 9, 17). Due to Corollary 5.1, the multigrid solver with subspace corrected mass smoother is robust with respect to the spline degree. The tables do reflex this. The rates slightly improve when the spline degree is increased. This might be due to the fact that the constants for $L^2 - H^1$-approximation errors estimate (like in Theorem 4.3) decrease if p is increased, cf. Ref. 5, while the constant in the inverse estimate (Theorem 4.5) is uniformly bounded. The product of these constants enters the convergence estimate. However, the iteration numbers

$\ell \diagdown p$	3	4	5	6	7	8	9
Symmetric Gauss-Seidel							
5	10	16	28	45	71	120	210
6	10	16	27	44	71	119	209
7	10	16	27	44	72	117	212
8	11	16	27	45	72	120	221
Subspace corrected mass smoother, $\sigma_0^{-1} = 0.02$							
5	126	122	114	105	98	93	85
6	131	129	123	116	110	105	100
7	132	133	127	121	116	110	106
8	133	134	130	124	118	114	110

Table 1. Iteration counts for 2D parametric domain, $\beta = 1$

$\ell \diagdown p$	3	4	5	6	7	8	9
Symmetric Gauss-Seidel							
5	10	16	28	45	71	119	211
6	10	16	27	44	71	118	208
7	10	16	27	44	72	117	212
8	11	16	27	45	72	119	221
Subspace corrected mass smoother, $\sigma_0^{-1} = 0.02$							
5	124	121	113	104	96	92	85
6	131	129	123	116	110	105	99
7	132	133	127	120	116	110	106
8	133	134	130	124	116	118	114

Table 2. Iteration counts for 2D parametric domain, $\beta = 10^7$

are relatively high. Table 3 shows the iteration numbers when using an uniform grid with spacing 1/4 on the coarsest level ($\ell = 0$), rather than the grid (6.1). The numbers in Table 3 are significantly smaller. This implies that the subspace corrected mass smoother is sensitive to the quasi-uniformity constant c_q.

6.2. *Numerical experiments on physical domain*

Now, we consider a domain with a nontrivial geometry transformation as displayed in Figures 1 and 2. The convergence of the subspace corrected mass smoother degrades significantly due to the nontrivial geometry mapping. To mitigate this, we consider the hybrid smoother described in Section 5.2. Table 4 and Table 5 display the iteration numbers for the 2D and 3D physical domains, respectively. These iteration numbers are relatively small and seam to be robust with respect to both

$\ell \setminus p$	3	4	5	6	7	8	9
Subspace corrected mass smoother, $\sigma_0^{-1} = 0.015$							
5	41	40	39	37	35	34	33
6	41	41	39	37	36	35	34
7	42	42	40	39	37	35	35
8	42	42	41	39	37	37	35

Table 3. Iteration counts for 2D parametric domain with uniform grid, $\beta = 1$

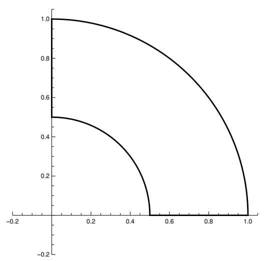

Fig. 1. The two-dimensional domain

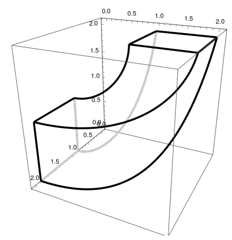

Fig. 2. The three-dimensional domain

grid size and spline degree. Although the hybrid smoother is more expensive, as one smoothing step can be view as two smoothing steps, the reduction of iteration numbers outweigh this cost for larger spline degrees $p > 4$. For smaller spline degrees, the symmetric Gauss-Seidel smoother is ideal choice.

$\ell \setminus p$	3	4	5	6	7	8	9
Hybrid smoother, $\beta = 1$							
5	28	23	23	24	26	27	27
6	28	23	22	25	24	26	26
7	29	23	22	23	24	24	24
8	28	22	21	21	22	22	22
Hybrid smoother, $\beta = 10^7$							
5	27	23	23	24	26	27	28
6	28	23	22	25	25	26	26
7	29	23	22	23	24	24	24
8	28	22	21	21	22	22	22

Table 4. Iteration counts for 2D Physical domain, $\sigma_0^{-1} = 0.015$, $\tau = 0.1$

Remark 6.1. All experiments presented so far, have also been performed for the

$\ell \setminus p$	3	4	5	6	7
Hybrid smoother, $\beta = 1$					
1	16	18	21	27	30
2	31	28	26	29	32
3	46	37	33	33	35
4	50	41	34	34	mem
Hybrid smoother, $\beta = 10^7$					
1	10	11	13	17	20
2	12	16	20	25	29
3	16	19	22	24	28
4	29	28	28	28	mem

Table 5. Iteration counts for 3D Physical domain, $\sigma_0^{-1} = 0.020$, $\tau = 0.1$

choice $\beta = 0$. In this case, one obtains iteration numbers that are identical to those obtained for $\beta = 1$. Therefore, we chose to only display the results for $\beta = 1$.

6.3. *Numerical experiments with singular mass matrix*

As mentioned in the introduction, for optimal control problems with limited observation, it is of interest to solve (1.1), where βu is multiplied with the characteristic function for the observation domain $\mathcal{O} \subsetneq \Omega$. In this case, the variational problem (2.2) takes the form

$$\beta(u,v)_{L^2(\mathcal{O})} + (\Delta u, \Delta v)_{L^2(\Omega)} = (f,v)_{L^2(\Omega)} \quad \forall\, v \in V. \tag{6.2}$$

Since $\mathcal{O} \subsetneq \Omega$, the resulting mass is singular. In general, it is not easy to apply the subspace corrected mass smoother since the first term in (6.2) would have to be approximated by the full mass matrix. As a consequence, we do not consider the subspace corrected mass smoother or hybrid smoother and we only consider the symmetric Gauss-Seidel smoother. Figure 3 displays the computational domain

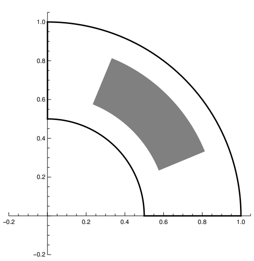

Fig. 3. The two-dimensional domain with limited observation

where the limited observation is marked in gray. Table 6 show the iteration numbers

$\ell \setminus \beta$	10^0	10^3	10^5	10^7
Full observation $\mathcal{O} = \Omega$				
5	28	28	28	28
6	29	29	29	29
7	29	29	29	29
8	28	28	28	28
Limited observation $\mathcal{O} \subsetneq \Omega$				
5	28	28	28	36
6	29	29	29	33
7	29	29	29	32
8	28	28	28	30

Table 6. Iteration counts for 2D Physical domain with full and limited observation, $p = 3$.

for both full oberservation and limited observation. The iteration counts are similar and has only a small increase for large values of β. We note that the theory does not cover the case of limited observation.

Remark 6.2. The multigrid solvers presented in this paper only consider single-patch discretizations. In many practical applications, the representation of the computational domain is only viable using multiple patches. Multigrid solvers for multi-patch discretizations of second order elliptic equations are considered in Ref. 31. For fourth order problems, the setup of $H^2(\Omega)$-conforming discretizations is a challenging and active research topic, see, e.g., Refs. 19, 18, 33. Multigrid methods for such discretizations might be considered. One viable alternative is to consider domain decomposition methods based on non-conforming coupling of the patches. The multigrid solvers can then be used to efficiently approximate the action of the inverse of the local stiffness matrices, see, e.g. Ref. 27 and references therein on inexact domain decomposition methods in the context of IgA.

Acknowledgment

This research was funded by the Austrian Science Fund (FWF): P31048. The first author acknowledges support from the Research Council of Norway, grant 301013.

References

1. A. Beigl, J. Sogn and W. Zulehner, Robust preconditioners for multiple saddle point problems and applications to optimal control problems, *SIAM Journal on Matrix Analysis and Applications* **41** (2020) 1590–1615.
2. J. H. Bramble, *Multigrid methods* (Routledge, 2018).

3. J. H. Bramble, J. E. Pasciak, J. P. Wang and J. Xu, Convergence estimates for multigrid algorithms without regularity assumptions, *Mathematics of Computation* **57** (1991) 23–45.
4. S. C. Brenner, An optimal-order nonconforming multigrid method for the biharmonic equation, *SIAM Journal on Numerical Analysis* **26** (1989) 1124–1138.
5. A. Bressan and E. Sande, Approximation in fem, dg and iga: a theoretical comparison, *Numer. Math.* (2019) 923–942.
6. L. Chen, J. Hu and X. Huang, Multigrid methods for Hellan–Herrmann–Johnson Mixed Method of Kirchhoff Plate Bending Problems, *Journal of Scientific Computing* (2015) 1–24.
7. P. G. Ciarlet, *The finite element method for elliptic problems* (SIAM, 2002).
8. A. P. de la Riva, F. J. Gaspar and C. Rodrigo, On the robust solution of an isogeometric discretization of bilaplacian equation by using multigrid methods, *Computers & Mathematics with Applications* **80** (2020) 386–394.
9. K. P. S. Gahalaut, J. K. Kraus and S. K. Tomar, Multigrid methods for isogeometric discretization, *Computer methods in applied mechanics and engineering* **253** (2013) 413–425.
10. V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations: theory and algorithms*, volume 5 (Springer Science & Business Media, 2012).
11. P. Grisvard, *Singularities in boundary value problems* (Springer, Berlin, 1992).
12. P. Grisvard, *Elliptic Problems in Nonsmooth Domains. Reprint of the 1985 hardback ed.* (Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2011).
13. W. Hackbusch, *Multi-Grid Methods and Applications*, volume 4 (Springer Science & Business Media, 2013).
14. M. R. Hanisch, Multigrid preconditioning for the biharmonic Dirichlet problem, *SIAM Journal on Numerical Analysis* **30** (1993) 184–214.
15. C. Hofreither and S. Takacs, Robust multigrid for isogeometric analysis based on stable splittings of spline spaces, *SIAM Journal on Numerical Analysis* **4** (2017) 2004–2024.
16. C. Hofreither, S. Takacs and W. Zulehner, A robust multigrid method for isogeometric analysis in two dimensions using boundary correction, *Computer Methods in Applied Mechanics and Engineering* **316** (2017) 22–42.
17. C. Hofreither and W. Zulehner, Spectral analysis of geometric multigrid methods for isogeometric analysis, in *International Conference on Numerical Methods and Applications* (Springer, 2014), pp. 123–129.
18. T. J. Hughes, G. Sangalli, T. Takacs and D. Toshniwal, Smooth multi-patch discretizations in isogeometric analysis, in *Handbook of Numerical Analysis* (Elsevier, 2021), volume 22, pp. 467–543.
19. M. Kapl, G. Sangalli and T. Takacs, An isogeometric c1 subspace on unstructured multi-patch planar domains, *Computer Aided Geometric Design* **69** (2019) 55–75.
20. A. Mantzaflaris, J. Sogn, S. Takacs and others (see website), G+Smo, https://github.com/gismo/gismo, 2021.
21. K.-A. Mardal, B. F. Nielsen and M. Nordaas, Robust preconditioners for PDE-constrained optimization with limited observations, *BIT* **57** (2017) 405–431.
22. K.-A. Mardal, J. Sogn and S. Takacs, Robust preconditioning and error estimates for optimal control of the convection–diffusion–reaction equation with limited observation in isogeometric analysis, *SIAM Journal on Numerical Analysis* **60** (2022) 195–221.
23. M. A. Olshanskii and A. Reusken, On the convergence of a multigrid method for linear reaction-diffusion problems, *Computing* **65** (2000) 193–202.
24. K. Rafetseder and W. Zulehner, A decomposition result for Kirchhoff plate bending problems and a new discretization approach, *SIAM Journal on Numerical Analysis*

56 (2018) 1961–1986.

25. E. Sande, C. Manni and H. Speleers, Sharp error estimates for spline approximation: Explicit constants, n-widths, and eigenfunction convergence, *Mathematical Models and Methods in Applied Sciences* **29** (2019) 1175–1205.
26. E. Sande, C. Manni and H. Speleers, Explicit error estimates for spline approximation of arbitrary smoothness in isogeometric analysis, *Numerische Mathematik* (2020) 1–41.
27. R. Schneckenleitner and S. Takacs, Inexact IETI-DP for conforming isogeometric multi-patch discretizations, in *Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2021* (2023), to appear. This paper is also available as arXiv preprint (arXiv:2110.06087).
28. J. Sogn, Schur complement preconditioners for multiple saddle point problems and applications, Ph.D. thesis, Johannes Kepler University Linz, 2018.
29. J. Sogn and S. Takacs, Robust multigrid solvers for the biharmonic problem in isogeometric analysis, *Computers & Mathematics with Applications* **77** (2019) 105–124.
30. J. Sogn and W. Zulehner, Schur complement preconditioners for multiple saddle point problems of block tridiagonal form with application to optimization problems, *IMA Journal of Numerical Analysis* **39** (2018) 1328–1359.
31. S. Takacs, Robust approximation error estimates and multigrid solvers for isogeometric multi-patch discretizations, *Mathematical Models and Methods in Applied Sciences* **28** (2018) 1899–1928.
32. S. Takacs and T. Takacs, Approximation error estimates and inverse inequalities for B-splines of maximum smoothness, *Mathematical Models and Methods in Applied Sciences* **26** (2016) 1411–1445.
33. P. Weinmüller and T. Takacs, An approximate C1 multi-patch space for isogeometric analysis with a comparison to Nitsche's method, *arXiv preprint arXiv:2202.04516* .
34. S. Zhang and J. Xu, Optimal solvers for fourth-order pdes discretized on unstructured grids, *SIAM Journal on Numerical Analysis* **52** (2014) 282–307.