

Constraints for eliminating the Gibbs phenomenon in finite element approximation spaces

Marco F. P. ten Eikelder 

*Institute for Mechanics, Computational Mechanics Group,
Technical University of Darmstadt,
Franziska-Braun-Straße 7, 64287 Darmstadt, Germany
marco.eikelder@tu-darmstadt.de

Stein K. F. Stoter 

*Department of Mechanical Engineering,
Eindhoven University of Technology,
P.O. Box 513, 5600 MB Eindhoven, Netherlands
k.f.s.stoter@tue.nl*

Yuri Bazilevs 

*School of Engineering,
Brown University,
184 Hope Street, Providence, RI 02912, USA
yuri_bazilevs@brown.edu*

Dominik Schillinger 

*Institute for Mechanics, Computational Mechanics Group,
Technical University of Darmstadt,
Franziska-Braun-Straße 7, 64287 Darmstadt, Germany
dominik.schillinger@tu-darmstadt.de*

Received 17 February 2023

Revised 30 October 2023

Accepted 2 November 2023

Published 24 January 2024

Communicated by F. Brezzi

One of the major challenges in finite element methods is the mitigation of spurious oscillations near sharp layers and discontinuities known as the Gibbs phenomenon. In this paper, we propose a set of functionals to identify spurious oscillations in best approximation problems in finite element spaces. Subsequently, we adopt these functionals in the formulation of constraints in an effort to eliminate the Gibbs phenomenon. By enforcing these constraints in best approximation problems, we can entirely eliminate over- and

*Corresponding author

undershoot in one-dimensional continuous approximations, and significantly suppress them in one- and higher-dimensional discontinuous approximations.

Keywords: Gibbs phenomenon; finite element methods; constrained optimization; isogeometric analysis; discontinuous Galerkin.

AMS Subject Classification: 65N30, 65K10, 35L67

1. Introduction

1.1. *Historical overview*

The discovery of the Gibbs phenomenon may be traced back to Henry Wilbraham (1848), and the phenomenon was rediscovered by J. Willard Gibbs (1898–1899), in their studies on Fourier series.^{10, 15} It is traditionally described as the inability to recover point values of a discontinuous function by a truncated Fourier expansion. Near the discontinuity, the error does not vanish as the number of terms in the expansion is increased, and the magnitude of the over- and undershoots tends to a fixed limit. The limiting value is known as the Gibbs constant. It is less well known that the Gibbs phenomenon also occurs in truncated expansions of other sets of orthogonal functions.^{26, 28} In fact, the associated Gibbs constants are often identical, as is the case for expansions with Legendre, Hermite, or Laguerre polynomials.^{6, 8}

Fundamentally, the Gibbs phenomenon has, however, little to do with Fourier series or expansions in orthogonal polynomials. The effect arises from the best approximation in a square integral metric, of which these expansions are examples.⁹ As such, it also occurs in best approximation problems in the L^2 -metric by piecewise linear polynomials⁹ or splines.³⁵ The role of the metric herein is crucial: spurious oscillations that appear in the L^2 -metric are significantly more severe than those that occur in the L^q -metric when q tends to 1, for which they are in some cases even completely absent.¹¹ A detailed study on the possible elimination of the Gibbs phenomenon in L^q -best approximation by piecewise linear finite element shape functions is presented in Ref. 17. In the last few decades, the $L^1(\Omega)$ functional settings has hence been explored as the point of departure for approximating solutions to partial differential equations (PDEs).^{11, 27, 30, 31, 34} The main challenge behind these approaches is that they require the minimization of a non-differentiable functional, which leads to a poorly behaved nonlinear problem. As a consequence, there is a lack of practical algorithms for solving even standard problems in computational mathematics. Additionally, even though approximations in subspaces of $L^1(\Omega)$ reduce spurious oscillations, on some meshes these do not vanish in general.^{16, 17}

The more conventional (Bubnov-)Galerkin method produces solutions that are optimal in an inner product-induced norm (associated with subspaces of $L^2(\Omega)$). As such, approximations of interior and boundary layers indeed tend to suffer from spurious oscillations. This issue is well known in the finite element community,

and many attempts at resolutions have been proposed. Arguably the most successful is the class of residual-based stabilized methods,^{4, 21, 22} which are primarily adopted for fluid mechanics-related applications. Residual-based stabilization significantly improves the solution quality in regions free of abrupt changes, but the Gibbs phenomenon still occurs in regions with sharp layers. As a remedy, the finite element formulation is often augmented with a nonlinear stabilization mechanism that locally introduces artificial diffusion.^{3, 24, 43} These methods are referred to as shock or discontinuity capturing methods.

In the case of nonlinear (hyperbolic) evolution equations, the above stabilization methods do still not suffice. In order to enhance the quality of numerical approximations for these types of problems, algorithms have been designed to inherit certain stability properties of the underlying PDE. The prevalent example is the entropy stability property possessed by entropy solutions. Weak solutions of nonlinear evolution equations are not unique and the entropy stability property singles out the entropy solution as the physically relevant solution.²⁹ The entropy stability concept, which reduces for many physical systems to an energy-dissipation property, has frequently been used in the construction of stable finite element methods.^{12, 32, 33, 39, 40, 42} Even though the solution quality enhances significantly, numerical solutions that inherit entropy stability do not preclude spurious oscillations. For particular variable sets, the Galerkin method may even exactly satisfy the entropy stability condition but still exhibit spurious oscillations.²³ Evidently, the entropy stability concept is not inextricably linked to the Gibbs phenomenon. It does, however, seem to be a good indicator for the identification of shock waves, and thus as an indicator of where the Gibbs phenomenon might manifest.

A stability concept that is more directly targeted at removing the Gibbs phenomenon, is the total variation diminishing (TVD) property introduced by Harten.^{13, 14} Solutions with the TVD property preclude the growth of the total variation of the solution. The design of numerical schemes with the TVD property still is an active area of research. The incentive for the design of TVD schemes is the desire to produce numerical approximations that satisfy the maximum principle, as well as certain monotonicity properties. Despite its success, particularly in the finite difference and finite volume communities, the applicability of TVD schemes is limited. Namely, it is solely suitable for time-dependent and scalar conservation laws and does not provide any information on local solution quality. Moreover, its introduction in the discrete setting relies on a Cartesian grid and lacks frame-invariance.

The above two observations, namely (i) the occurrence of the Gibbs phenomenon in entropy stable discrete solutions, and (ii) the limitations of TVD schemes, have incentivized the design of a novel stability concept called *variation entropy theory*.⁴¹ This theory provides a local continuous generalization of the TVD stability condition for general conservation laws in an entropy framework. Similar to classical entropy solutions, variation entropy solutions satisfy an underlying stability condition. This stability condition serves in the discrete setting as an indicator of the Gibbs phenomenon. It has successfully been employed in the variation multiscale

(VMS) paradigm^{2, 18, 20} to design a framework for discontinuity capturing methods.⁴³

1.2. Objective

Despite the significant attention it has gained, a precise mathematical definition of the Gibbs phenomenon does not exist. Any attempt at eliminating the Gibbs phenomenon thus first requires an identification strategy. The identifier that we develop is rooted in variation entropy theory. We propose to eliminate the Gibbs phenomenon via enforcement of constraints.

This brings us to the main objective of this paper: *to identify a set of practical constraints that aim to eliminate the Gibbs phenomenon in the approximation of sharp layers and discontinuities in finite element spaces.* To facilitate the analysis, we discuss our results in the isogeometric analysis framework, which we think of as a generalization of C^0 - and C^{-1} -finite element spaces to higher-order continuity.

Some remarks are in order. First, it may seem feasible to construct constraints that remove oscillations by explicitly choosing the coefficients of the basis functions such that the numerical approximation does not exceed bounds of the analytical profile. This is however not a strategy that is realizable in practical computations. The challenge is thus to establish a set of constraints that can be adopted in practice. Second, the idea of *a priori* enforcing constraints in numerical methods is not new. A notable contribution in this regard is the work of Evans *et al.*,⁷ in which a framework for the enforcement of constraints in the VMS framework is presented.

1.3. Main results

The main result of this paper is a set of integral constraints that aim to identify and eliminate the Gibbs phenomenon. The occurrence of the Gibbs phenomenon in a certain approximation cannot solely be inferred from the approximation itself. Rather, it stands in relation to the function being approximated, and depends on the (sub)domain of interest. We propose an indicator of the form

$$\mathcal{G}_{\phi,\omega}(\phi^*) \leq 0, \quad (1.1)$$

where the function $\phi^* \in H^1(\tilde{\Omega})$ is an approximation of the function $\phi \in H^1(\Omega)$ on $\omega \subset \Omega$. Here, $H^1(\tilde{\Omega})$ is a broken Sobolev space and $\tilde{\Omega}$ is a collection of disjoint subdomains (precise definitions are provided in Sec. 4). We call this constraint the *Gibbs constraint*. It follows from the *Gibbs functional*, which we define as

$$\mathcal{G}_{\phi,\omega}(\phi^*) := \int_{\omega} g_{\phi}(\phi^*) \, dx, \quad (1.2)$$

with g_{ϕ} as

$$g_{\phi}(\phi^*) := \begin{cases} \|\nabla \phi^*\|_2^{-1} \nabla \phi^* \cdot \nabla(\phi^* - \phi) & \text{for } \nabla \phi^* \neq 0, \\ 0 & \text{for } \nabla \phi^* = 0. \end{cases} \quad (1.3)$$

We search functions ϕ^* as an approximation of ϕ for which the Gibbs constraint is satisfied on predetermined sets of subdomains ω . We study the application of the Gibbs constraints in the context of finite element best approximation problems. In particular, we consider the constrained best approximation problem

$$\phi^h = \operatorname{arginf}_{\theta^h \in \mathcal{K}_{p,\alpha}} \|\phi - \theta^h\|_{\mathcal{H}}, \quad (1.4)$$

where $\|\cdot\|_{\mathcal{H}}$ is a norm induced by a certain inner product and the feasible set is given by

$$\mathcal{K}_{p,\alpha} := \{\phi^h \in \mathcal{V}_{D,p,\alpha}^h : \mathcal{G}_{\phi,\omega_j}(\phi^*) \leq 0, j = 1, \dots, J, \omega_j \in \mathcal{T}_{\omega}\}. \quad (1.5)$$

Precise definitions of $\mathcal{V}_{D,p,\alpha}^h$ and $\mathcal{T}_{\omega}$ are provided in later in the paper. We demonstrate for sharp layers that finite element approximations of arbitrary degree and continuity are free of over- and undershoots when they satisfy the Gibbs constraint (for one-dimensional continuous approximations), or these oscillations are significantly suppressed (for discontinuous approximation).

The choice of subdomains ω depends on the regularity α of the finite element approximation space and the dimension of the domain. In one dimension, the constraints may be applied elementwise ($\omega_j = K_j$) when the finite element space is either discontinuous or $\mathcal{C}^0$ -continuous. For higher regularity finite element spaces ($\alpha \geq 1$), the subdomains ω need to be collections of neighboring elements, and the number of collected elements increases with the regularity. In higher dimensions, the Gibbs constraints are too restrictive for continuous finite element spaces, limiting its applicability to discontinuous finite element spaces.

1.4. Outline

The remainder of the paper is structured as follows. First, in Sec. 2, we provide preliminaries concerning function spaces and projectors. Then, in Sec. 3, we present an overview of the Gibbs phenomenon for best approximations in finite element spaces of arbitrary degree and continuity. In Sec. 4, we present the identification of the constraints in one spatial dimension. Next, we extend our construction to higher dimensions in Sec. 5. Finally, we provide a summary and outlook in Sec. 6.

2. Preliminaries

2.1. Function spaces

We adopt the standard functional analysis setting. We denote by $\Omega \subset \mathbb{R}^d$ the bounded, open and connected domain with spatial dimension d , and with boundary $\partial\Omega$. $L^2(\Omega)$ is the Lebesgue space of 2-integrable functions on Ω . Furthermore, $H^1(\Omega) \subset L^2(\Omega)$ is the Sobolev space of $L^2(\Omega)$ -functions with their gradient also in $[L^2(\Omega)]^d$. The subspace $H_0^1(\Omega) \subset H^1(\Omega)$ consists of functions with zero trace on $\partial\Omega$. The associated norms are denoted as $\|\cdot\|_{L^2(\Omega)} = \|\cdot\|_{\Omega}$ and $\|\cdot\|_{H_0^1(\Omega)}$. Furthermore, we use standard notation for the L^2 -inner product on Ω , $(\cdot, \cdot)_{L^2(\Omega)} = (\cdot, \cdot)_{\Omega}$ and

write $\langle \cdot, \cdot \rangle_D$ for the duality pairing $H^{-1/2}(D) \times H^{1/2}(D) \rightarrow \mathbb{R}$ on some boundary domain D .

In this paper, we consider finite element spaces of arbitrary degree and continuity. As such, we make use of the *isogeometric analysis* framework.^{1, 5, 19} We introduce knot vectors, univariate and multivariate B-splines, geometrical mappings and the physical mesh. The ordered knot vector Ξ is defined for degree p and dimensionality n as

$$\Xi := \{-1 = \xi_1, \xi_2, \dots, \xi_{n+p+1} = 1\}, \quad (2.1)$$

where $\xi_i \in \mathbb{R}$ represents the i th knot with $i = 1, \dots, n + p + 1$. We adopt the convention that $p = 0, 1, 2, \dots$ refers to piecewise constants, linears, quadratics, etc. In this work, we restrict ourselves to *open* knot vectors, meaning that the first and last knot appear $p + 1$ times. The univariate B-spline basis functions are defined recursively for $p = 0, 1, 2, \dots$. Starting with piecewise constant functions, we have

$$N_{i,0}(\xi) = \begin{cases} 1 & \text{if } \xi_i \leq \xi < \xi_{i+1}, \\ 0 & \text{otherwise,} \end{cases} \quad (2.2)$$

whereas for $p = 1, 2, \dots$, the B-spline basis functions are given by

$$N_{i,p}(\xi) = \frac{\xi - \xi_i}{\xi_{i+p} - \xi_i} N_{i,p-1}(\xi) + \frac{\xi_{i+p+1} - \xi}{\xi_{i+p+1} - \xi_{i+1}} N_{i+1,p-1}(\xi). \quad (2.3)$$

This definition is augmented with the convention that if a denominator (i.e. $\xi_{i+p} - \xi_i$ or $\xi_{i+p+1} - \xi_{i+1}$) is zero, that fraction is taken as zero. B-spline basis functions coincide with standard finite element Lagrange basis functions for $p = 0$ and 1, and differ for $p \geq 2$. The set of B-spline basis functions of degree p consists of nonnegative piecewise p th-order polynomial functions with local support, that form a partition of unity. Linear combinations of B-spline basis functions are referred to as B-splines. We introduce the vector $\zeta = \{\zeta_1, \dots, \zeta_m\}$ consisting of all knots without repetitions. The open knot vector implies that the basis functions are interpolatory at the ends of the interval. Inside a knot interval B-spline basis functions are smooth, whereas the repetition of a knot reduces the continuity of the B-spline basis function at that knot. More precisely, a B-spline basis function of degree p at a knot ξ_i with multiplicity k_i has $\alpha_i := p - k_i$ continuous derivatives at ξ_i (note that $\alpha_1 = \alpha_m = -1$). We denote the space of B-splines of polynomial degree p and regularity $\alpha = \{\alpha_1, \dots, \alpha_m\}$ as

$$S_{\alpha}^p := \text{span}\{N_{i,p}\}_{i=1}^n. \quad (2.4)$$

B-spline basis functions of degree p with uniform internal multiplicity p are interpolatory and span the same space as standard $\mathcal{C}^0$ -Lagrange basis functions. Similarly, B-spline basis functions of degree p with uniform internal multiplicity $p + 1$ are discontinuous and span the same space as discontinuous Lagrange basis functions.

The construction of multivariate B-splines follows from taking a tensor-product of the univariate B-splines. We introduce the open knot vectors

$$\Xi_l := \{\xi_{1,l}, \xi_{2,l}, \dots, \xi_{n_l+p_l+1,l}\}, \quad (2.5)$$

for polynomial degrees p_l and dimensionality integers n_l for $l = 1, \dots, d$. We define for each knot vector Ξ_l univariate B-spline basis functions $N_{i_l, p_l, l}$ of polynomial degree p_l for $i_l = 1, \dots, n_l$. Again we introduce the vector of knots with repetition $\zeta_l = \{\zeta_{1,l}, \dots, \zeta_{m_l, l}\}$ and regularity vector $\alpha_l = \{\alpha_{1,l}, \dots, \alpha_{m_l, l}\}$. The Cartesian mesh on the parametric domain $\hat{\Omega} = (-1, 1)^d \subset \mathbb{R}^d$ is now given by

$$\hat{\mathcal{T}} = \{Q = \otimes_{l=1, \dots, d} (\zeta_{i_l, l}, \zeta_{i_l+1, l}), 1 \leq i_l \leq m_l - 1\}. \quad (2.6)$$

The boundary of an open element $Q \in \hat{\mathcal{T}}$ is denoted ∂Q . The multivariate tensor-product B-spline basis functions are defined on the parametric mesh $\hat{\mathcal{T}}$ as

$$N_{i_1, \dots, i_d, p_1, \dots, p_d} = N_{i_1, p_1, 1} \otimes \dots \otimes N_{i_d, p_d, d}. \quad (2.7)$$

The associated tensor-product B-spline function space on $\hat{\mathcal{T}}$ is given by

$$S_{\alpha_1, \dots, \alpha_d}^{p_1, \dots, p_d} := \text{span}\{N_{i_1, \dots, i_d, p_1, \dots, p_d}\}_{i_1=1, \dots, i_d=1}^{n_1, \dots, n_d}. \quad (2.8)$$

Throughout this paper, we restrict ourselves to a uniform regularity vector $\alpha = \alpha_l$ in the interior, i.e. $\alpha_{2,l} = \dots = \alpha_{m-1,l} = \alpha$, and use equal polynomial degrees $p_1 = \dots = p_d = p$. We assume that the physical domain can be exactly described by the continuously differentiable geometrical map (with continuously differentiable inverse) $\mathbf{F} : \hat{\Omega} \rightarrow \Omega$. The physical mesh on Ω follows by applying the geometrical map $\mathbf{F}$ on elements of the parametric mesh

$$\mathcal{T} = \{K : K = \mathbf{F}(Q), Q \in \hat{\mathcal{T}}\}. \quad (2.9)$$

As usual, we demand the element $K \in \mathcal{T}$ to be shape-regular. The boundary of an element $K \in \mathcal{T}$ is denoted as ∂K . We define the Jacobian of the mapping $\mathbf{F}$ as $\mathbf{J} = \partial \mathbf{x} / \partial \xi$. Lastly, we introduce the finite element approximation space $\mathcal{V}_{p, \alpha}^h = \{\mathbf{F}(S_{\alpha}^p)\} := \{\mathbf{F}(S_{\alpha, \dots, \alpha}^{p, \dots, p})\}$, and its subspaces $\mathcal{V}_{0, p, \alpha}^h \subset \mathcal{V}_{p, \alpha}^h$ and $\mathcal{V}_{D, p, \alpha}^h \subset \mathcal{V}_{p, \alpha}^h$ consisting of those functions that satisfy homogeneous and inhomogeneous boundary conditions on $\partial\Omega$, respectively.

2.2. Projection operators

In this section, we introduce some orthogonal projection operators $\mathcal{P} : \mathcal{V} \rightarrow \mathcal{V}_{D, p, \alpha}^h$. Consider first the (constrained) $\mathcal{H}$ -best approximation problem

$$\phi^h = \underset{\theta^h \in \mathcal{V}_{p, \alpha}^h}{\operatorname{arginf}} \|\phi - \theta^h\|_{\mathcal{H}}, \quad (2.10)$$

subject to the trace equality

$$\phi^h|_{\partial\Omega} = \phi|_{\partial\Omega}, \quad (2.11)$$

where $\|\cdot\|_{\mathcal{H}}$ is a norm induced by the inner product $(\cdot, \cdot)_{\mathcal{H}}$. The constraint (2.11) may be homogenized via the adoption of a lift argument, as is standard in finite element methods, whereby the approximation space becomes $\mathcal{V}_{0; p, \alpha}^h \subset \mathcal{V}_{p, \alpha}^h$. The $\mathcal{H}$ -best approximation may be determined by solving the first-order optimality

conditions obtained from taking the Gateaux derivative in (2.10):

$$\begin{aligned} \text{find } \phi^h \in \mathcal{V}_{D;p,\alpha}^h \text{ such that for all } w^h \in \mathcal{V}_{0;p,\alpha}^h \\ (\mathcal{P}_{\mathcal{H}}\phi - \phi, w^h)_{\mathcal{H}} = (\phi^h - \phi, w^h)_{\mathcal{H}} = 0. \end{aligned} \quad (2.12)$$

For approximation spaces $\mathcal{V}_{p,\alpha}^h$ consisting of continuous functions (i.e. $\alpha \geq 0$), we introduce the L^2 - and H_0^1 -orthogonal projectors, respectively, as

$$(\mathcal{P}_{L^2}\phi - \phi, w^h)_{\Omega} = (\phi^h - \phi, w^h)_{\Omega} = 0, \quad (2.13a)$$

$$(\mathcal{P}_{H_0^1}\phi - \phi, w^h)_{H_0^1(\Omega)} = (\phi^h - \phi, w^h)_{H_0^1(\Omega)} = 0. \quad (2.13b)$$

Next, we consider approximation spaces $\mathcal{V}_{p,\alpha}^h$ consisting of discontinuous functions (i.e. $\alpha = -1$). In this context, the standard H_0^1 -norm is not a suitable best approximation problem. In order to introduce a suitable alternative to the H_0^1 -best approximation problem, we first introduce some additional notation. We define the union of (n_{el}) open element domains and the associated interface skeleton as

$$\tilde{\Omega} = \bigcup_{i=1}^{n_{\text{el}}} K_i, \quad (2.14a)$$

$$\Gamma = \bigcup_{i=1}^{n_{\text{el}}} \partial K_i, \quad (2.14b)$$

and introduce $\Gamma^0 = \Gamma \setminus \partial\Omega$ as the interior part of the interface skeleton. Next, we introduce some trace operators that are convenient in the context of discontinuous basis functions. For an interior edge e , shared by elements K^+ and K^- , we define the outward pointing unit normal vectors on e as $\mathbf{n}^+$ and $\mathbf{n}^-$, respectively. Denoting $\phi^+ = \phi|_{\partial K^+}$ and $\phi^- = \phi|_{\partial K^-}$ of a scalar quantity ϕ , we define the average $\{\!\{\phi\}\!\}$ and jump $\llbracket\phi\rrbracket$ on Γ^0 as

$$\{\!\{\phi\}\!\} = \frac{1}{2}(\phi^+ + \phi^-), \quad (2.15a)$$

$$\llbracket\phi\rrbracket = \phi^+ \mathbf{n}^+ + \phi^- \mathbf{n}^-. \quad (2.15b)$$

For a vector-valued quantity ψ on e we define ψ^+ and ψ^- analogously and introduce the average $\{\!\{\psi\}\!\}$ on Γ^0 as

$$\{\!\{\psi\}\!\} = \frac{1}{2}(\psi^+ + \psi^-). \quad (2.16)$$

We do not require the jump of a vector quantity and leave it undefined.

We now provide an alternative to the H_0^1 -best approximation problem that is well-posed. This alternative is closely related the well-known interior penalty method. This method requires the evaluation of boundary flux $\partial_n \phi$ on Γ^0 , which is an unbounded operator and yields a double-valued function for $\phi \in \mathcal{V} = H^1(\Omega)$. To circumvent this issue we first introduce the broken space

$$H^1(\tilde{\Omega}) = \{\phi \in L^2(\Omega) : \phi|_K \in H^1(K) \text{ for all } K \in \mathcal{T}\}, \quad (2.17)$$

and consider solution functions $\phi \in \tilde{\mathcal{V}} = H^1(\tilde{\Omega})$. To mitigate the issue of the unbounded boundary flux operator, we introduce a suitable additional function space. Recalling that the boundary flux is double-valued, we introduce the function space of the boundary fluxes as the product space $\mathcal{Q} \times \mathcal{Q}$, where $\mathcal{Q} = H^{-1/2}(\Gamma^0)$. Consider then the following operator:

$$\begin{aligned} \mathcal{P}_{\text{IP}} : \tilde{\mathcal{V}} \times \mathcal{Q} \times \mathcal{Q} &\longrightarrow \text{ran } \mathcal{P}_{\text{IP}}, \\ (\phi, \mu^+, \mu^-) &\longrightarrow (\phi^h, \partial_n \phi^{+,h}, \partial_n \phi^{-,h}), \end{aligned} \quad (2.18)$$

with

$$\begin{aligned} \phi^h &= \underset{\theta^h \in \mathcal{V}_{D,p,\alpha}^h}{\text{arginf}} \frac{1}{2} \|\phi - \theta^h\|_{H_0^1(\tilde{\Omega})}^2 - \langle \llbracket \mu \mathbf{n} - \nabla \theta^h \rrbracket, \llbracket \phi - \theta^h \rrbracket \rangle_{\Gamma^0} \\ &\quad + \frac{1}{2} \langle \eta \llbracket \phi - \theta^h \rrbracket, \llbracket \phi - \theta^h \rrbracket \rangle_{\Gamma^0}. \end{aligned} \quad (2.19)$$

The range of $\mathcal{P}_{\text{IP}}$ is given by

$$\text{ran } \mathcal{P}_{\text{IP}} = \{(w^h, \partial_n w^{+,h}, \partial_n w^{-,h}) : w^h \in \mathcal{V}_{D,p,\alpha}^h\}, \quad (2.20)$$

with dimension $\dim(\text{ran } \mathcal{P}_{\text{IP}}) = \dim \mathcal{V}_{D,p,\alpha}^h$. Additionally, the mapping $\mathcal{P}_{\text{IP}}$ is idempotent, and is a linear and bounded operator on the space $\tilde{\mathcal{V}} \times \mathcal{Q} \times \mathcal{Q}$. As a consequence, $\mathcal{P}_{\text{IP}}$ is a projector, and we refer to it as the *interior penalty projector*.^{37, 38} The penalty parameter η penalizes mismatches of interface jumps. Well-posedness is ensured when the penalty parameter satisfies a certain lower bound. In this work we base the value of the penalty parameter η on the work of Shahbazi.³⁶

Remark 2.1. (Interpretation boundary flux) The double-valued quantity μ acts as a surrogate to $\partial_n \phi$ in (2.19), and is introduced to make the projector a bounded operator. In practice, we simply use $\nabla \phi$ in place of $\mu \mathbf{n}$ in (2.19). To clarify the consistency of this replacement, we expand a part of the second integrand in (2.19)

$$\begin{aligned} \llbracket \mu \mathbf{n} \rrbracket \cdot \llbracket \phi - \phi^h \rrbracket &= \frac{1}{2} (\mu^+ (\phi - \phi^h)^+ - \mu^+ (\phi - \phi^h)^- \\ &\quad - \mu^- (\phi - \phi^h)^+ + \mu^- (\phi - \phi^h)^-). \end{aligned} \quad (2.21)$$

By then replacing μ by $\partial_n \phi$, we arrive at

$$\begin{aligned} \llbracket \partial_n \phi \mathbf{n} \rrbracket \cdot \llbracket \phi - \phi^h \rrbracket &= \frac{1}{2} (\mathbf{n}^+ \cdot \nabla \phi^+ (\phi - \phi^h)^+ - \mathbf{n}^+ \cdot \nabla \phi^+ (\phi - \phi^h)^- \\ &\quad - \mathbf{n}^- \cdot \nabla \phi^- (\phi - \phi^h)^+ + \mathbf{n}^- \cdot \nabla \phi^- (\phi - \phi^h)^-) \\ &= \llbracket \nabla \phi \rrbracket \cdot \llbracket \phi - \phi^h \rrbracket, \end{aligned} \quad (2.22)$$

which reveals the connection between $\mu \mathbf{n}$ and $\nabla \phi$.

To employ the interior penalty projector for the best approximation of sufficiently smooth functions, we may replace $\mu \mathbf{n}$ by $\nabla \phi$ (see Remark 2.1). By taking the

Gateaux derivative of (2.19), and using this substitution, we obtain the following first-order optimality condition:

$$\begin{aligned} \text{find } \phi^h \in \mathcal{V}_{D;p,0}^h \text{ such that for all } w^h \in \mathcal{V}_{0;p,0}^h \\ (\phi^h - \phi, w^h)_{H_0^1(\tilde{\Omega})} - \langle \llbracket \phi^h - \phi \rrbracket, \{\!\{ \nabla w^h \}\!\} \rangle_{\Gamma^0} - \langle \{\!\{ \nabla \phi^h - \nabla \phi \}\!\}, \llbracket w^h \rrbracket \rangle_{\Gamma^0} \\ + \langle \eta \llbracket \phi^h - \phi \rrbracket, \llbracket w^h \rrbracket \rangle_{\Gamma^0} = 0. \end{aligned} \quad (2.23)$$

With this substitution it is easy to see that the interior penalty projector is indeed associated with a best approximation problem

$$\phi^h = \operatorname{arginf}_{\theta^h \in \mathcal{V}_{D;p,-1}^h} \|\phi - \theta^h\|_{\text{IP}(\Omega)}, \quad (2.24)$$

where the norm is defined as

$$\|v\|_{\text{IP}(\Omega)}^2 := \|v\|_{H_0^1(\tilde{\Omega})}^2 - 2\langle \{\!\{ \nabla v \}\!\}, \llbracket v \rrbracket \rangle_{\Gamma^0} + \langle \eta \llbracket v \rrbracket, \llbracket v \rrbracket \rangle_{\Gamma^0}. \quad (2.25)$$

3. An Exposition of the Gibbs Phenomenon for Best Approximations in Finite Element Spaces

In this section we demonstrate the occurrence of the Gibbs phenomenon in best approximation problems that involve finite element approximation spaces of arbitrary continuity and degree. For simplicity, we work with B-splines with equal knot spacing. We consider the one-dimensional case in Sec. 3.1 and the two-dimensional case in Sec. 3.2.

3.1. The Gibbs phenomenon in one dimension

We consider best approximations of a step function $\phi \in L^2(\Omega)$ defined as

$$\phi = \phi_a(x) = \begin{cases} 1 & x > a, \\ -1 & x < a, \end{cases} \quad (3.1)$$

where a denotes the location of the jump discontinuity. As some best approximation statements involve weak derivatives, we wish to work with solution functions in $H^1(\Omega)$. Therefore, we introduce the following smooth (differentiable) approximation $\phi \in H^1(\Omega) =: \mathcal{V}$ of the step function:

$$\phi = \phi_a^\epsilon(x) = \tanh\left(\frac{x-a}{\epsilon}\right), \quad (3.2)$$

where $\epsilon \ll 1$ is a smoothing parameter.

We start off with the case in which the approximation space $\mathcal{V}_{p,\alpha}^h$ consists of continuous functions, i.e. $\alpha \geq 0$. The L^2 -best approximation $\phi^h \in \mathcal{V}_{D;1,0}^h$ (the space spanned by continuous piecewise linear basis functions) of the smooth step function is illustrated in Fig. 1. We observe that the numerical approximation ϕ^h contains over- and undershoots near the sharp layer. These oscillations do not vanish when

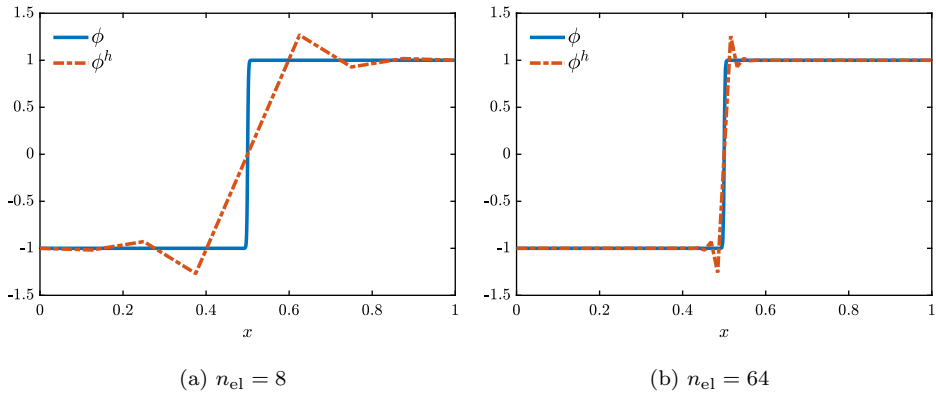


Fig. 1. The L^2 -best approximation $\phi^h \in \mathcal{V}_{D;1,0}^h$ of the smooth step function $\phi = \phi_{0.5}$.

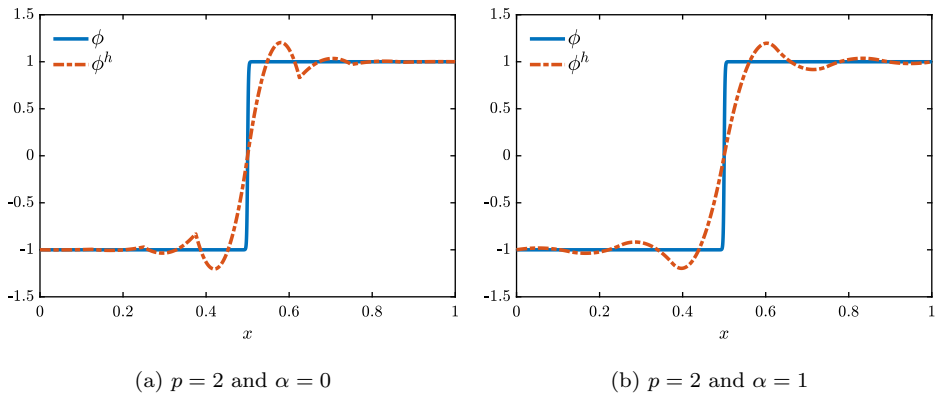


Fig. 2. The L^2 -best approximations $\phi^h \in \mathcal{V}_{D;2,0}^h$ (left) and $\phi^h \in \mathcal{V}_{D;2,1}^h$ (right), both with $n_{\text{el}} = 8$, of the smooth step function $\phi = \phi_{0.5}$.

the number of elements is increased. In fact, the over- and undershoots on each side of the discontinuity converge to the value $1 - \sqrt{3}/2 \approx 0.13$ as the number of elements is increased (assuming that the layer is “sufficiently sharp”).⁹ We note that the Gibbs phenomenon is often mistakenly interpreted as related to approximation with higher-order basis functions. This example illustrates that this is *not* the case.

Figure 2 shows the approximations in $\mathcal{V}_{D;2,0}^h$ and $\mathcal{V}_{D;2,1}^h$, the spaces of continuous quadratic finite elements. The figure shows over- and undershoots of roughly the same magnitude as those for the linear approximation. For the quadratic B-splines the over- and undershoots on each side of the discontinuity converge with the number of elements to a value of approximately 0.10.³⁵ The Gibbs phenomenon persists when the polynomial order p of the maximum regularity B-spline basis functions is increased. Moreover, the magnitude of the over- and undershoots converges to the same value as that of a truncated Fourier series. This value is approximately 0.09 and is known as the *Gibbs constant*.

Remark 3.1. (Different degrees of freedom) It is important to realize that the number of degrees of freedom (dofs) is significantly different for results with the same number of elements and polynomial degree p but with different regularity α . In this situation, we have 21 dofs for $\phi^h \in \mathcal{V}_{D;2,0}^h$ and only 10 dofs for $\phi^h \in \mathcal{V}_{D;2,1}^h$ (the count includes boundary dofs).

In Fig. 3, we visualize the H_0^1 -best approximation $\phi^h \in \mathcal{V}_{D;1,0}^h$ (the space spanned by continuous piecewise linear basis functions) for $a = 0.5$ and $a = 0.58$. We observe nodally exact numerical approximations for both cases. The combination of the linear basis functions with the nodal exactness implies that the numerical approximations are free of over- and undershoots, i.e. the Gibbs phenomenon is not present.

In Fig. 4, we illustrate the H_0^1 -best approximations for the approximation spaces of quadratic finite elements $\mathcal{V}_{D;2,0}^h$ and $\mathcal{V}_{D;2,1}^h$. We observe over- and undershoots for both the quadratic Lagrange polynomials and the B-spline functions. Again, these oscillations persist with mesh refinement. For the case of the Lagrange basis functions, we have nodal exactness at element boundary nodes, but the monotonicity property is lost in the element interiors.²⁵

Lemma 3.1. (Nodal interpolant) *The H_0^1 -best approximation in the space $\mathcal{V}_{D;p,0}^h$ in one dimension is nodally interpolatory at the element boundary nodes. For linear elements ($p = 1$) this best approximation is monotonicity preserving, while for higher-order basis functions ($p > 1$) monotonicity inside the elements is in general lost.*

Next, we turn our attention to discontinuous approximations, i.e. $\alpha = -1$. We select as penalty parameter $\eta = 6(p+1)^2/h$. The interior penalty-best approximation of the smooth step function is illustrated in Fig. 5 for the approximation spaces $\mathcal{V}_{D;1,-1}^h$ (discontinuous piecewise linears) and $\mathcal{V}_{D;2,-1}^h$ (discontinuous

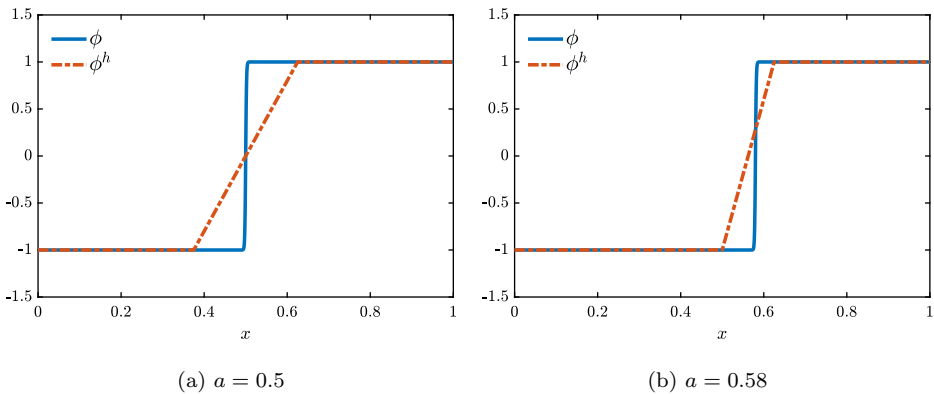


Fig. 3. The H_0^1 -best approximations $\phi^h \in \mathcal{V}_{D;1,0}^h$, with $n_{\text{el}} = 8$, of the smooth step function $\phi = \phi_a$ for different a .

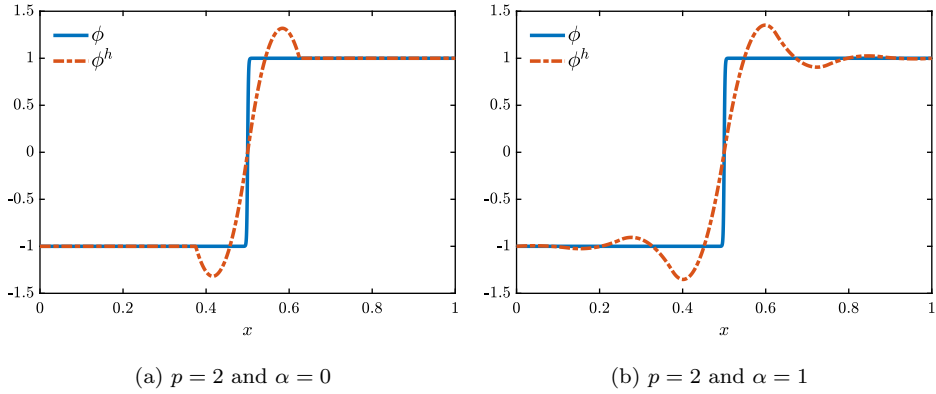


Fig. 4. The H_0^1 -best approximation $\phi^h \in \mathcal{V}_{D;2,0}^h$ (left) and $\phi^h \in \mathcal{V}_{D;2,1}^h$ (right), with $n_{\text{el}} = 8$, of the smooth step function $\phi = \phi_{0.5}$.

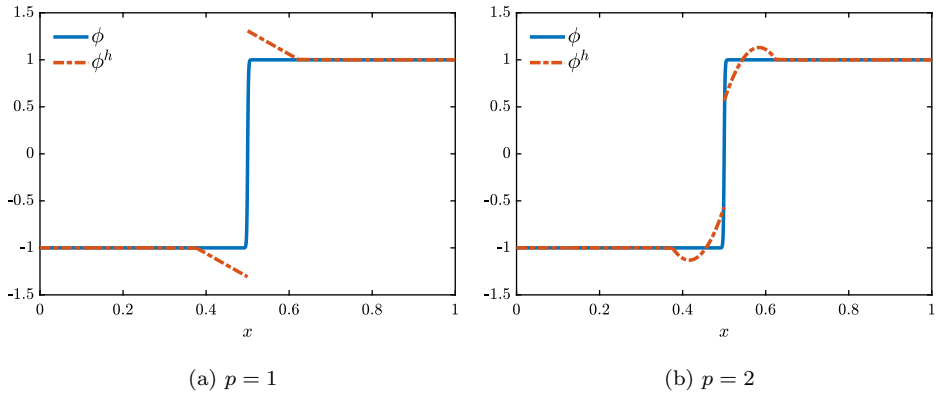


Fig. 5. The interior penalty-best approximation $\phi^h \in \mathcal{V}_{D;p,-1}^h$, with $n_{\text{el}} = 8$, of the smooth step function $\phi = \phi_{0.5}$, for different p .

piecewise quadratics). For this best approximation problem, we see that the numerical approximations ϕ^h contain over- and undershoots near the sharp layer; the nodal exactness of the H_0^1 -best approximation problem is not inherited. Furthermore, we observe that the average of the approximation ϕ^h at the element boundaries coincides with the value of ϕ . This is a property of the interior penalty projector.³⁸

Proposition 3.1. (Vanishing average error on element boundaries) *The interior penalty-best approximation $\phi^h \in \mathcal{V}_{D,p,-1}^h$ of ϕ in one-dimension satisfies the property*

$$\{\!\!\{\phi^h - \phi\}\!\!\}_{|\Gamma^0} = 0, \quad (3.3)$$

for all polynomial orders p .

3.2. The Gibbs phenomenon in two dimensions

We consider best approximations of a two-dimensional step function $\phi \in L^2(\Omega)$ on the square domain $\Omega = (-1, 1)^2$

$$\phi = \phi(x, y) = \begin{cases} 1 & x - y > 0, \\ -1 & x - y < 0. \end{cases}$$

Again, we work with a smooth approximation $\phi \in H^1(\Omega) =: \mathcal{V}$ of the step function

$$\phi = \phi^\epsilon(x, y) = \tanh\left(\frac{y - x}{\epsilon}\right),$$

where $\epsilon \ll 1$ is a smoothing parameter.

Analogous to the one-dimensional case, we begin with approximation spaces $\mathcal{V}_{D,p,\alpha}^h$ consisting of continuous functions ($\alpha \geq 0$). Recall from the one-dimensional case that the L^2 -best approximation contains over- and undershoots. This is also the case in higher dimensions, and we omit the visualization. We display the H_0^1 -best approximation for the continuous finite element approximation spaces $\mathcal{V}_{D,1,0}^h$, $\mathcal{V}_{D,2,0}^h$ and $\mathcal{V}_{D,2,1}^h$ in Fig. 6. We observe the occurrence of over- and undershoots for each of the approximations.

Next, we consider discontinuous approximations ($\alpha = -1$). Analogous to the one-dimensional case, we solely consider the interior penalty-best approximation. We select as a penalty parameter $\eta = 2(2p + 1)(2p + 2)/h$. In Fig. 7, we visualize the interior penalty-best approximations for approximation spaces $\mathcal{V}_{1,-1}^h$ and $\mathcal{V}_{2,-1}^h$, i.e. for linear and quadratic discontinuous basis functions. Again, we observe over- and undershoots of the finite element approximation in both cases.

4. Eliminating the Gibbs Phenomenon in One Dimension

In this section, we present the constraints for the elimination of the Gibbs phenomenon in finite element spaces in one dimension. We call these constraints the *Gibbs constraints*. To this purpose, we first describe the construction of our proposed Gibbs constraints in general approximation spaces in Sec. 4.1, and present the properties of the constraints in Sec. 4.2. Next, we advance the discussion to best approximation problems in Sec. 4.3. Finally, in Sec. 4.4, we apply the Gibbs constraints to finite element spaces of arbitrary continuity, and perform numerical experiments.

4.1. Gibbs constraints

One of the challenges of dealing with the Gibbs phenomenon is the uncertainty in the level of locality that is required to identify the phenomenon. Pointwise evaluations of functions, or their derivatives, carry insufficient information to be able to infer the occurrence of the Gibbs phenomenon. On the other hand, the information that may be deduced from global evaluations, such as global integrals, is too

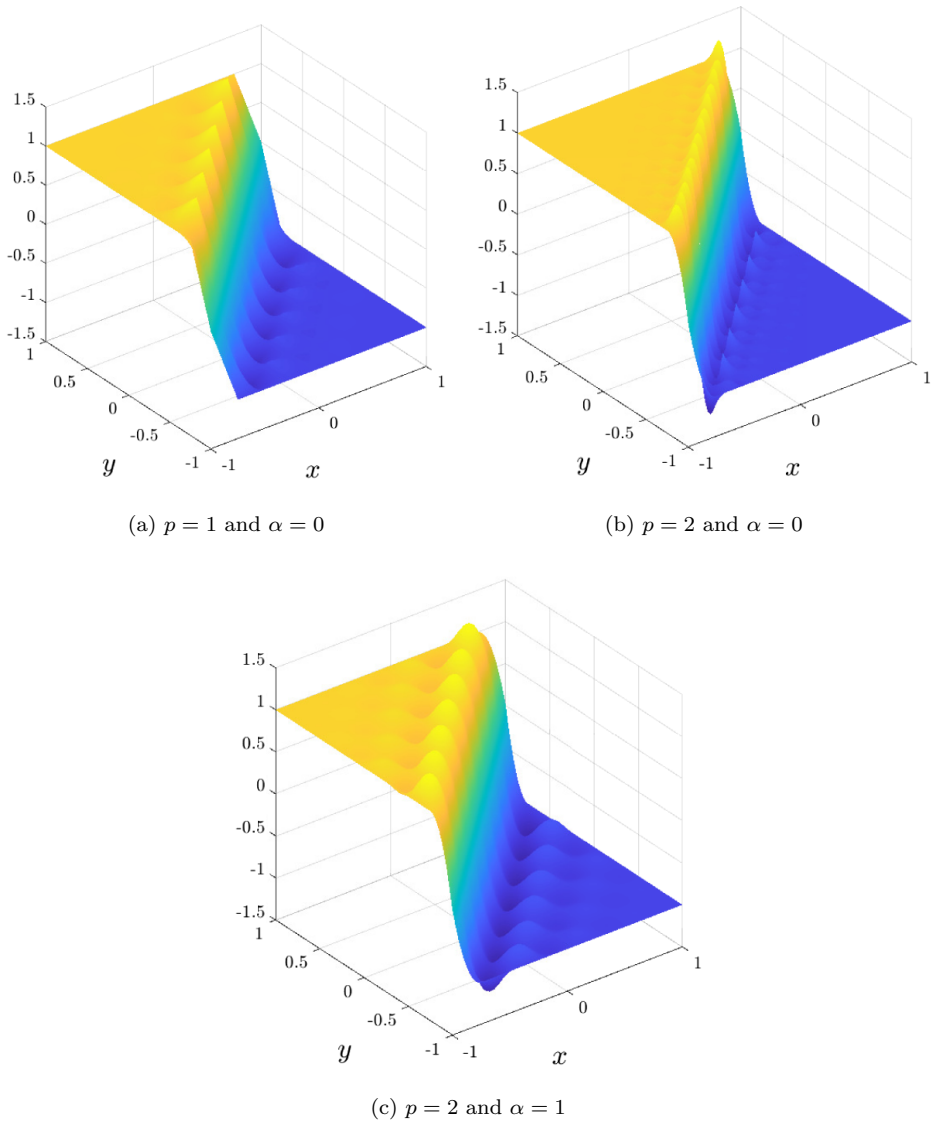


Fig. 6. The H_0^1 -best approximation $\phi^h \in \mathcal{V}_{D;p,\alpha}^h$, with $n_{\text{el}} = 8 \times 8$, of the smooth two-dimensional step function, for different p and α .

coarse-grained to establish the existence of spurious oscillations on a local scale. In this section, we construct a constraint for the elimination of the Gibbs phenomenon on a *given subdomain*. The selection of the subdomains then remains an important matter, which we discuss extensively in Sec. 4.4 in the context of finite element approximations. We first introduce some well-known but important theories associated with the total variation concept. The constraint is specified in Definitions 4.2 and 4.3.

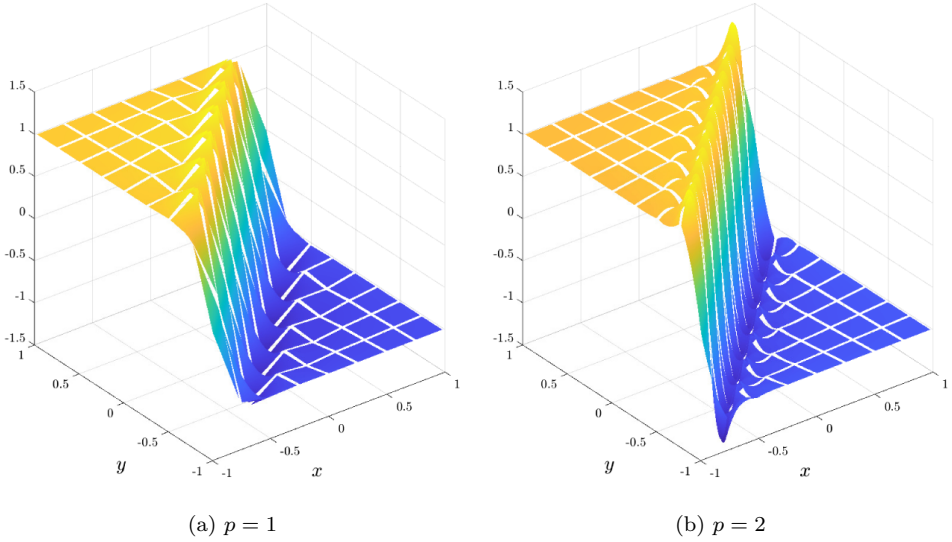


Fig. 7. The interior penalty-best approximation $\phi^h \in \mathcal{V}_{D;p,-1}^h$, with $n_{\text{el}} = 8 \times 8$, of the smooth two-dimensional step function, for different p .

Consider the one-dimensional simply connected domain $\Omega \subset \mathbb{R}$, and let $\phi^* : \Omega \rightarrow \mathbb{R}$ denote an approximation of $\phi : \Omega \rightarrow \mathbb{R}$. The main ingredient in the elimination of the Gibbs phenomenon relies on the *fundamental theorem of Lebesgue integral calculus*, which we recall here.

Theorem 4.1. (Fundamental theorem of Lebesgue integral calculus) *Let $\theta : \Omega \rightarrow \mathbb{R}$ be an absolutely continuous function, then θ is differentiable almost everywhere and for each $\omega = [x_L, x_R] \subset \Omega$ we have*

$$\int_{x_L}^{x_R} D\theta \, dx = \theta_R - \theta_L, \quad (4.1)$$

with trace equalities $\theta(x_L) = \theta_L$ and $\theta(x_R) = \theta_R$, and $D\theta \in L^1(\Omega)$.

The fundamental theorem communicates that the trace values (the right-hand side in (4.1)) are controlled by the integral. Still, the trace values do not provide any information on the oscillatory behavior of the function θ inside ω . In contrast, the *total variation* is a concept that does incorporate this.

Definition 4.1. (Total variation) Let $\theta : \Omega \rightarrow \mathbb{R}$ be a given function, and let $P = \{x_L = x_0, x_1, \dots, x_{N-1}, x_N = x_R\}$ denote a partition of $\omega = [x_L, x_R] \subset \Omega$. The variation of θ with respect to partition P is defined as

$$V_{\omega,P}(\theta) := \sum_{i=0}^N |\theta(x_{i+1}) - \theta(x_i)|. \quad (4.2)$$

Denote by $\mathfrak{P}$ the set of all possible partitions of ω . The total variation is given by

$$V_\omega(\theta) := \sup_{P \in \mathfrak{P}} V_{\omega,P}(\theta). \quad (4.3)$$

The total variation $V_\omega(\theta)$ represents a measure of the fluctuations of θ on ω . A function θ with the property $V_\omega(\theta) < \infty$ is said to have *bounded variation* and we write $\theta \in \text{BV}(\omega)$. An absolutely continuous function has bounded variation. In case θ is a continuous differentiable function, the total variation $V_\omega(\theta)$ may be evaluated as follows.

Lemma 4.1. (Total variation continuous differentiable function) *Let $\theta \in \mathcal{C}^1(\Omega)$. The total variation of θ on $\omega \subset \Omega$ is given by*

$$V_\omega(\theta) = \int_\omega |D\theta| dx. \quad (4.4)$$

If the function θ is only piecewise continuously differentiable, the expression in Lemma 4.1 needs to be augmented with jump terms, as is expressed in Lemma 4.2.

Lemma 4.2. (Total variation piecewise-continuous function) *Let $\theta \in \text{BV}(\Omega)$ be a function that has a continuous derivative on each $(a_i, a_{i+1}) \subset \Omega, i = 0, \dots, M$ and jump discontinuities at $a_i, i = 1, \dots, M$. Denote with a_i^-, a_i^+ the left and right limits of the discontinuity at a_i . The total variation of θ on $\omega \subset \Omega$ is given by*

$$V_\omega(\theta) = \sum_{i=0}^M \int_{a_i}^{a_{i+1}} |D\theta| dx + \sum_{i=1}^M |\theta(a_i) - \theta(a_i^-)| + |\theta(a_i^+) - \theta(a_i)|. \quad (4.5)$$

Next, we note that V_ω is convex and satisfies a homogeneity property.

Proposition 4.1. (Convexity and homogeneity of the total variation) *Given the same assumptions on θ as in Lemma 4.2, the functional V_ω is convex and satisfies the homogeneity property*

$$dV_\omega(\theta)(\theta) = V_\omega(\theta). \quad (4.6)$$

Remark 4.1. A convex functional that satisfies the homogeneity property is termed a *variation entropy*.⁴¹ Variation entropy functionals form the basis of an entropy stability theory for (hyperbolic) conservation laws called variation entropy theory.

The total variation of a monotonic function is related to its trace values in the following way.

Lemma 4.3. (Total variation monotonic function) *Let $\omega = [x_L, x_R] \subset \Omega$ and let $\mathcal{C}_D^0(\Omega)$ denote the space of continuous functions θ that satisfy trace equalities $\theta(x_L) = \theta_L$ and $\theta(x_R) = \theta_R$. For $\theta \in \mathcal{C}_D^0(\Omega)$ we have*

$$V_\omega(\theta) \geq |\theta_R - \theta_L|, \quad (4.7)$$

where equality only holds when θ is monotonic.

This is an important ingredient in the design of the Gibbs constraints. It communicates that the jump of the trace values are controlled by the total variation.

Let us introduce the quantity

$$\mathcal{V}_{\phi,\omega}(\phi^*) := V_\omega(\phi^*) - V_\omega(\phi). \quad (4.8)$$

With the aim of eliminating the Gibbs phenomenon, one could introduce the constraint $\mathcal{V}_{\phi,\omega}(\phi^*) \leq 0$. The disadvantage is that total variation does not take into account the sign of its argument

$$V_\omega(-\theta) = V_\omega(\theta). \quad (4.9)$$

As a consequence, we have

$$\mathcal{V}_{\phi,\omega}(-\phi) = 0. \quad (4.10)$$

We now generalize the element-based definition of the subdivision of domain Ω , (2.14a), to the union of (J) disjoint general subdomains, i.e.

$$\tilde{\Omega} := \bigcup_{j=1}^J \omega_j, \quad (4.11)$$

with $\omega_j \cap \omega_k = \emptyset$ for $j \neq k$. Furthermore, we redefine the broken space (2.17) for this subdivision as

$$H^1(\tilde{\Omega}) := \{v \in L^2(\Omega) : v|_\omega \in H^1(\omega) \text{ for all } \omega \in \mathcal{T}_\omega\}, \quad (4.12)$$

with $\mathcal{T}_\omega$ the collection of $\omega_j, j = 1, \dots, J$.

Motivated by the above observation, we wish to find approximations that, besides bounding the size of the jump of the approximation, also carry information about the direction of the analytical solution ϕ . To this purpose we now propose the *Gibbs functional* and the associated *Gibbs constraint*.

Definition 4.2. (Gibbs functional) The Gibbs functional of the function $\phi^* \in H^1(\tilde{\Omega})$, on $\omega = \omega_j$ (for some j) and with respect to the given function $\phi \in H^1(\Omega)$, is defined as

$$\mathcal{G}_{\phi,\omega}(\phi^*) := \int_\omega g_\phi(\phi^*) dx, \quad (4.13)$$

where the functional g_ϕ is defined as

$$g_\phi(\phi^*) := |D\phi^*| - \text{sgn}(D\phi^*)D\phi = -\text{sgn}(D\phi^*)D\phi'. \quad (4.14)$$

Here, sgn is the sign function (i.e. $\text{sgn}(t) = t/|t|$ for $t \neq 0$ and $\text{sgn}(0) = 0$), and $\phi' : \Omega \rightarrow \mathbb{R}$ defined as $\phi' := \phi - \phi^*$ denotes the error function.

Definition 4.3. (Gibbs constraint) The Gibbs constraint of the function $\phi^* \in H^1(\tilde{\Omega})$, on ω and with respect to the given function $\phi \in H^1(\Omega)$, is defined as

$$\mathcal{G}_{\phi,\omega}(\phi^*) \leq 0. \quad (4.15)$$

Note that the incorporation of the correct sign may be recognized in the Gibbs constraint via the equivalence

$$\left. \begin{aligned} |\mathrm{D}\phi^*| - |\mathrm{D}\phi| &\leq 0 \\ \mathrm{sgn}(\mathrm{D}\phi^*) - \mathrm{sgn}(\mathrm{D}\phi) &= 0 \end{aligned} \right\} \Leftrightarrow g_\phi(\phi^*) \leq 0, \quad (4.16)$$

the integration of which provides the Gibbs constraint.

The objective is now to search for functions ϕ^* as approximations of ϕ that satisfy the Gibbs constraint on certain ω .

4.2. Properties of the Gibbs constraint

It is the purpose of this section to discuss the properties of the Gibbs constraint and to establish its connection with the well-known concepts of monotonic solutions and the maximum principle.

We have the simple but important property that ϕ as an approximation of itself satisfies the Gibbs constraint.

Proposition 4.2. (Perfect approximation) *The Gibbs functional vanishes for a perfect approximation ($\phi^* = \phi$)*

$$\mathcal{G}_{\phi,\omega}(\phi) = 0. \quad (4.17)$$

Furthermore, we have the following lower bound of the Gibbs functional.

Lemma 4.4. (Gibbs functional bound) *The Gibbs functional satisfies the lower bound*

$$\mathcal{G}_{\phi,\omega}(\phi^*) \geq \mathcal{V}_{\phi,\omega}(\phi^*). \quad (4.18)$$

We now proceed with establishing connections between the Gibbs constraint and certain properties of the function approximation. We first provide a characterization of the Gibbs functional.

Lemma 4.5. (Characterization Gibbs functional) *Let $\phi^* \in H^1(\tilde{\Omega})$ and $\phi \in H^1(\Omega)$, and let $\omega = [x_L, x_R] \subset \Omega$ be given. Denote the locations of sign changes of $\mathrm{D}\phi^*$ by $x_i, i = 1, \dots, N$ with $x_i < x_{i+1}$. The form of Gibbs functional $\mathcal{G}_{\phi,\omega}$ depends on the sign of $\mathrm{D}\phi^*$ on $[x_L, x_1]$ and the number of sign changes N*

(1) *for N odd and $\mathrm{D}\phi^* \geq 0$ on $[x_L, x_1]$*

$$\mathcal{G}_{\phi,\omega}(\phi^*) = \phi'(x_L) + 2 \sum_{i=1}^N (-1)^i \phi'(x_i) + \phi'(x_R), \quad (4.19)$$

(2) *for N odd and $\mathrm{D}\phi^* \leq 0$ on $[x_L, x_1]$*

$$\mathcal{G}_{\phi,\omega}(\phi^*) = -\phi'(x_L) - 2 \sum_{i=1}^N (-1)^i \phi'(x_i) - \phi'(x_R), \quad (4.20)$$

(3) for N even and $D\phi^* \geq 0$ on $[x_L, x_1]$

$$\mathcal{G}_{\phi, \omega}(\phi^*) = \phi'(x_L) + 2 \sum_{i=1}^N (-1)^i \phi'(x_i) - \phi'(x_R), \quad (4.21)$$

(4) for N even and $D\phi^* \leq 0$ on $[x_L, x_1]$

$$\mathcal{G}_{\phi, \omega}(\phi^*) = -\phi'(x_L) - 2 \sum_{i=1}^N (-1)^i \phi'(x_i) + \phi'(x_R), \quad (4.22)$$

where we recall $\phi' = \phi - \phi^*$.

A direct consequence of this characterization is the following lemma.

Lemma 4.6. (Interpolatory monotonic approximation) *An interpolatory monotonic approximation $\phi^* \in H^1(\tilde{\Omega})$ of $\phi \in H^1(\Omega)$ on $\omega \subset \Omega$ satisfies the Gibbs constraint*

$$\mathcal{G}_{\phi, \omega}(\phi^*) \leq 0. \quad (4.23)$$

Additionally, an approximation of a monotonic analytical profile that is free of the Gibbs phenomenon satisfies a bound on the trace values.

Lemma 4.7. (Approximation of monotonic function) *Suppose that $\phi \in H^1(\Omega)$ is monotonically increasing (decreasing) on $\omega = [x_L, x_R] \subset \Omega$ and the approximation $\phi^* \in H^1(\tilde{\Omega})$ satisfies the Gibbs constraint*

$$\mathcal{G}_{\phi, \omega}(\phi^*) \leq 0, \quad (4.24)$$

then the increase (decrease) of ϕ^ is bounded by the increase (decrease) of ϕ , i.e.*

$$\phi^*(x_R) - \phi^*(x_L) \leq \phi(x_R) - \phi(x_L) \quad (\text{if } \phi \text{ is monotonically increasing}), \quad (4.25a)$$

$$\phi^*(x_L) - \phi^*(x_R) \leq \phi(x_L) - \phi(x_R) \quad (\text{if } \phi \text{ is monotonically decreasing}). \quad (4.25b)$$

The equality in (4.25) holds when (4.24) holds with equality.

Proof. We omit the proof of the general case and consider instead two simple cases. Without loss of generality, assume that ϕ is monotonically increasing. Suppose first that $D\phi^*$ has no sign changes, i.e. ϕ^* is monotonically increasing on $\omega = [x_L, x_R]$. Then, (4.21) reduces to

$$\mathcal{G}_{\phi, \omega}(\phi^*) = (\phi^*(x_R) - \phi^*(x_L)) - (\phi(x_R) - \phi(x_L)), \quad (4.26)$$

which is negative if and only if (4.25a) holds. Now, suppose that $D\phi^* > 0$ has a single change of sign, say at x_1 , and that $D\phi^* > 0$ for $x < x_1$. Additionally, since we may shift ϕ^* by $\phi'(x_L)$, we take $\phi'(x_L) = 0$. It is easy to verify that $\phi^*(x_1) \leq \phi(x_R)$, as otherwise $\mathcal{G}_{\phi, \omega}(\phi^*) \geq 0$. Since ϕ^* is decreasing on $[x_1, x_R]$ and ϕ is increasing, we immediately have $\phi'(x_R) \geq 0$. The case $D\phi^* < 0$ for $x < x_1$ follows from a similar argument. $\square$

Next, we introduce the classical definition of the maximum principle.

Definition 4.4. (Maximum principle) An approximation $\phi^* : \Omega \rightarrow \mathbb{R}$ of $\phi : \Omega \rightarrow \mathbb{R}$ satisfies the maximum principle on $\omega \subset \Omega$ if and only if it does not exceed the bounds of ϕ on ω

$$\inf_{\omega} \phi^* \geq \inf_{\omega} \phi, \quad (4.27a)$$

$$\sup_{\omega} \phi^* \leq \sup_{\omega} \phi. \quad (4.27b)$$

Theorem 4.2. (Interpolatory approximation of monotonic function) Suppose that $\phi \in H^1(\Omega)$ is monotonic on $\omega = [x_L, x_R] \subset \Omega$ and $\phi^* \in H^1(\tilde{\Omega})$ is an interpolatory approximation, i.e. $\phi^*(x_L) = \phi(x_L)$ and $\phi^*(x_R) = \phi(x_R)$. We have the following results:

- (1) $\mathcal{G}_{\phi, \omega}(\phi^*) \geq 0$,
- (2) ϕ^* is a monotonic function if and only if $\mathcal{G}_{\phi, \omega}(\phi^*) = 0$,
- (3) if $\mathcal{G}_{\phi, \omega}(\phi^*) = 0$ then ϕ^* satisfies the maximum principle on ω .

Proof. (1) Without loss of generality, assume ϕ is increasing. In case $D\phi^*$ does not change sign, we invoke (4.26) of Lemma 4.7 and obtain $\mathcal{G}_{\phi, \omega}(\phi^*) = 0$. In the other case, denote the locations of sign changes of $D\phi^*$ as $x_i \in \omega, i = 1, \dots, N$ with $x_i \leq x_{i+1}$. Lemma 4.5 provides the following:

- for $D\phi^* \geq 0$ on $[x_L, x_1]$

$$\mathcal{G}_{\phi, \omega}(\phi^*) = 2 \sum_{i=1}^N (-1)^i \phi'(x_i), \quad (4.28)$$

- for $D\phi^* \leq 0$ on $[x_L, x_1]$

$$\mathcal{G}_{\phi, \omega}(\phi^*) = -2 \sum_{i=1}^N (-1)^i \phi'(x_i), \quad (4.29)$$

where we recall $\phi' = \phi - \phi^*$. To show nonnegativity of (4.28) and (4.29), one has to consider various cases. For example, for $N = 1$ we have $\mathcal{G} = 2|\phi'(x_1)| \geq 0$. We omit a detailed proof of the general case.

(2) “ $\Rightarrow$ ”: Suppose ϕ^* is a monotonic function. Without loss of generality, assume ϕ is increasing. Since ϕ^* is a monotonic function, the sums in (4.29) and (4.28) are empty and the expressions vanish.

“ $\Leftarrow$ ”: Suppose that $\mathcal{G}_{\phi, \omega}(\phi^*) = 0$. Lemma 4.4 implies $\mathcal{V}_{\phi, \omega}(\phi^*) \leq 0$. Without loss of generality, suppose that ϕ is monotonically increasing. We arrive at

$$V_{\omega}(\phi^*) \leq \phi(x_R) - \phi(x_L). \quad (4.30)$$

Since ϕ^* is interpolatory, Lemma 4.3 implies that ϕ^* is monotonically increasing.

(3) If $\mathcal{G}_{\phi,\omega}(\phi^*) = 0$ then ϕ^* is monotonic via Claim 2, and (4.25a) provides the maximum principle via

$$\sup_{\omega} \phi^* = \phi^*(x_R) = \phi(x_R) = \sup_{\omega} \phi, \quad (4.31a)$$

$$\inf_{\omega} \phi^* = \phi^*(x_L) = \phi(x_L) = \inf_{\omega} \phi. \quad (4.31b)$$

□

We have the following connection between the Gibbs constraints and the maximum principle.

Lemma 4.8. (Gibbs constraints and maximum principle) *Let the analytical profile $\phi^* \in H^1(\Omega)$ be monotonic on each $\omega_j, j = 1, \dots, J$, and let $\tilde{\phi}^* \in H^1(\tilde{\Omega})$ be an approximation that satisfies $\mathcal{G}_{\phi,\omega_j}(\tilde{\phi}^*) \leq 0, j = 1, \dots, J$, such that $\tilde{\phi}^*$ is interpolatory on $\partial\Omega$. Then, $\tilde{\phi}^*$ satisfies the maximum principle on Ω .*

In order to preclude the Gibbs phenomenon on the entire domain, the strategy is to require $\mathcal{G}_{\phi,\omega_j}(\phi^*) \leq 0$, for $j = 1, \dots, J$. We note that the practical applicability of this strategy relies on an appropriately chosen subdivision of Ω . In the subsequent section, we discuss the Gibbs constraints in the context of best approximation problems. We return to the domain subdivision problem in the context of finite elements in Sec. 4.4.

4.3. Best approximation problems under Gibbs constraints

Let now $\phi \in H^1(\Omega)$ be given and consider the best approximation problem

$$\phi^* = \operatorname{arginf}_{\theta^* \in \mathcal{K}} \|\phi - \theta^*\|_{\mathcal{H}}, \quad (4.32a)$$

where the feasible set is defined as

$$\mathcal{K} := \{\phi^* \in H^1(\tilde{\Omega}) : \mathcal{G}_{\phi,\omega_j}(\phi^*) \leq 0, j = 1, \dots, J\}. \quad (4.32b)$$

We note that, in general $\mathcal{K}$, has no strictly feasible point. For example, in the case $\phi \equiv 0$ in Ω , we have $\mathcal{G}_{\phi,\omega_j}(\phi^*) = 0$ for $j = 1, \dots, J$. In the following we exclude this trivial case.

The standard techniques to study best approximation problems are the gradient-based methods. We remark, however, that the Gateaux derivative of $\phi^* \rightarrow \mathcal{G}_{\phi,\omega_j}(\phi^*)$ does not exist due to the occurrence of the sign function. To permit the adoption of standard gradient methods, we regularize the nondifferentiable constraint function as follows. We introduce the parameter $\varepsilon \in \mathbb{R}$ and define the differentiable regularizations $|\cdot|_{\varepsilon} : \mathbb{R} \rightarrow \mathbb{R}_+$ and $\operatorname{sgn}_{\varepsilon} : \mathbb{R} \rightarrow \mathbb{R}$ as

$$|r|_{\varepsilon} := |r^2 + \varepsilon^2|^{1/2}, \quad (4.33a)$$

$$\operatorname{sgn}_{\varepsilon}(r) := r|r|_{\varepsilon}^{-1}, \quad (4.33b)$$

which satisfy the relation

$$D|r|_\varepsilon = \operatorname{sgn}_\varepsilon(r). \quad (4.34)$$

Next, we introduce the regularized functional

$$\mathcal{G}_{\phi,\omega}^\varepsilon(\phi^*) := \int_\omega g_\phi^\varepsilon(\phi^*) \, dx, \quad (4.35)$$

where the functional g_ϕ^ε is defined as

$$g_\phi^\varepsilon(\phi^*) := |D\phi^*|_\varepsilon - \operatorname{sgn}_\varepsilon(D\phi^*)D\phi = -\operatorname{sgn}_\varepsilon(D\phi^*)D\phi' + \varepsilon^2|D\phi^*|_\varepsilon^{-1}. \quad (4.36)$$

It is now clear that the regularized functional $\mathcal{G}_{\phi,\omega}^\varepsilon(\phi^*)$ is differentiable in the Gateaux sense. The Gateaux derivative of $\mathcal{G}_{\phi,\omega}^\varepsilon(\phi^*)$, denoted $d\mathcal{G}_{\phi,\omega}^\varepsilon(\phi^*)(w)$, is given by

$$d\mathcal{G}_{\phi,\omega}^\varepsilon(\phi^*)(w) = \int_{\omega_j} |D\phi^*|_\varepsilon^{-1} D\phi^* Dw - \varepsilon^2 |D\phi^*|_\varepsilon^{-3} D\phi Dw \, dx. \quad (4.37)$$

We have the following property regarding the convexity of the regularized Gibbs functional.

Proposition 4.3. (Convexity of Gibbs functional) *The functional g_ϕ^ε is quasi-convex*

$$g_{\phi,\omega}^\varepsilon(\zeta\phi_1^* + (1-\zeta)\phi_2^*) \leq \max\{g_{\phi,\omega}^\varepsilon(\phi_1^*), g^\varepsilon(\phi_2^*)\}, \quad (4.38)$$

for all $\phi_1^*, \phi_2^* \in H^1(\tilde{\Omega})$, $\zeta \in [0, 1]$. The functional $\mathcal{V}_{\phi,\omega}$ is convex, but $\mathcal{G}_{\phi,\omega}^\varepsilon$ is in general not (quasi-)convex.

We now consider the best approximation problem

$$\phi^* = \operatorname{arginf}_{\theta^* \in \mathcal{K}^\varepsilon} \|\phi - \theta^*\|_{\mathcal{H}}, \quad (4.39a)$$

where the regularized feasible set is defined as

$$\mathcal{K}^\varepsilon := \{\phi^* \in H^1(\tilde{\Omega}) : \mathcal{G}_{\phi,\omega_j}^\varepsilon(\phi^*) \leq 0, j = 1, \dots, J\}. \quad (4.39b)$$

A consequence of Proposition 4.3 is that the feasible set $\mathcal{K}^\varepsilon$ is not convex. This is a result of the occurrence of the (regularized) sign function in g_ϕ^ε .

Remark 4.2. The feasible set determined by (a regularized) functional $\mathcal{V}_{\phi,\omega}(\phi^*) \leq 0$, i.e.

$$\{\phi^* \in H^1(\tilde{\Omega}) : \mathcal{V}_{\phi,\omega_j}^\varepsilon(\phi^*) \leq 0, j = 1, \dots, J\}, \quad (4.40)$$

with

$$\mathcal{V}_{\phi,\omega_j}^\varepsilon(\phi^*) = \int_\omega |D\phi^*|_\varepsilon - |D\phi|_\varepsilon \, dx, \quad (4.41)$$

is convex. Furthermore, note that quasi-convexity of the functional is sufficient for the associated feasible set to be convex.

We now introduce the first-order optimality Karush–Kuhn–Tucker (KKT) conditions of the constrained best approximation problem (4.39). We note that the solution of the KKT conditions is only guaranteed to be a local optimum due to the lack of convexity of the feasible set $\mathcal{K}^\varepsilon$. Replacing $\mathcal{K}^\varepsilon$ by the feasible set from (4.40) would imply global optimality.

Theorem 4.3. (KKT conditions) *The function $\phi^* \in \mathcal{K}$ is a local optimum of the problem (4.39) if and only if there exist Lagrange multipliers $\lambda_j \in \mathbb{R}$ $[j=1, \dots, J]$ such that the following KKT conditions hold:*

Stationarity:

Find $\phi^* \in H^1(\tilde{\Omega})$, $\lambda_j \in \mathbb{R}$ for $j = 1, \dots, J$ such that

$$(\phi^* - \phi, w)_{\mathcal{H}} + \sum_{j=1}^J \lambda_j d\mathcal{G}_{\phi, \omega_j}^\varepsilon(\phi^*)(w) = 0 \quad \text{for all } w \in H^1(\tilde{\Omega}), \quad (4.42a)$$

Primal feasibility:

$$\mathcal{G}_{\phi, \omega_j}^\varepsilon(\phi^*) \leq 0 \quad \text{for } j = 1, \dots, J, \quad (4.42b)$$

Dual feasibility:

$$\lambda_j \geq 0 \quad \text{for } j = 1, \dots, J, \quad (4.42c)$$

Complementary slackness:

$$\lambda_j \mathcal{G}_{\phi, \omega_j}^\varepsilon(\phi^*) = 0 \quad \text{for } j = 1, \dots, J. \quad (4.42d)$$

Proposition 4.4. (Homogeneity Gibbs functional) *The functional $\mathcal{G}_{\phi, \omega}^\varepsilon$ satisfies the property*

$$\lim_{\varepsilon \rightarrow 0} d\mathcal{G}_{\phi, \omega}^\varepsilon(\phi^*)(\phi^*) = V_\omega(\phi^*) \geq 0. \quad (4.43)$$

Proof. Substitution of $w = \phi^* \in H^1(\tilde{\Omega})$ into (4.37) provides

$$d\mathcal{G}_{\phi, \omega}^\varepsilon(\phi^*)(\phi^*) = \int_{\omega} dg_{\phi}^\varepsilon(\phi^*)(\phi^*) dx, \quad (4.44a)$$

$$dg_{\phi}^\varepsilon(\phi^*)(\phi^*) = |D\phi^*|_\varepsilon - \varepsilon^2 |D\phi^*|_\varepsilon^{-3} (|D\phi^*|_\varepsilon^2 + D\phi D\phi^*). \quad (4.44b)$$

The integrand $dg_{\phi}^\varepsilon(\phi^*)(\phi^*)$ vanishes for $D\phi^* = 0$. On the other hand, for $D\phi^* \neq 0$ we have

$$\lim_{\varepsilon \rightarrow 0} dg_{\phi}^\varepsilon(\phi^*)(\phi^*) = |D\phi^*| \geq 0. \quad (4.45)$$

□

Lemma 4.9. (Positivity) *The solution of the optimization problem (4.32) has the following inner product form that is positive:*

$$(\phi', \phi^*)_{\mathcal{H}} \geq 0. \quad (4.46)$$

Proof. Substitution of $w = \phi^*$ into the stationarity condition (4.42a) provides

$$-(\phi', \phi^*)_{\mathcal{H}} + \sum_{j=1}^J \lambda_j d\mathcal{G}_{\phi, \omega}^{\varepsilon}(\phi^*)(\phi^*) = 0, \quad (4.47)$$

where we have employed the substitution $\phi' = \phi - \phi^*$. The result now follows from taking the limit $\varepsilon \rightarrow 0$, utilizing Proposition 4.4 and invoking the dual feasibility property (4.42c). $\square$

4.4. Best approximations in finite element spaces

In this section, we seek for finite element approximations that satisfy the Gibbs constraints. Consider the best approximation problem

$$\phi^h = \operatorname{arginf}_{\theta^h \in \mathcal{K}_{p, \alpha}} \|\phi - \theta^h\|_{\mathcal{H}}, \quad (4.48a)$$

where the feasible set $\mathcal{K}_{p, \alpha}$ for a finite element approximation space of polynomial degree p and regularity α is defined as

$$\mathcal{K}_{p, \alpha} := \{\phi^h \in \mathcal{V}_{D; p, \alpha}^h : \mathcal{G}_{\phi, \omega_j}(\phi^h) \leq 0, j = 1, \dots, J\}. \quad (4.48b)$$

To proceed, we regularize the nondifferentiable constraint according to (4.35)-(4.36) and introduce the KKT conditions of the regularized problem.

Stationarity:

Find $\phi^h \in \mathcal{V}_{D; p, \alpha}^h, \lambda_j \in \mathbb{R}$ for $j = 1, \dots, J$ such that

$$(\phi^h - \phi, w^h)_{\mathcal{H}} + \sum_{j=1}^J \lambda_j d\mathcal{G}_{\phi, \omega_j}^{\varepsilon}(\phi^h)(w^h) = 0 \quad \text{for all } w^h \in \mathcal{V}_{0, p, \alpha}^h, \quad (4.49a)$$

Primal feasibility:

$$\mathcal{G}_{\phi, \omega_j}^{\varepsilon}(\phi^h) \leq 0 \quad \text{for } j = 1, \dots, J, \quad (4.49b)$$

Dual feasibility:

$$\lambda_j \geq 0 \quad \text{for } j = 1, \dots, J, \quad (4.49c)$$

Complementary slackness:

$$\lambda_j \mathcal{G}_{\phi, \omega_j}^{\varepsilon}(\phi^h) = 0 \quad \text{for } j = 1, \dots, J. \quad (4.49d)$$

The problem (4.49) takes the following algebraic form:

Find ϕ^h, λ such that

$$\mathbf{M}\phi^h = \mathbf{M}\phi - \lambda^T \mathbf{G}(\phi^h), \quad (4.50a)$$

$$\mathbf{g}(\phi^h) \leq 0, \quad (4.50b)$$

$$\lambda \geq 0, \quad (4.50c)$$

$$\lambda^T \mathbf{g}(\phi^h) = 0, \quad (4.50d)$$

where the matrices $\mathbf{M} = [M_{AB}] \in \mathbb{R}^{n_{\text{dof}} \times n_{\text{dof}}}$ and $\mathbf{G}(\phi^h) = [G_{jB}] \in \mathbb{R}^{n_{\text{el}} \times n_{\text{dof}}}$ are given by

$$M_{AB} = (N_A, N_B)_{\mathcal{H}}, \quad (4.51a)$$

$$G_{jB} = \int_{\omega_j} |\mathrm{D}\phi^h|_{\varepsilon}^{-1} \mathrm{D}\phi^h \mathrm{D}N_B - \varepsilon^2 |\mathrm{D}\phi^h|_{\varepsilon}^{-3} \mathrm{D}\phi \mathrm{D}N_B \, dx, \quad (4.51b)$$

and the vectors are $\mathbf{g}(\phi^h) = [\mathcal{G}_{\phi, \omega_j}(\phi^h)] \in \mathbb{R}^{n_{\text{el}}}$ and $\boldsymbol{\lambda} = [\lambda_j]$. Here, n_{dof} denotes the number of dofs.

Remark 4.3. (Computation constrained solutions) In this paper, we use carefully selected examples that allow the construction of solutions of the constrained best approximation problem (4.48). The computation of constrained solutions with standard gradient-based methods is in general very difficult. There are a number of challenges. First, in some situations the number of dofs in the feasible solution set may be very small. It can even occur that the feasible set consists of a single function (see Remark 4.4). This is extremely difficult to find with gradient-based methods. Second, the problem is in general nonconvex. As a consequence, gradient-based methods may get stuck in local optima. This excludes a large class of powerful methodologies from convex optimization (that often rely on the KKT conditions). Third, the problem is highly nonlinear. This means that standard Newton-Raphson linearization methods often do not converge, and one has to work with less efficient approaches such as quasi-Newton-type methods. Finally, we note that similar issues occur in the computation of L^q -best approximations when taking $q \rightarrow 1$, and especially when $q = 1$.

We start with discontinuous approximation spaces ($\alpha = -1$). In order to apply the Gibbs constraints in practice, the subdomains ω_j need to be selected. We select the subsets $\omega_j \subset \Omega, j = 1, \dots, J$ as the finite elements: $\omega_j = K_j$. Note that the set $\mathcal{K}_{p,-1}$ is not empty (it contains at least all piecewise constants). The feasible set is convex for discontinuous piecewise linear basis functions (recall that it is general not convex).

Theorem 4.4. (Convexity of $\mathcal{K}_{1,-1}$) *The feasible set $\mathcal{K}_{1,-1}$ is convex.*

Proof. Let $\phi_1^h, \phi_2^h \in \mathcal{K}_{1,-1}$ be given. Fix an element $K_i = (x_i, x_{i+1})$; on this element the functions ϕ_1^h and ϕ_2^h have a representation $\phi_1^h = a_1x + b_1$ and $\phi_2^h = a_2x + b_2$ for some scalars $a_1, a_2, b_1, b_2 \in \mathbb{R}$. Since $\phi_1^h, \phi_2^h \in \mathcal{K}_{1,-1}$, we now have

$$\begin{aligned} \mathcal{G}_{\phi, K_i}(\phi_1^h) &= \int_{x_i}^{x_{i+1}} |\mathrm{D}\phi_1^h| - \operatorname{sgn}(\mathrm{D}\phi_1^h) \mathrm{D}\phi \, dx \\ &= (x_{i+1} - x_i)|a_1| - \operatorname{sgn}(a_1)(\phi(x_{i+1}) - \phi(x_i)) \leq 0, \end{aligned} \quad (4.52a)$$

$$\mathcal{G}_{\phi, K_i}(\phi^h) = \int_{x_i}^{x_{i+1}} |D\phi_2^h| - \operatorname{sgn}(D\phi_2^h) D\phi dx \quad (4.52b)$$

$$= (x_{i+1} - x_i) |a_2| - \operatorname{sgn}(a_2) (\phi(x_{i+1}) - \phi(x_i)) \leq 0.$$

Without loss of generality, we assume $\phi(x_{i+1}) - \phi(x_i) > 0$. From (4.52), we find $a_1 > 0$ and $a_2 > 0$, and we can write

$$(x_{i+1} - x_i) a_1 - (\phi(x_{i+1}) - \phi(x_i)) \leq 0, \quad (4.53a)$$

$$(x_{i+1} - x_i) a_2 - (\phi(x_{i+1}) - \phi(x_i)) \leq 0. \quad (4.53b)$$

Now, let $\zeta \in [0, 1]$ and define $\phi_\zeta^h = \zeta \phi_1^h + (1 - \zeta) \phi_2^h$. Convexity is then a direct consequence of (4.53)

$$\begin{aligned} \mathcal{G}_{\phi, K_i}(\phi_\zeta^h) &= \int_{x_i}^{x_{i+1}} |D\phi_\zeta^h| - \operatorname{sgn}(D\phi_\zeta^h) D\phi dx \\ &= (x_{i+1} - x_i) (\zeta a_1 + (1 - \zeta) a_2) - (\phi(x_{i+1}) - \phi(x_i)) \\ &\leq 0. \end{aligned} \quad (4.54)$$

□

Consider now the best approximation problem (4.48a)–(4.48b) with the interior penalty optimality (2.18)–(2.19), subject to the elementwise Gibbs constraints. Note that the Gateaux derivatives $d\mathcal{G}_{\phi, \omega_j}^\varepsilon$ are linearly independent. Therefore, the Lagrange multiplier λ_j solely depends on quantities defined on K_j . We visualize the interior penalty-best approximation $\phi^h \in \mathcal{V}_{D;p,-1}^h$ of the smooth step function, subject to the elementwise Gibbs constraints, in Fig. 8 for $p = 1, 2$.

We observe that the constrained interior penalty-best approximations do not show over- or undershoots. Note that this property is not valid in general. Not all feasible solutions are monotonic, since the addition of a piecewise constant does

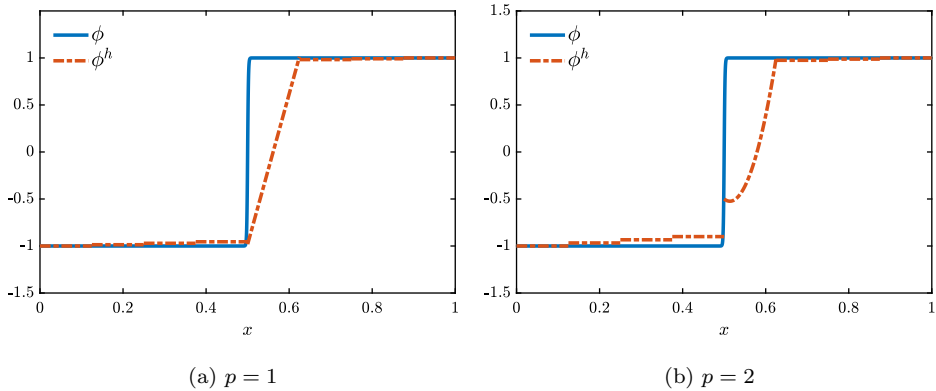


Fig. 8. The interior penalty-best approximation $\phi^h \in \mathcal{V}_{D;p,-1}^h$, with $n_{\text{el}} = 8$, of the smooth step function $\phi = \phi_{0.5}$, subject to elementwise Gibbs constraints, for different p .

not alter the Gibbs functional. Furthermore, both approximations deviate by a small piecewise constant away from the sharp layer. This behavior diminishes as the penalty parameter η is increased.

Next, we focus on the case $\alpha = 0$, i.e. continuous approximation spaces. Again, we take $\omega_j = K_j$. It is easy to see that $\mathcal{K}_{p,0}$ is not empty. Namely, $\mathcal{K}_{1,0}$ contains the piecewise linear interpolant, and $\mathcal{K}_{1,0} \subset \mathcal{K}_{p,0}$ for $p \geq 1$. Considering the best approximation problem (4.48a)–(4.48b) with $\mathcal{H} = H_0^1(\Omega)$, we have the following lemma.

Lemma 4.10. (H_0^1 -orthogonality continuous linear) *Let $\phi^h \in \mathcal{V}_{D,1,0}^h$ be the H_0^1 -best approximation of $\phi \in \mathcal{V}$. We have the property*

$$\int_{K_i} D\phi^h D\phi' \, dx = 0, \quad (4.55)$$

for all elements $K_i = (x_{L,i}, x_{R,i}), i = 1, \dots, n_{\text{el}}$.

Proof. Applying the first Green's identity we find

$$\int_{K_i} D\phi^h D\phi' \, dx = \int_{\partial K_i} n D\phi^h \phi' \, da - \int_{K_i} D^2 \phi^h \phi' \, dx, \quad (4.56)$$

with n the outward unit normal. Noting that for piecewise linear polynomials the second term vanishes and we are left with

$$\int_{K_i} D\phi^h D\phi' \, dx = \phi'(x_{R,i}) D\phi^h(x_{R,i}) - \phi'(x_{L,i}) D\phi^h(x_{L,i}). \quad (4.57)$$

Recalling now from Lemma 3.1 that the H_0^1 -projector provides nodally exact solutions (i.e. $\phi'(x_{R,i}) = \phi'(x_{L,i}) = 0$) completes the proof. $\square$

Lemma 4.11. (Vanishing Lagrange multipliers) *Suppose that the approximation satisfies homogeneous boundary conditions, $\phi^h \in \mathcal{V}_{0,1,0}^h$. The Lagrange multiplier λ_j in the KKT conditions (4.49) vanishes if ϕ^h is not constant on ω_j , for $\varepsilon \rightarrow 0$.*

Proof. Noting that the H_0^1 -best approximation is monotonic and interpolatory, Lemma 4.6 ensures satisfaction of the Gibbs constraints $\mathcal{G}_{\phi,K_i}(\phi^h) \leq 0$ (for $\varepsilon \rightarrow 0$). As a consequence, the first term in (4.49a) vanishes (consistent with Lemma 4.10), and we are left with

$$\sum_{j=1}^J \lambda_j d\mathcal{G}_{\phi,K_j}^\varepsilon(\phi^h)(w^h) = 0, \quad (4.58)$$

for all $w^h \in \mathcal{V}_{0,1,0}^h$. Substituting $w^h = \phi^h$, taking the limit $\varepsilon \rightarrow 0$ and invoking the homogeneity property of Proposition 4.4 yields

$$\sum_{j=1}^J \lambda_j V_{\omega_j}(\phi^h) = 0. \quad (4.59)$$

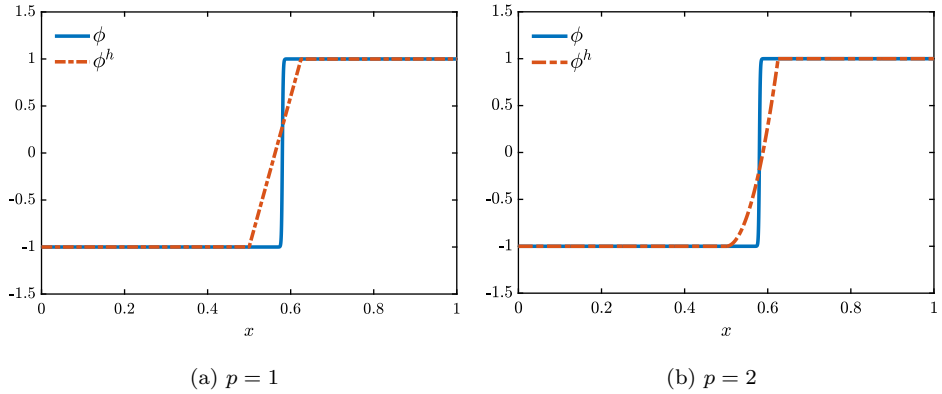


Fig. 9. The H_0^1 -best approximation $\phi^h \in \mathcal{V}_{D;p,0}^h$, with $n_{\text{el}} = 8$, of the smooth step function $\phi = \phi_{0.58}$, subject to elementwise Gibbs constraints, for different p .

Noting that $\lambda_j \geq 0$ and $V_\omega(\phi^h) \geq 0$, we get $\lambda_j V_{\omega_j}(\phi^h) = 0$, for $j = 1, \dots, J$. If ϕ^h is not constant on ω_j , we have $V_{\omega_j}(\phi^h) > 0$ and thus $\lambda_j = 0$. $\square$

Note that, as a consequence of the overlapping support of the basis functions, the Gateaux derivatives $d\mathcal{G}_{\phi,K_j}^\varepsilon$ are linearly dependent. Therefore, the Lagrange multipliers $\lambda_j, j = 1, \dots, n_{\text{el}}$ are *nonlocal*, in the sense that λ_j does not solely depend on quantities defined on ω_j . The H_0^1 -best approximation $\phi^h \in \mathcal{V}_{D;p,0}^h$, $p = 1, 2$, of the smooth step function (with $a = 0.58$) subject to the elementwise Gibbs constraints is illustrated in Fig. 9. We observe in both figures an interpolatory monotonic approximation. Moreover, we have $\mathcal{G}_{\phi,K_j}(\phi^h) = 0$ for all element numbers $j = 1, \dots, J$. In general, the following theorem holds.

Theorem 4.5. (Monotonic interpolant for regularity $\alpha = 0$) *The constrained best approximation $\phi^h \in \mathcal{V}_{D;p,0}^h$ defined by the problem (4.48a)–(4.48b) of a monotonic profile ϕ is a monotonic interpolant.*

Remark 4.4. (Uniqueness) By Theorem 4.5, the best approximation for polynomial degree $p = 1$ is the sole feasible solution in $\mathcal{K}_{1,0}$. This function is thus independent of the optimality condition $\mathcal{H}$.

We now turn our attention to higher-order smooth ($\alpha \geq 1$) approximation spaces $\mathcal{V}_{D;p,\alpha}^h$. The following proposition precludes existence of a solution for $\omega_j = K_j$ in general.

Proposition 4.5. (Infeasible elementwise constraints) *The constrained best approximation problem (4.48a)–(4.48b) with $\omega_j = K_j$ and $a = 0.58$ has in general no solution for $\alpha \geq 1$.*

This is a consequence of the observation that interpolatory solutions exist for $\alpha = -1$ and $\alpha = 0$, but not for $\alpha \geq 1$. We illustrate this for the smooth step

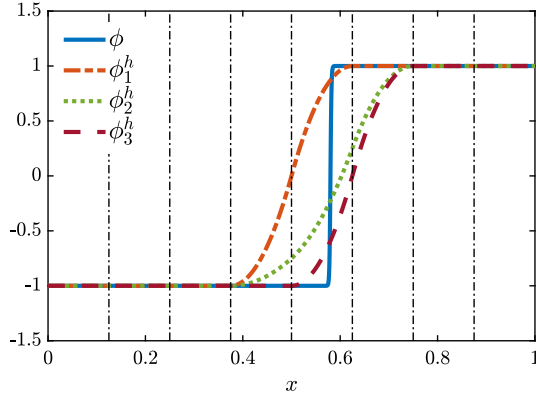


Fig. 10. Sharp finite element approximations $\phi_1^h, \phi_2^h, \phi_3^h \in \mathcal{V}_{D,2,1}^h$, with $n_{\text{el}} = 8$, of the smooth one-dimensional step function $\phi = \phi_{0.58}$. The vertical dashed lines indicate the element boundaries.

function (3.2) with parameter $\alpha = 0.58$ in Fig. 10. Here, we display several sharp finite element approximations with quadratic basis functions ($p = 2$) and regularity $\alpha = 1$. The finite element approximation requires at least 2 elements to capture the sharp layer. There thus necessarily exists an element K_i for which $\mathcal{G}_{\phi, K_i}(\phi^h) > 0$, and the constrained best approximation problem (4.48a)–(4.48b) therefore has no solution.

In general, B-spline basis functions of degree p have the following local support:

$$\text{supp}(N_{i,p}) = (\xi_i, \xi_{i+p+1}), \quad (4.60)$$

which depends on the regularity α . The B-spline basis function $N_{i,p}$ has support in at most $p - k + 2 = \alpha + 2$ elements. Neighboring B-spline basis functions share the support

$$\text{supp}(N_{i,p}) \cap \text{supp}(N_{i+1,p}) = (\xi_{i+1}, \xi_{i+p+1}), \quad (4.61)$$

which consists of at most $\alpha + 1$ elements. With the aim of finding a subdivision of Ω that permits the sharpest approximations, we select ω_j as the union of $\alpha + 1$ neighboring elements. We subdivide domain Ω into disjoint subdomains $\omega_j = \cup_{i \in \mathcal{I}_j} K_i$, where $\mathcal{I}_j$ denotes an index set. Note that shifting groups of elements yields a different subdivision (for $\alpha > 0$). The construction is therefore not unique and we consider $\alpha + 1$ possibilities. For the first index set $\mathcal{I}_1$ we have the options $\{1\}, \dots, \{1, \dots, \alpha + 1\}$. The consecutive index sets contain the next $\alpha + 1$ consecutive numbers, where the last set terminates with the last element number. We visualize two possible subdivisions in Fig. 11.

In Fig. 12, we visualize the H_0^1 -best approximation $\phi^h \in \mathcal{V}_{D,2,1}^h$ of the smooth step function subject to Gibbs constraints on the subdomains of both subdivisions. We see that both approximations are completely free of over- and undershoots. In general, the existence of a feasible solution of the constrained best approximation problem with $\alpha \geq 1$ is an open question.

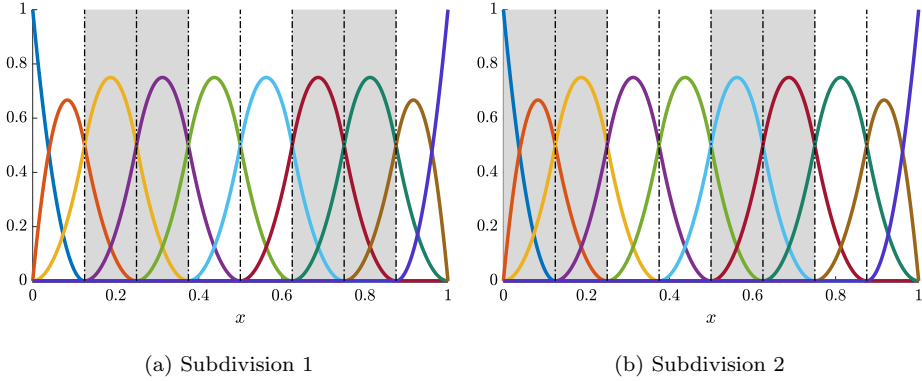


Fig. 11. The two possible subdivisions for $\mathcal{V}_{2,1}^h$ when $n_{\text{el}} = 8$. The vertical dashed lines represent the element boundaries.

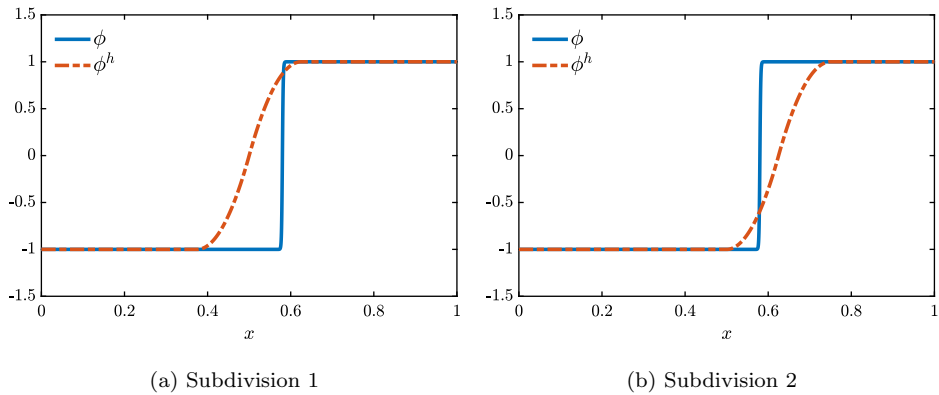


Fig. 12. The H_0^1 -best approximation $\phi^h \in \mathcal{V}_{D;2,1}^h$, with $n_{\text{el}} = 8$, of the smooth step function $\phi = \phi_{0.58}$, subject to Gibbs constraints on each of the subdomains illustrated in Fig. 11.

5. Eliminating the Gibbs Phenomenon in Higher Dimensions

In this section, we extend the strategy to eliminate the Gibbs phenomenon presented in Sec. 4 to higher dimensions. We first introduce the Gibbs constraints for the elimination of the Gibbs phenomenon in general function spaces in Sec. 5.1. After that, in Sec. 5.2, we discuss the Gibbs constraints in the context of best approximations in finite element spaces.

5.1. Gibbs constraints

We present the construction of Gibbs constraints on a multidimensional given subdomain ω , by building upon the one-dimensional framework presented in Sec. 4. The main distinguishing feature in the design of Gibbs constraints in higher dimensions is the notion of directionality. As a consequence, the occurrence of over- and undershoots now depends on the point of view, meaning that there exists no precise

mathematical definition of monotonicity in higher dimensions. To construct a set of constraints, we instead aim to preclude the Gibbs phenomenon in a particular direction. We start by assuming an arbitrary direction, $\mathbf{e}$, and afterwards suggest a particular closed form expression for $\mathbf{e}$.

Definition 5.1. (Directional Gibbs functional) The directional Gibbs functional of the function $\phi^* \in H^1(\tilde{\Omega})$, with respect to the given function $\phi \in H^1(\Omega)$ and on $\omega \subset \Omega$, is defined as

$$\mathcal{G}_{\phi,\omega}(\phi^*; \mathbf{e}) := \int_{\omega} g_{\phi}(\phi^*, \mathbf{e}) d\Omega, \quad (5.1)$$

where the functional g_{ϕ} is defined as

$$g_{\phi}(\phi^*, \mathbf{e}) := |\mathbf{e} \cdot \nabla \phi^*| - \text{sgn}(\mathbf{e} \cdot \nabla \phi^*) \mathbf{e} \cdot \nabla \phi = -\text{sgn}(\mathbf{e} \cdot \nabla \phi^*) \mathbf{e} \cdot \nabla \phi', \quad (5.2)$$

and $\mathbf{e} \in \mathbb{R}^d$ is a unit vector (i.e. $\|\mathbf{e}\|_2 = 1$).

In the following remark, we provide a motivation for this generalization of the one-dimensional Gibbs functional.

Remark 5.1. (Pointwise motivation Gibbs constraint) Define the projection operator $\mathcal{P}_{\mathbf{e}} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ in the direction $\mathbf{e}$ by $\mathcal{P}_{\mathbf{e}} \mathbf{v} := \mathbf{P}_{\mathbf{e}} \mathbf{v}$ with projection matrix $\mathbf{P}_{\mathbf{e}} = \mathbf{e} \otimes \mathbf{e}$. The projections of $\nabla \phi^*$ and $\nabla \phi$ in direction $\mathbf{e}$ are thus given by $\mathcal{P}_{\mathbf{e}} \nabla \phi^*$ and $\mathcal{P}_{\mathbf{e}} \nabla \phi$, respectively. The form of the local functional $g_{\phi}(\phi^*, \mathbf{e})$ is now motivated by the following equivalence:

$$\left. \begin{aligned} \|\mathcal{P}_{\mathbf{e}} \nabla \phi^*\|_2 - \|\mathcal{P}_{\mathbf{e}} \nabla \phi\|_2 \leq 0 &\Leftrightarrow |\mathbf{e} \cdot \nabla \phi^*| \leq |\mathbf{e} \cdot \nabla \phi| \\ \text{sgn}(\mathcal{P}_{\mathbf{e}} \nabla \phi^* \cdot \mathbf{e}) = \text{sgn}(\mathcal{P}_{\mathbf{e}} \nabla \phi \cdot \mathbf{e}) &\Leftrightarrow \text{sgn}(\mathbf{e} \cdot \nabla \phi^*) = \text{sgn}(\mathbf{e} \cdot \nabla \phi) \end{aligned} \right\} \Leftrightarrow g_{\phi}(\phi^*, \mathbf{e}) \leq 0. \quad (5.3)$$

We have the following simple, but important, property.

Proposition 5.1. (Invariance of directional Gibbs functional opposite direction) *The directional Gibbs functional is invariant with respect to flipping the sign of the direction $\mathbf{e}$*

$$\mathcal{G}_{\phi,\omega}(\phi^*; -\mathbf{e}) = \mathcal{G}_{\phi,\omega}(\phi^*; \mathbf{e}). \quad (5.4)$$

To proceed, we provide a short study on the functional $g_{\phi} = g_{\phi}(\phi^*, \mathbf{e})$. In the following proposition, we comment on the sign thereof, depending on the unit vector $\mathbf{e}$.

Proposition 5.2. (Sign of directional g_{ϕ}) *The sign of the functional $g_{\phi}(\phi^*, \mathbf{e})$ depends in the following way on the unit vector $\mathbf{e}$:*

$$g_{\phi}(\phi^*, \mathbf{e}) < 0 \Leftrightarrow \begin{cases} \mathbf{e} \cdot \nabla \phi' > 0 & \text{and} & \mathbf{e} \cdot \nabla \phi^* > 0 & \text{if } \mathbf{e} \cdot \nabla \phi > 0, \\ \mathbf{e} \cdot \nabla \phi' < 0 & \text{and} & \mathbf{e} \cdot \nabla \phi^* < 0 & \text{if } \mathbf{e} \cdot \nabla \phi < 0, \end{cases} \quad (5.5a)$$

$$g_{\phi}(\phi^*, \mathbf{e}) = 0 \Leftrightarrow \mathbf{e} \cdot \nabla \phi^* = 0 \quad \text{or} \quad \mathbf{e} \cdot \nabla \phi' = 0, \quad (5.5b)$$

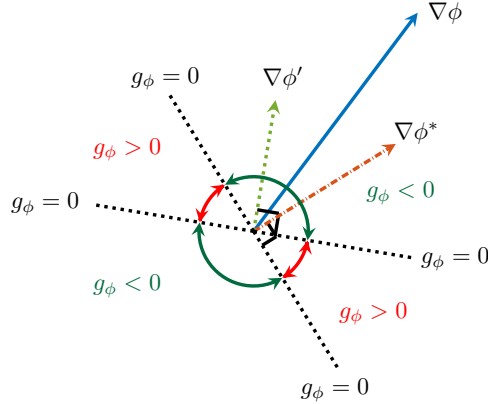


Fig. 13. (Color online) Two-dimensional visualization of the sign of $g_\phi = g_\phi(\phi^*, \mathbf{e})$ for a specific case. A unit vector $\mathbf{e}$ in the green region corresponds to $g_\phi(\phi^*, \mathbf{e}) < 0$, in the red region to $g_\phi(\phi^*, \mathbf{e}) > 0$, and at the intersections to $g_\phi(\phi^*, \mathbf{e}) = 0$.

$$g_\phi(\phi^*, \mathbf{e}) > 0 \Leftrightarrow \begin{cases} \mathbf{e} \cdot \nabla \phi' < 0 & \text{or} & \mathbf{e} \cdot \nabla \phi^* < 0 & \text{if } \mathbf{e} \cdot \nabla \phi > 0, \\ \mathbf{e} \cdot \nabla \phi' > 0 & \text{or} & \mathbf{e} \cdot \nabla \phi^* > 0 & \text{if } \mathbf{e} \cdot \nabla \phi < 0. \end{cases} \quad (5.5c)$$

We provide a visualization of the sign of $g_\phi(\phi^*, \mathbf{e})$ for varying direction $\mathbf{e}$ for a two-dimensional scenario in Fig. 13.

In general, there always exists a direction $\mathbf{e}$ for which $g_\phi(\phi^*, \mathbf{e}) > 0$, as expressed in the following proposition.

Proposition 5.3. (Supremum of $g_\phi(\phi^*, \mathbf{e})$) *The supremum of $g_\phi(\phi^*, \mathbf{e})$ is given by*

$$\sup_{\|\mathbf{e}\|_2=1} g_\phi(\phi^*, \mathbf{e}) = \begin{cases} \mathbf{b} \cdot \nabla \phi' & \text{if } \nabla \phi^* \cdot \nabla \phi' \geq 0, \\ \|\nabla \phi'\|_2 & \text{if } \nabla \phi^* \cdot \nabla \phi' < 0, \end{cases} \quad (5.6)$$

where $\mathbf{b} \in \{\mathbf{a} \in \mathbb{R}^d : \|\mathbf{a}\|_2 = 1, \mathbf{a} \cdot \nabla \phi^* = 0\}$.

Proof. Distinguish the two cases: (i) $\nabla \phi^* \cdot \nabla \phi' < 0$ and (ii) $\nabla \phi^* \cdot \nabla \phi' \geq 0$.

- (i) If $\mathbf{e} \cdot \nabla \phi^* \geq 0$, then the supremum of g_ϕ is $\|\nabla \phi'\|_2$, which is attained at $\mathbf{e} = -\nabla \phi' / \|\nabla \phi'\|_2$. Similarly, if $\mathbf{e} \cdot \nabla \phi^* < 0$, then the supremum of g_ϕ is again $\|\nabla \phi'\|_2$, which is then attained at $\mathbf{e} = \nabla \phi' / \|\nabla \phi'\|_2$.
- (ii) If $\mathbf{e} \cdot \nabla \phi^* \geq 0$, then g_ϕ is a decreasing function in $\mathbf{e} \cdot \nabla \phi^*$. Therefore, g_ϕ attains its supremum when $\mathbf{e} \cdot \nabla \phi^* = 0$. Similarly, if $\mathbf{e} \cdot \nabla \phi^* < 0$, then g_ϕ is an increasing function in $\mathbf{e} \cdot \nabla \phi^*$. Again, g_ϕ attains its supremum when $\mathbf{e} \cdot \nabla \phi^* = 0$. $\square$

Propositions 5.2 and 5.3 show that demanding the elimination of the Gibbs functional in each possible direction, which might intuitively yield the most desirable result, is not prudent: selecting the direction $\mathbf{e}$ to maximize $g_\phi = g_\phi(\phi^*, \mathbf{e})$ and subsequently using this in the Gibbs constraint would only yield (i) the zero

approximation $\phi^* = 0$ and (ii) the perfect approximation $\phi^* = \phi$ as feasible solutions. To proceed, we must thus select a direction $\mathbf{e}$ that leads to both suitable and practical constraints. The relevant directions to consider are functions of the (approximate) solution gradients: (i) $\mathbf{e} = \nabla\phi/\|\nabla\phi\|_2$, (ii) $\mathbf{e} = \nabla\phi^*/\|\nabla\phi^*\|_2$ and (iii) $\mathbf{e} = \nabla\phi'/\|\nabla\phi'\|_2$ leading to the expressions

$$g_\phi\left(\phi^*, \frac{\nabla\phi}{\|\nabla\phi\|_2}\right) = -|\nabla\phi \cdot \nabla\phi^*|^{-1} \|\nabla\phi\|_2^{-1} (\nabla\phi \cdot \nabla\phi) (\nabla\phi \cdot \nabla\phi'), \quad (5.7a)$$

$$g_\phi\left(\phi^*, \frac{\nabla\phi^*}{\|\nabla\phi^*\|_2}\right) = -\|\nabla\phi^*\|_2^{-1} \nabla\phi^* \cdot \nabla\phi', \quad (5.7b)$$

$$g_\phi\left(\phi^*, \frac{\nabla\phi'}{\|\nabla\phi'\|_2}\right) = -\|\nabla\phi'\|_2^{-1} \nabla\phi^* \cdot \nabla\phi' \|\nabla\phi^*\|_2^{-1} \nabla\phi^* \cdot \nabla\phi'. \quad (5.7c)$$

Insisting compatibility with the one-dimensional case requires the choice of direction $\mathbf{e} = \nabla\phi^*/\|\nabla\phi^*\|_2$, which we focus on exclusively in the following.

Definition 5.2. (Gibbs functional and constraint) The Gibbs functional of the function $\phi^* \in H^1(\tilde{\Omega})$, with respect to the given function $\phi \in H^1(\Omega)$ and on $\omega \subset \Omega$, is defined as

$$\mathcal{G}_{\phi,\omega}(\phi^*) := \int_{\omega} g_\phi(\phi^*) d\Omega, \quad (5.8)$$

where the functional g_ϕ is defined as

$$g_\phi(\phi^*) := \begin{cases} -\|\nabla\phi^*\|_2^{-1} \nabla\phi^* \cdot \nabla\phi' & \text{if } \nabla\phi^* \neq 0, \\ 0 & \text{if } \nabla\phi^* = 0. \end{cases} \quad (5.9)$$

The associated Gibbs constraint reads

$$\mathcal{G}_{\phi,\omega}(\phi^*) \leq 0. \quad (5.10)$$

Note that the Gibbs functional is frame-invariant, and that the properties of Proposition 4.2 and Lemma 4.4 are inherited from the one-dimensional case.

5.2. Gibbs constraints for finite element best approximations

In this section, we apply the Gibbs constraints to finite element best approximations. Again, we consider the following general form of the best approximation problem:

find $\phi^h \in \mathcal{V}_{D;p,\alpha}^h$ such that

$$\phi^h = \underset{\theta^h \in \mathcal{K}_{p,\alpha}}{\operatorname{arginf}} \|\phi - \theta^h\|_{\mathcal{H}}, \quad (5.11a)$$

where the feasible set $\mathcal{K}_{p,\alpha}$ is defined as

$$\mathcal{K}_{p,\alpha} := \{\phi^h \in \mathcal{V}_{D;p,\alpha}^h : \mathcal{G}_{\phi,\omega_j}(\phi^h) \leq 0, j = 1, \dots, J\}. \quad (5.11b)$$

We first consider discontinuous finite element approximation spaces ($\alpha = -1$). Analogous to the one-dimensional case, we consider (5.11a)–(5.11b) with $\mathcal{H} = \text{IP}$, i.e. the interior penalty-best approximation. We select the subdomains ω_j as the finite elements $K_j, j = 1, \dots, n_{\text{el}}$. Feasibility of the constrained optimization problem is a consequence of the existence of an elementwise constant solution ϕ^h , for which $\mathcal{G}_{\phi, K_j}(\phi^h) = 0$. Again, we may analyze the solution properties via the KKT conditions (4.49). In case of homogeneous boundary conditions, we have $(\phi^h, \phi')_{\mathcal{H}} \geq 0$ (Lemma 4.9). In Fig. 14, we visualize the interior penalty-best approximation $\phi^h \in \mathcal{V}_{D;p,-1}^h, p = 1, 2$ of the profile of example 1, subject to the elementwise Gibbs constraints. We observe that the constrained interior penalty-best approximations display a significant reduction of over- and undershoots compared to the nonconstrained solutions depicted in Sec. 3.2.

Next, we study the case with regularity $\alpha = 0$. Consider again the example with the sharp layer skew to the mesh. In Fig. 15, we visualize the sharpest possible approximation $\phi^h \in \mathcal{V}_{D,1,0}^h$, and plot the elementwise values of the Gibbs functionals $\mathcal{G}_{\phi, K_i}(\phi^h)$. We note that the sharpest approximation is interpolatory. Nevertheless, the Gibbs constraints are not fulfilled elementwise. Similar to the one-dimensional case, one could define ω_j as the collection of several elements. However, a careful examination reveals that the only possible choice is the collection of all elements $\omega = \bigcup_{j=1}^{n_{\text{el}}} K_j$, which would constitute a rather weak condition. It appears impossible to obtain a feasible solution when using subdomains ω_j that consist of a few elements. This is a consequence of the support of the basis functions, which contain multiple elements. The Gibbs constraints for any collection of elements K_i on the off-diagonal

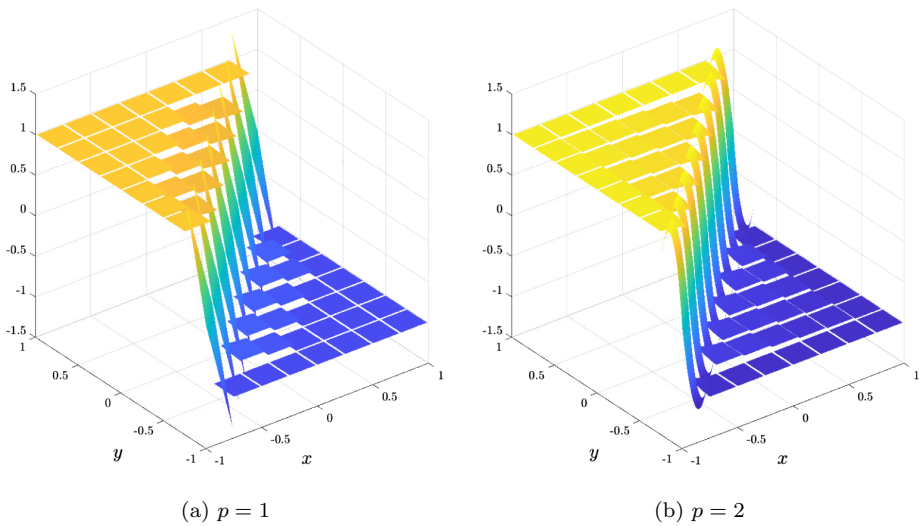


Fig. 14. The interior penalty-best approximation $\phi^h \in \mathcal{V}_{D;p,-1}^h$ of the smooth step function, with $n_{\text{el}} = 8 \times 8$, subject to elementwise Gibbs constraints, for different p .

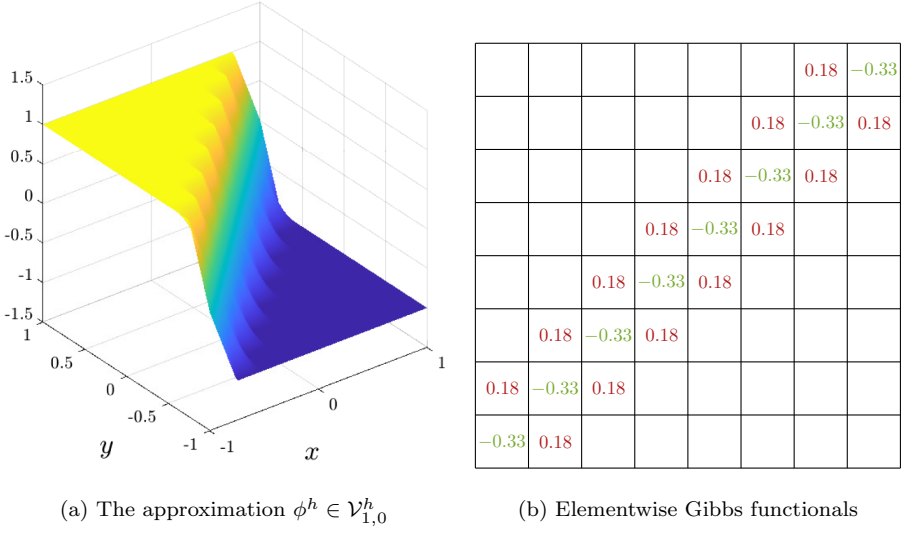


Fig. 15. The sharpest possible approximation $\phi^h \in \mathcal{V}_{1,0}^h$ (a) and the corresponding elementwise values of the Gibbs functionals $\mathcal{G}_{\phi,K_i}(\phi^h)$ (b).

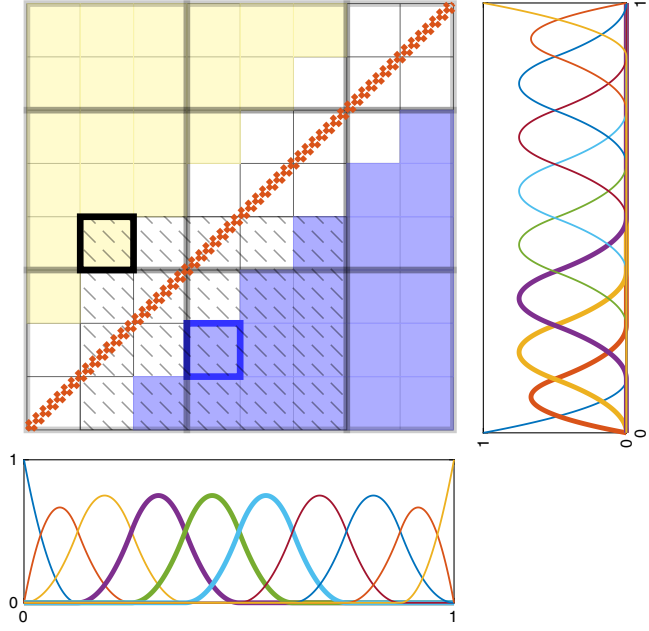


Fig. 16. Subdivision of an 8×8 element domain into groups of maximum 3×3 elements with approximation space $\mathcal{V}_{D,2,1}^h$. The right and bottom inset figures display the univariate basis functions along the corresponding axes.

with $x \geq y$ ($x \leq y$) would enforce $\phi_{K_i}^h$ to be close to 1 (and -1). This is in conflict with the continuity requirement of the finite element approximation space $V_{D,1,0}^h$, and persists when using higher-order polynomials.

Lastly, we consider higher-order smooth finite element approximation spaces ($\alpha \geq 1$). The above discussed infeasibility of the optimization problem also applies to higher-order smooth approximations. We illustrate this for the smooth step function with a quadratic B-spline approximation space $\mathcal{V}_{D,2,1}^h$. Consider a subdivision into groups of maximum 3×3 elements, as visualized in Fig. 16. In order to satisfy the Gibbs constraint on the element group on the middle-bottom we must have $\phi^h = -1$ in the blue boxed element. Similarly, in the black boxed element we require $\phi^h = 1$. These two conditions are incompatible.

6. Conclusions

In this paper, we constructed a set of integral constraints with the aim of eliminating the Gibbs phenomenon in finite element best approximations. We first provided an overview of the Gibbs phenomenon for best approximations in finite element spaces. We illustrated with computational examples that spurious oscillations occur in one and two dimensions for standard projections onto finite element spaces of arbitrary degree and regularity (with the exception of the one-dimensional H_0^1 -projection with linear continuous finite elements). The proposed constraints build onto the concept of total variation. In this regard, we established in one dimension the interrelation between the Gibbs constraint and interpolatory and monotonic approximations, as well as the maximum principle. Furthermore, we displayed in one dimension that the proposed constraints may be applied elementwise when the finite element space is either discontinuous or $\mathcal{C}^0$ -continuous. For higher regularity finite element spaces, the integration domains of the constraints depend on multiple elements. We showed that enforcing the constraints removes over- and undershoots for continuous finite element spaces, and suppresses them for discontinuous finite element spaces. In higher dimensions, the constraints act in the direction of the solution gradient. The applicability of the constraints is then limited to discontinuous finite element spaces. We demonstrated that also in this case over- and undershoots are severely reduced.

We recognize two open problems that require further investigation. The first is linked to the last observation, namely the extension of the set of constraints to continuous finite element spaces in higher dimensions. As a consequence of the overlapping support of the basis functions, it is hard to identify and eliminate the Gibbs phenomenon on a local level. The second open problem is the construction of a finite element method that incorporates these constraints in practical computations. We conjecture that a possible resolution lies in the variational multiscale framework, in particular through the work of Evans *et al.*⁷ and ten Eikelder.⁴¹

Acknowledgments

The authors are grateful to Thomas J. R. Hughes for insightful discussions on the topic. M. F. P. ten Eikelder was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) via the Walter Benjamin project EI 1210/1-1. Dominik Schillinger gratefully acknowledges support from the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) via the Emmy Noether Award SCH 1249/2-1.

ORCID

Marco F. P. ten Eikelder  <https://orcid.org/0000-0002-1153-146X>

Stein K. F. Stoter  <https://orcid.org/0000-0001-5845-4900>

Yuri Bazilevs  <https://orcid.org/0000-0003-4022-5039>

Dominik Schillinger  <https://orcid.org/0000-0002-9068-6311>

References

1. Y. Bazilevs, L. Beirao da Veiga, J. A. Cottrell, T. J. R. Hughes and G. Sangalli, Iso-geometric analysis: Approximation, stability and error estimates for h -refined meshes, *Math. Models Methods Appl. Sci.* **16** (2006) 1031–1090.
2. Y. Bazilevs, V. M. Calo, J. A. Cottrell, T. J. R. Hughes, A. Reali and G. Scovazzi, Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows, *Comput. Methods Appl. Mech. Engrg.* **197** (2007) 173–201.
3. Y. Bazilevs, V. M. Calo, T. E. Tezduyar and T. J. R. Hughes, $YZ\beta$ discontinuity-capturing for advection-dominated processes with application to arterial drug delivery, *Internat. J. Numer. Methods Fluids* **54** (2007) 593–608.
4. A. N. Brooks and T. J. R. Hughes, Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations, *Comput. Methods Appl. Mech. Engrg.* **32** (1982) 199–259.
5. J. A. Cottrell, T. J. R. Hughes and Y. Bazilevs, *Isogeometric Analysis: Toward Integration of CAD and FEA* (Wiley, 2009).
6. J. M. Davis and P. Hagelstein, Gibbs phenomena for some classical orthogonal polynomials, *J. Math. Anal. Appl.* **505** (2022) 125574.
7. J. A. Evans, T. J. R. Hughes and G. Sangalli, Enforcement of constraints and maximum principles in the variational multiscale method, *Comput. Methods Appl. Mech. Engrg.* **199** (2009) 61–76.
8. T. H. Fay and P. H. Kloppers, The Gibbs phenomenon for series of orthogonal polynomials, *Internat. J. Math. Ed. Sci. Tech.* **37** (2006) 973–989.
9. J. Foster and F. B. Richards, The Gibbs phenomenon for piecewise-linear approximation, *Amer. Math. Monthly* **98** (1991) 47–49.
10. D. Gottlieb and C.-W. Shu, On the Gibbs phenomenon and its resolution, *SIAM Rev.* **39** (1997) 644–668.
11. J.-L. Guermond, A finite element technique for solving first-order PDEs in L^p , *SIAM J. Numer. Anal.* **42** (2004) 714–737.
12. J.-L. Guermond, R. Pasquetti and B. Popov, Entropy viscosity method for nonlinear conservation laws, *J. Comput. Phys.* **230** (2011) 4248–4267.
13. A. Harten, High resolution schemes for hyperbolic conservation laws, *J. Comput. Phys.* **49** (1983) 357–393.

14. A. Harten, On a class of high resolution total-variation-stable finite-difference schemes, *SIAM J. Numer. Anal.* **21** (1984) 1–23.
15. E. Hewitt and R. E. Hewitt, The Gibbs-Wilbraham phenomenon: An episode in Fourier analysis, *Arch. Hist. Exact Sci.* **21** (1979) 129–160.
16. P. Houston, S. Roggendorf and K. G. van der Zee, Eliminating Gibbs phenomena: A non-linear Petrov–Galerkin method for the convection–diffusion–reaction equation, *Comput. Math. Appl.* **80** (2020) 851–873.
17. P. Houston, S. Roggendorf and K. G. van der Zee, Gibbs phenomena for L^q -best approximation in finite element spaces, *ESAIM Math. Model. Numer. Anal.* **56** (2022) 177–211.
18. T. J. R. Hughes, Multiscale phenomena: Green’s functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, *Comput. Methods Appl. Mech. Engrg.* **127** (1995) 387–401.
19. T. J. R. Hughes, J. A. Cottrell and Y. Bazilevs, Isogeometric analysis: CAD, finite elements, NURBS, exact geometry, and mesh refinement, *Comput. Methods Appl. Mech. Engrg.* **194** (2005) 4135–4195.
20. T. J. R. Hughes, G. Feijóo, L. Mazzei and J. B. Quincy, The variational multiscale method — A paradigm for computational mechanics, *Comput. Methods Appl. Mech. Engrg.* **166** (1998) 3–24.
21. T. J. R. Hughes, L. P. Franca and M. Balestra, A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška–Brezzi condition: A stable Petrov–Galerkin formulation of the Stokes problem accommodating equal-order interpolations, *Comput. Methods Appl. Mech. Engrg.* **59** (1986) 85–99.
22. T. J. R. Hughes, L. P. Franca and G. M. Hulbert, A new finite element formulation for fluid dynamics: VIII. The Galerkin/least-squares method for advective–diffusive equations, *Comput. Methods Appl. Mech. Engrg.* **73** (1989) 173–189.
23. T. J. R. Hughes, L. P. Franca and M. Mallet, A new finite element formulation for computational fluid dynamics: I. Symmetric forms of the compressible Euler and Navier–Stokes equations and the second law of thermodynamics, *Comput. Methods Appl. Mech. Engrg.* **54** (1986) 223–234.
24. T. J. R. Hughes, M. Mallet and M. Akira, A new finite element formulation for computational fluid dynamics: II. Beyond SUPG, *Comput. Methods Appl. Mech. Engrg.* **54** (1986) 341–355.
25. T. J. R. Hughes and G. Sangalli, Variational multiscale analysis: The fine-scale Green’s function, projection, optimization, localization, and stabilized methods, *SIAM J. Numer. Anal.* **45** (2007) 539–557.
26. M. M. Jacob, Sur le phénomène de Gibbs dans les développements de séries de polynômes d’Hermite, *C. R. Acad. Sci. Paris* **204** (1937) 1540–1543.
27. B.-N. Jiang, Non-oscillatory and non-diffusive solution of convection problems by the iteratively reweighted least-squares finite element method, *J. Comput. Phys.* **105** (1993) 108–121.
28. S.-M. Kaber, The Gibbs phenomenon for Jacobi expansions, *Commun. Appl. Anal.* **10** (2006) 133–148.
29. S. N. Kružkov, First order quasilinear equations in several independent variables, *Math. USSR Sb.* **10** (1970) 217.
30. J. E. Lavery, Nonoscillatory solution of the steady-state inviscid Burgers’ equation by mathematical programming, *J. Comput. Phys.* **79** (1988) 436–448.
31. J. E. Lavery, Solution of steady-state, two-dimensional conservation laws by mathematical programming, *SIAM J. Numer. Anal.* **28** (1991) 141–155.
32. J. Liu, H. Gomez, J. A. Evans, T. J. R. Hughes and C. M. Landis, Functional entropy variables: A new methodology for deriving thermodynamically consistent

- algorithms for complex fluids, with particular reference to the isothermal Navier–Stokes–Korteweg equations, *J. Comput. Phys.* **248** (2013) 47–86.
33. J. Liu, C. M. Landis, H. Gomez and T. J. R. Hughes, Liquid–vapor phase transition: Thermomechanical theory, entropy stable numerical formulation, and boiling simulations, *Comput. Methods Appl. Mech. Engrg.* **297** (2015) 476–553.
 34. R. B. Lowrie and P. L. Roe, On the numerical solution of conservation laws by minimizing residuals, *J. Comput. Phys.* **113** (1994) 304–308.
 35. F. B. Richards, A Gibbs phenomenon for spline functions, *J. Approx. Theory* **66** (1991) 334–351.
 36. K. Shahbazi, An explicit expression for the penalty parameter of the interior penalty method, *J. Comput. Phys.* **205** (2005) 401–407.
 37. S. K. F. Stoter, B. Cockburn, T. J. R. Hughes and D. Schillinger, Discontinuous Galerkin methods through the lens of variational multiscale analysis, *Comput. Methods Appl. Mech. Engrg.* **388** (2022) 114220.
 38. S. K. F. Stoter, M. F. P. ten Eikelder, F. de Prenter, I. Akkerman, E. H. van Brummelen, C. V. Verhoosel and D. Schillinger, Nitsche’s method as a variational multiscale formulation and a resulting boundary layer fine-scale model, *Comput. Methods Appl. Mech. Engrg.* **382** (2021) 113878.
 39. M. F. P. ten Eikelder and I. Akkerman, Correct energy evolution of stabilized formulations: The relation between VMS, SUPG and GLS via dynamic orthogonal small-scales and isogeometric analysis. I: The convective–diffusive context, *Comput. Methods Appl. Mech. Engrg.* **331** (2018) 259–280.
 40. M. F. P. ten Eikelder and I. Akkerman, Correct energy evolution of stabilized formulations: The relation between VMS, SUPG and GLS via dynamic orthogonal small-scales and isogeometric analysis. II: The incompressible Navier–Stokes equations, *Comput. Methods Appl. Mech. Engrg.* **340** (2018) 1135–1159.
 41. M. F. P. ten Eikelder and I. Akkerman, Variation entropy: A continuous local generalization of the TVD property using entropy principles, *Comput. Methods Appl. Mech. Engrg.* **355** (2019) 261–283.
 42. M. F. P. ten Eikelder and I. Akkerman, A novel diffuse-interface model and a fully-discrete maximum-principle-preserving energy-stable method for two-phase flow with surface tension and non-matching densities, *Comput. Methods Appl. Mech. Engrg.* **379** (2021) 113751.
 43. M. F. P. ten Eikelder, Y. Bazilevs and I. Akkerman, A theoretical framework for discontinuity capturing: Joining variational multiscale analysis and variation entropy theory, *Comput. Methods Appl. Mech. Engrg.* **359** (2020) 112664.