



# Residual Minimization for Isogeometric Analysis in Reduced and Mixed Forms

Victor M. Calo<sup>1,2</sup>, Quanling Deng<sup>1</sup>(✉), Sergio Rojas<sup>1</sup>, and Albert Romkes<sup>3</sup>

<sup>1</sup> Department of Applied Geology, Curtin University, Kent Street, Bentley, Perth, WA 6102, Australia

{Victor.Calo,Quanling.Deng}@curtin.edu.au, srojash@gmail.com

<sup>2</sup> Mineral Resources, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Kensington, Perth, WA 6152, Australia

<sup>3</sup> Department of Mechanical Engineering, South Dakota School of Mines and Technology, 501 E. St. Joseph Street, Rapid City, SD 57701, USA  
Albert.Romkes@sdsmt.edu

**Abstract.** Most variational forms of isogeometric analysis use highly-continuous basis functions for both trial and test spaces. Isogeometric analysis results in excellent discrete approximations for differential equations with regular solutions. However, we observe that high continuity for test spaces is not necessary. In this work, we present a framework which uses highly-continuous B-splines for the trial spaces and basis functions with minimal regularity and possibly lower order polynomials for the test spaces. To realize this goal, we adopt the residual minimization methodology. We pose the problem in a mixed formulation, which results in a system governing both the solution and a Riesz representation of the residual. We present various variational formulations which are variationally-stable and verify their equivalence numerically via numerical tests.

**Keywords:** Isogeometric analysis · Finite elements · Discontinuous Petrov-Galerkin · Mixed formulation

## 1 Introduction

Isogeometric analysis, introduced in 2005 [12,26], is a widely-used numerical method for solving partial differential equations (PDEs). This analysis unifies the finite element methods with computer-aided design tools. Within the framework of the classic Galerkin finite element methods, isogeometric analysis uses as basis functions the functions that describe the geometry in the computer-aided design (CAD) and engineering (CAE) technologies, namely, B-splines or non-uniform rational B-splines (NURBS) and their generalizations. These CAD/CAE basis functions may possess higher continuity. This smoothness improves the numerical approximations of PDEs which have highly regular solutions. Bazilevs *et al.* established the approximation, stability, and error estimates in [2].

Cottrell *et al.* in [13] first used isogeometric analysis to study the structural vibrations and wave propagation problems. Spectral analysis shows that isogeometric elements significantly improve the accuracy of the spectral approximation when compared with classical finite elements. In [27], the authors explored the additional advantages of isogeometric analysis on the spectral approximation over finite elements. Moreover, the work [6, 21, 22, 35] minimized the spectral approximation errors for isogeometric elements. The minimized spectral errors possess superconvergence (two extra orders) as the mesh is refined.

Galerkin isogeometric analysis uses highly-continuous B-splines or NURBS for both trial and test spaces. For PDEs with smooth solutions (high regularities), isogeometric elements based on smooth functions render solutions which have a better physical interpretation. For example, for a PDE with a solution in  $H^2$ , both the exact solution field and exact flux field are expected to be globally continuous. Isogeometric analysis with  $C^1$  B-spline basis functions produces a globally continuous flux field while the classical finite element method does not. Therefore, for smooth problems, it makes sense to use highly-continuous B-splines for the trial spaces. However, it is not necessary to use highly-continuous B-splines for the test spaces. In the view of the work [11] which explores the cost of continuity, we expect an extra cost for solving the resulting system when we apply unnecessary high continuities for the basis functions in the test spaces. This extra cost per degree of freedom for highly-continuous discretizations was verified for direct [9, 11, 33] and iterative solvers [10]. Additionally, the work [24, 25] shows that a reduction in the continuity of the trial and test spaces may lead to faster and more accurate solutions. Thus, we seek to develop an isogeometric analysis which uses minimal regularity for the test spaces. In this work, we establish a framework which uses highly-continuous B-splines for the trial spaces and basis functions with minimal regularity for the test spaces.

To realize this goal, we adopt the residual minimization methodology sharing with the Discontinuous Petrov-Galerkin (DPG) method, the idea of a Riesz representation for the residual with respect to a stabilizing norm (c.f., [17, 18, 20, 38]). The main idea of DPG is the use of discontinuous or broken basis functions for the test spaces within the Petrov-Galerkin framework. The method computes the optimal test function by using a trial-to-test operator. The goal is to stabilize the discrete formulation [5, 30–32] without parameter tuning. The DPG method can be interpreted as a minimum residual method in which the residual is measured in the dual test norm [19]. By introducing an auxiliary unknown representing the Riesz representation of residual, the method is cast into a mixed problem. Effectively, we extend to highly continuous isogeometric discretizations the method described in [7] where standard  $C^0$  finite element basis functions build the trial space while the test space uses broken polynomials. The proposed method has a feature which distinguishes us from DPG. We use highly-continuous B-splines for space that we seek for the solution, while DPG uses discontinuous basis functions. Consequently, DPG introduces additional unknowns that live on the mesh skeleton; see equation (10) or (39) in [19]. We do not introduce any additional unknowns. Instead, we use basis functions with minimal regularity in the sense

that no inner products are introduced on the mesh skeleton (thus, the bilinear form only involves element integrations) while maintaining the inf-sup condition for the resulting variational formulation.

The rest of this paper is organized as follows. Section 2 describes the problem under consideration and introduces various variational formulations. Section 3 presents isogeometric formulations at the discrete level, while Sect. 4 shows numerical examples to demonstrate the performance of the formulations. Concluding remarks are given in Sect. 5.

## 2 Problem Statement

Let  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , be an open bounded domain with Lipschitz boundary  $\partial\Omega$ . We consider the advection-diffusion-reaction equation: Find  $u$  such that

$$\begin{aligned} -\nabla \cdot (\kappa \nabla u - \beta u) + \gamma u &= f & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega, \end{aligned} \quad (1)$$

where  $\kappa$  is the diffusion coefficient,  $\beta$  is the advective velocity vector,  $\gamma \geq 0$  is the reaction coefficient,  $u$  is the function to be found, and  $f$  is a forcing function. This problem can be cast into an equivalent system of two first-order equations

$$\begin{aligned} \mathbf{q} - \kappa \nabla u + \beta u &= \mathbf{0} & \text{in } \Omega, \\ -\nabla \cdot \mathbf{q} + \gamma u &= f & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega, \end{aligned} \quad (2)$$

where  $\mathbf{q}$  is an auxiliary variable standing for the flux.

### 2.1 General Setting

We present the general setting for Eqs. (1) and (2) as they have distinguishing features. Let  $V$  and  $W$  denote the trial and test space, respectively. For (2) we let  $(u, \mathbf{q})$  be the solution pair in its trial space  $V^u \times V^q$  and let its corresponding test space be  $W^u \times W^q$ . We specify these spaces in each of the variational formulations. We denote by  $\|\cdot\|_{0,\Omega}$  the  $L^2$  norm. Let  $H^1(\Omega)$  be the Hilbert space and  $H_0^1(\Omega)$  be the space of all functions in  $H^1(\Omega)$  vanishing at the boundary  $\partial\Omega$ . The  $H^1$  seminorm is defined as  $|w|_{1,\Omega} = \|\nabla w\|_{0,\Omega}$  and  $H^1$  norm is defined as  $\|w\|_{1,\Omega}^2 = \|w\|_{0,\Omega}^2 + |w|_{1,\Omega}^2$ . Finally,  $H(\text{div}, \Omega)$  consists of all square integrable vector-valued fields on  $\Omega$  whose divergence is a function that is also square integrable. Similarly, the  $H(\text{div}, \Omega)$  norm is defined as  $\|\mathbf{p}\|_{\text{div},\Omega}^2 = \|\mathbf{p}\|_{0,\Omega}^2 + \|\nabla \cdot \mathbf{p}\|_{0,\Omega}^2$ .

Let  $\mathcal{T}_h$  be a partition of  $\Omega$  into non-overlapping mesh elements. For simplicity and the purpose of using B-splines in multiple dimensions, we assume the tensor-product structure. Let  $K \in \mathcal{T}_h$  be a generic element and denote by  $\partial K$  its boundary. Let  $\mathbf{n}$  be the outward unit normal vector. Let  $(\cdot, \cdot)_S$  denote the  $L^2(S)$  inner product where  $S$  is a  $d$  or  $d - 1$  dimensional domain.

At discrete level, we define the finite spaces, namely the finite subspaces of  $V, W, V^u \times V^q$ , and  $W^u \times W^q$ . Since isogeometric analysis adopts highly-continuous basis functions, we specify the corresponding finite spaces with both the polynomial order  $p$  and the order  $k$  of global continuity (that is,  $C^k$ ). Let us denote by  $V_{p,k}^h, W_{p,k}^h$  and  $V_{p,k}^{u,h}, V_{p,k}^{q,h}, W_{p,k}^{u,h}, W_{p,k}^{q,h}$  the finite-dimensional subspaces of  $V, W$  for (1) and  $V^u, V^q, W^u, W^q$  for (2), respectively. In this work, we construct the test spaces  $W_{q,l}^h$  and  $W_{q,l}^{u,h}$  using basis functions with lower order polynomials as well as lower regularity.

**2.2 Various Variational Formulations at Continuous Level**

In this section, we present six variational formulations for both (1) and (2) following closely the six formulations described within the DPG framework in [19]. In particular, we add one more formulation which resembles the isogeometric collocation method.

We start with the abstract variational formulations of (1) and (2). The variational weak formulations of (1) can be written as: Find  $u \in V$  such that

$$b(w, u) = \ell(w), \quad \forall w \in W, \tag{3}$$

where  $b(\cdot, \cdot)$  and  $\ell(\cdot)$  are bilinear and linear forms, respectively. Similarly, the variational weak formulations of (2) can be written: Find  $(u, \mathbf{q}) \in V^u \times V^q$  such that

$$b((w, \mathbf{p}), (u, \mathbf{q})) = \ell((w, \mathbf{p})), \quad \forall (w, \mathbf{p}) \in W^u \times W^q, \tag{4}$$

where  $b(\cdot, \cdot)$  and  $\ell(\cdot)$  are bilinear and linear forms defined over two fields. Herein, we refer to  $(u, \mathbf{q})$  as the solution (trial) pair while to  $(w, \mathbf{p})$  as the weighting (test) function pair. Equations (1) and (3) are the primal forms of the PDE and the variational formulation while Eqs. (2) and (4) are their mixed forms. We keep this structure for the forms at discrete level in Sect. 3. All these forms and their associated spaces are to be specified in each particular formulation as follows.

1. Primal trivial formulation: Let  $V = H_0^1(\Omega) \cap C^1(\Omega), W = L^2(\Omega)$ . Find  $u \in V$  satisfying (3) with

$$\begin{aligned} b(w, u) &:= b_1(w, u) = (w, -\nabla \cdot (\kappa \nabla u - \beta u) + \gamma u)_\Omega, \\ \ell(w) &:= \ell_1(w) = (w, f)_\Omega. \end{aligned} \tag{5}$$

2. Primal classical (FEM) formulation: Let  $V = H_0^1(\Omega), W = H_0^1(\Omega)$ . Find  $u \in V$  satisfying (3) with

$$\begin{aligned} b(w, u) &:= b_2(w, u) = (\nabla w, \kappa \nabla u - \beta u)_\Omega + (w, \gamma u)_\Omega, \\ \ell(w) &:= \ell_2(w) = (w, f)_\Omega. \end{aligned} \tag{6}$$

3. Mixed trivial formulation I: Let  $V^u = H_0^1(\Omega), V^q = H(\text{div}, \Omega), W^u = L^2(\Omega), W^q = (L^2(\Omega))^d$ . Find  $(u, \mathbf{q}) \in V^u \times V^q$  satisfying (4) with

$$\begin{aligned} b((w, \mathbf{p}), (u, \mathbf{q})) &:= b_3((w, \mathbf{p}), (u, \mathbf{q})) \\ &= (w, -\nabla \cdot \mathbf{q} + \gamma u)_\Omega + (\mathbf{p}, \mathbf{q} - \kappa \nabla u + \beta u)_\Omega, \\ \ell((w, \mathbf{p})) &:= \ell_3((w, \mathbf{p})) = (w, f)_\Omega + (\mathbf{0}, \mathbf{q})_\Omega = (w, f)_\Omega. \end{aligned} \tag{7}$$

4. Mixed classical formulation II: Let  $V^u = L^2(\Omega), V^q = H(\operatorname{div}, \Omega), W^u = L^2(\Omega), W^q = H(\operatorname{div}, \Omega)$ . Find  $(u, \mathbf{q}) \in V^u \times V^q$  satisfying (4) with

$$\begin{aligned} b((w, \mathbf{p}), (u, \mathbf{q})) &:= b_4((w, \mathbf{p}), (u, \mathbf{q})) \\ &= (w, -\nabla \cdot \mathbf{q} + \gamma u)_\Omega + (\mathbf{p}, \mathbf{q} + \beta u)_\Omega + (\nabla \cdot (\kappa \mathbf{p}), u)_\Omega, \quad (8) \\ \ell((w, \mathbf{p})) &:= \ell_4((w, \mathbf{p})) = (w, f)_\Omega + (\mathbf{0}, \mathbf{q})_\Omega = (w, f)_\Omega. \end{aligned}$$

5. Mixed classical formulation: Let  $V^u = H_0^1(\Omega), V^q = (L^2(\Omega))^d, W^u = H_0^1(\Omega), W^q = (L^2(\Omega))^d$ . Find  $(u, \mathbf{q}) \in V^u \times V^q$  satisfying (4) with

$$\begin{aligned} b((w, \mathbf{p}), (u, \mathbf{q})) &:= b_5((w, \mathbf{p}), (u, \mathbf{q})) \\ &= (\nabla w, \mathbf{q})_\Omega + (w, \gamma u)_\Omega + (\mathbf{p}, \mathbf{q} - \kappa \nabla u + \beta u)_\Omega, \quad (9) \\ \ell((w, \mathbf{p})) &:= \ell_5((w, \mathbf{p})) = (w, f)_\Omega + (\mathbf{0}, \mathbf{q})_\Omega = (w, f)_\Omega. \end{aligned}$$

6. Mixed ultraweak formulation: Let  $V^u = L^2(\Omega), V^q = (L^2(\Omega))^d, W^u = H_0^1(\Omega), W^q = H(\operatorname{div}, \Omega)$ . Find  $(u, \mathbf{q}) \in V^u \times V^q$  satisfying (4) with

$$\begin{aligned} b((w, \mathbf{p}), (u, \mathbf{q})) &:= b_6((w, \mathbf{p}), (u, \mathbf{q})) \\ &= (\nabla w, \mathbf{q})_\Omega + (w, \gamma u)_\Omega + (\mathbf{p}, \mathbf{q} + \beta u)_\Omega + (\nabla \cdot (\kappa \mathbf{p}), u)_\Omega, \\ \ell((w, \mathbf{p})) &:= \ell_6((w, \mathbf{p})) = (w, f)_\Omega + (\mathbf{0}, \mathbf{q})_\Omega = (w, f)_\Omega. \end{aligned} \quad (10)$$

Herein, the primal trivial formulation reduces to the isogeometric collocation method when applying constant test functions; see [37].

### 3 Various Isogeometric Formulations

In this section, we present the isogeometric formulations at the discrete level for both (3) and (4). We first specify the basis functions for all the finite element spaces associated with the mesh configuration  $\mathcal{T}_h$ . For this purpose, we use the Cox-de Boor recursion formula [15, 34] on each dimension and then take tensor-product to obtain the necessary basis functions for multiple dimensions. The Cox-de Boor recursion formula generates  $C^k$  and  $p$ -th order B-spline basis functions, where  $p = 0, 1, 2, \dots$ , and  $k = 0, 1, \dots, p-1$ .  $C^k$  for  $k = 0, 1, \dots, p-1$  basis functions generate a finite-dimensional subspace of the  $H^1(\Omega)$  and  $H(\operatorname{div}, \Omega)$ . We also consider the polynomials which generate the broken space. These finite-dimensional subspaces consist of piecewise polynomials of order  $p = 0, 1, 2, \dots$  and of continuity order  $k = 0, 1, \dots, p-1$ .

The definition of the B-spline basis functions in one dimension is as follows. Let  $X = \{x_0, x_1, \dots, x_m\}$  be a knot vector with knots  $x_j$ , that is, a nondecreasing sequence of real numbers which are called knots. The  $j$ -th B-spline basis function of degree  $p$ , denoted as  $\theta_p^j(x)$ , is defined as [15, 34]

$$\begin{aligned} \theta_0^j(x) &= \begin{cases} 1, & \text{if } x_j \leq x < x_{j+1} \\ 0, & \text{otherwise} \end{cases} \\ \theta_p^j(x) &= \frac{x - x_j}{x_{j+p} - x_j} \theta_{p-1}^j(x) + \frac{x_{j+p+1} - x}{x_{j+p+1} - x_{j+1}} \theta_{p-1}^{j+1}(x). \end{aligned} \quad (11)$$

We then construct the finite-dimensional subspaces using the B-splines on uniform tensor-product meshes with non-repeating and repeating knots. For a  $p$ -th order B-spline, a repetition of  $k = 0, 1, \dots, p$  times of an internal node results in a function of  $C^{p-1-k}$  continuity; see [12, 34]. These B-spline basis functions characterize the finite-dimensional subspaces. For example,

$$V_{p,k}^h = \begin{cases} S_k^p = \text{span}\{\theta_j^p(x)\}_{j=1}^{N_x}, & \text{in 1D} \\ S_{k,k}^{p,p} = \text{span}\{\theta_i^p(x)\theta_j^p(y)\}_{i,j=1}^{N_x, N_y}, & \text{in 2D} \\ S_{k,k,k}^{p,p,p} = \text{span}\{\theta_i^p(x)\theta_j^p(y)\theta_l^p(z)\}_{i,j,l=1}^{N_x, N_y, N_z}, & \text{in 3D} \end{cases} \quad (12)$$

$$V_{p,k}^{q,h} = \begin{cases} S_{k-1}^{p-1}, & \text{in 1D} \\ S_{k-1,k-1}^{p-1,p-1} \times S_{k-1,k-1}^{p-1,p-1}, & \text{in 2D} \\ S_{k-1,k-1}^{p-1,p-1} \times S_{k-1,k-1}^{p-1,p-1} \times S_{k-1,k-1}^{p-1,p-1}, & \text{in 3D} \end{cases}$$

where  $p$  and  $k$  specify the approximation order and continuity order in each dimension (they can be different in general), respectively.  $N_x, N_y, N_z$  are the total numbers of basis functions in each dimension. Note that the space  $V_{p,k}^{q,h}$  collapse to standard Raviart-Thomas mixed finite elements [36]; see [23, Section 5] or [3, Section 3] for more details on the construction of these spaces using B-splines.

We now adopt the mixed formulation for (3) at discrete level: Find  $(u^h, \phi^h) \in V_{p,k}^h \times W_{q,l}^h$  such that

$$\begin{aligned} g(w^h, \phi^h) + b(w^h, u^h) &= \ell(w^h), & \forall w^h \in W_{q,l}^h, \\ b(\phi^h, v^h) &= 0, & \forall v^h \in V_{p,k}^h, \end{aligned} \quad (13)$$

where the spaces are chosen such that  $\dim(W_{q,l}^h) \geq \dim(V_{p,k}^h)$ . Herein,  $\phi^h$  is the negative of the Riesz representation of the residual. The auxiliary bilinear form  $g(\cdot, \cdot)$  is an inner product and it produces a Gramm matrix for the purpose of residual minimization. We define it generally as follows

$$g(v, w) = \sum_{K \in \mathcal{T}_h} \tau_0(v, w)_K + \tau_1 h^{\iota_1} (\nabla v, \nabla w)_K + \tau_2 h^{\iota_2} (\Delta v, \Delta w)_K, \quad \forall v, w \in W_{q,l}^h, \quad (14)$$

where  $\tau_i, \iota_j \in \mathbb{R}, i = 0, 1, 2, j = 1, 2$  are free parameters. The default setting is  $\tau_0 = 1, \tau_1 = 1, \tau_2 = 0, \iota_1 = 2, \iota_2 = 0$ . Once one specifies an inner product  $g(\cdot, \cdot)$ , then under inf-sup assumption on the discrete bilinear formulation in (13), the approximate solution  $u^h$  has a minimal error in the energy norm induced from  $g(\cdot, \cdot)$ ; see, for example [18, 19].

Similarly, we present the mixed formulation for (4) at discrete level: Find  $(u^h, \mathbf{q}^h) \in V_{p_1, k_1}^{u,h} \times V_{p_2, k_2}^{q,h}$  and  $(\phi^h, \boldsymbol{\psi}^h) \in W_{q_1, l_1}^{w,h} \times W_{q_2, l_2}^{q,h}$  such that

$$\begin{aligned} g((w^h, \mathbf{p}^h), (\phi^h, \boldsymbol{\psi}^h)) + b((w^h, \mathbf{p}^h), (u^h, \mathbf{q}^h)) &= \ell((w^h, \mathbf{p}^h)), \\ b((\phi^h, \boldsymbol{\psi}^h), (v^h, \mathbf{r}^h)) &= 0 \end{aligned} \quad (15)$$

with  $(w^h, \mathbf{p}^h) \in W_{q_1, l_1}^{w,h} \times W_{q_2, l_2}^{q,h}, (v^h, \mathbf{r}^h) \in V_{p_1, k_1}^{u,h} \times V_{p_2, k_2}^{q,h}$ , where the spaces are chosen such that  $\dim(W_{q_1, l_1}^{w,h}) \geq \dim(V_{p_1, k_1}^{u,h})$  and  $\dim(W_{q_2, l_2}^{q,h}) \geq \dim(V_{p_2, k_2}^{q,h})$ .

Herein, the continuity orders corresponding to solution  $u$  and flux  $\mathbf{q}$  can be different. Potentially, their polynomial orders can also be different. Similarly,  $g((\cdot, \cdot), (\cdot, \cdot))$  is an inner product for the purpose of residual minimization. We define the Gramm product generally as follows, for  $(v, \mathbf{r}), (w, \mathbf{p}) \in W_{q_1, l_1}^{w, h} \times W_{q_2, l_2}^{q, h}$ ,

$$g((v, \mathbf{r}), (w, \mathbf{p})) = \sum_{K \in \mathcal{T}_h} \tau_3(v, w)_K + \tau_4 h^{\ell_3} (\nabla v, \nabla w)_K + \tau_5(\mathbf{r}, \mathbf{p})_K + \tau_6 h^{\ell_4} (\nabla \cdot \mathbf{r}, \nabla \cdot \mathbf{p})_K, \tag{16}$$

where  $\tau_i, \ell_j \in \mathbb{R}, i = 3, 4, \dots, 6, j = 3, 4$  are free parameters. The default setting is  $\tau_3 = 1, \tau_4 = 1, \tau_5 = 1, \tau_6 = 1, \ell_3 = 2, \ell_4 = 2$ .

Within these formulations, once we specify all free parameters, the bilinear and linear forms, and space settings, we have a different method. We present the following discrete variational formulations

1. Let  $b(\cdot, \cdot) = b_1(\cdot, \cdot), \ell(\cdot) = \ell_1(\cdot)$  defined in (5). Let  $V_{p, k}^h$  consist of B-spline basis functions of continuity  $C^k, k \geq 1$  and  $W_{q, l}^h$  consist of discontinuous basis functions. The *discrete primal trivial formulation* is: Find  $(u^h, \phi^h) \in V_{p, k}^h \times W_{q, l}^h$  satisfying (13).
2. Let  $b(\cdot, \cdot) = b_2(\cdot, \cdot), \ell(\cdot) = \ell_2(\cdot)$  defined in (6). Let  $V_{p, k}^h$  and  $W_{q, l}^h$  consist of B-spline basis functions of continuity at least  $C^0$ . The *discrete primal classical formulation* is: Find  $(u^h, \phi^h) \in V_{p, k}^h \times W_{q, l}^h$  satisfying (13).
3. Let  $b(\cdot, \cdot) = b_3(\cdot, \cdot), \ell(\cdot) = \ell_3(\cdot)$  defined in (7). Let the solution space consist of B-spline basis functions of continuity at least  $C^0$  while the test space consist of discontinuous basis functions. The *discrete mixed trivial formulation* is: Find  $(u^h, \mathbf{q}^h) \in V_{p_1, k_1}^{u, h} \times V_{p_2, k_2}^{q, h}$  and  $(\phi^h, \psi^h) \in W_{q_1, l_1}^h \times W_{q_2, l_2}^{q, h}$  satisfying (15).
4. Let  $b(\cdot, \cdot) = b_4(\cdot, \cdot), \ell(\cdot) = \ell_4(\cdot)$  defined in (8). Let the solution space and test space for flux consist of B-spline basis functions of continuity at least  $C^0$  while the test space for  $u$  consist of discontinuous basis functions. The *discrete mixed classical formulation I* is: Find  $(u^h, \mathbf{q}^h) \in V_{p_1, k_1}^{u, h} \times V_{p_2, k_2}^{q, h}$  and  $(\phi^h, \psi^h) \in W_{q_1, l_1}^h \times W_{q_2, l_2}^{q, h}$  satisfying (15).
5. Let  $b(\cdot, \cdot) = b_5(\cdot, \cdot), \ell(\cdot) = \ell_5(\cdot)$  defined in (9). Let the solution space and test space for  $u$  consist of B-spline basis functions of continuity at least  $C^0$  while the test space for flux consist of discontinuous basis functions. The *discrete mixed classical formulation II* is: Find  $(u^h, \mathbf{q}^h) \in V_{p_1, k_1}^{u, h} \times V_{p_2, k_2}^{q, h}$  and  $(\phi^h, \psi^h) \in W_{q_1, l_1}^{w, h} \times W_{q_2, l_2}^{q, h}$  satisfying (15).
6. Let  $b(\cdot, \cdot) = b_6(\cdot, \cdot), \ell(\cdot) = \ell_6(\cdot)$  defined in (10). Let the solution and test spaces consist of B-spline basis functions of continuity at least  $C^0$ . The *discrete mixed ultraweak formulation* is: Find  $(u^h, \mathbf{q}^h) \in V_{p_1, k_1}^{u, h} \times V_{p_2, k_2}^{q, h}$  and  $(\phi^h, \psi^h) \in W_{q_1, l_1}^{w, h} \times W_{q_2, l_2}^{q, h}$  satisfying (15).

The primal classical formulation can be reduced to the standard finite element and isogeometric element methods. If we set  $k = 0, l = 0$  for  $p \geq 1$ , then we have  $\phi^h = 0$  due to the orthogonal condition (second equation) in (13). Thus, the method reduces to finite element method. Similarly, if we set  $k = l = 1, \dots, p - 1$

for  $p \geq 2$ , then we have  $\phi^h = 0$  in (13) as well and the method reduces to isogeometric analysis. For all other discrete variational formulations, we may constrain the solution and test spaces in such a way that the variational forms we discuss render standard discretization techniques when the Riesz representation of the residual is identically zero.

For the scenarios where we use different trial and test spaces, we obtain a non-zero discrete representation of the residual  $\Phi$ , which we use as an error estimator to guide the refinements of the meshes accordingly. The error estimators are defined in the sense of  $G$  which is a result of the bilinear form  $g(\cdot, \cdot)$ . We refer to the DPG work [17, 18, 20, 38] for more details in this direction.

These discrete variational formulations lead to a linear matrix problem

$$\begin{bmatrix} G & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} \Phi \\ U \end{bmatrix} = \begin{bmatrix} L \\ \mathbf{0} \end{bmatrix}, \tag{17}$$

where  $L$  is the corresponding forcing term arising from the linear form  $\ell(\cdot)$ ,  $\Phi, U$  are the solution pairs for the mixed formulations,  $G$  represents the Gramm product in which we minimize the residual that arises from the bilinear form  $g(\cdot, \cdot)$  and  $B$  is the matrix arising from the bilinear form  $b(\cdot, \cdot)$ . We solve the first equation in (17) and substitute to the second equation in (17) to obtain

$$B^T G^{-1} (L - BU) = B^T G^{-1} L - B^T G^{-1} BU = \mathbf{0}. \tag{18}$$

Herein,  $BU - L$  is the residual and (18) is a least-square type of problem.

At continuous level, the bilinear forms defined in the formulations (6)–(10) satisfy simultaneously the inf-sup condition, that is,

$$\inf_{u \neq 0} \sup_{w \neq 0} \frac{|b(w, u)|}{\|w\|_W \|u\|_V} \geq \alpha > 0, \tag{19}$$

where  $\alpha$  is a constant; see [16, 19]. The discrete inf-sup condition depends on the spaces. Using the coercivity of  $b(\cdot, \cdot)$ , the discrete primal classical formulation with  $q = p, 0 \leq l \leq k$  (hence  $V_{p,k}^h \subset W_{q,l}^h$ ) satisfies

$$\inf_{V_{p,k}^h \ni u^h \neq 0} \sup_{W_{q,l}^h \ni w^h \neq 0} \frac{|b(w^h, u^h)|}{\|w\|_W \|u\|_V} \geq \inf_{V_{p,k}^h \ni u^h \neq 0} \sup_{V_{p,k}^h \ni w^h \neq 0} \frac{|b(w^h, u^h)|}{\|w\|_V \|u\|_V} \geq \alpha_0, \tag{20}$$

where  $\alpha_0 > 0$  is a constant. The proof of discrete inf-sup condition for other formulations is more involved and we omit it here.

For all the discrete variational formulations, we verify the following optimal error convergence rates numerically in the Sect. 4:

$$|u - u^h|_{1,\Omega} \leq Ch^p, \tag{21}$$

where  $C$  is a constant independent of  $h$ . For the approximate fluxes, we define the following errors in  $L^2$  norm

$$\|\mathbf{q} - \mathbf{q}^h\|_{0,\Omega} = \|(\kappa \nabla u - \beta u) - \mathbf{q}^h\|_{0,\Omega}. \tag{22}$$



The optimal convergence rate is

$$\|\mathbf{q} - \mathbf{q}^h\|_{0,\Omega} \leq Ch^{p+1}, \quad (23)$$

where  $C$  is a constant independent of  $h$ .

## 4 Numerical Experiments

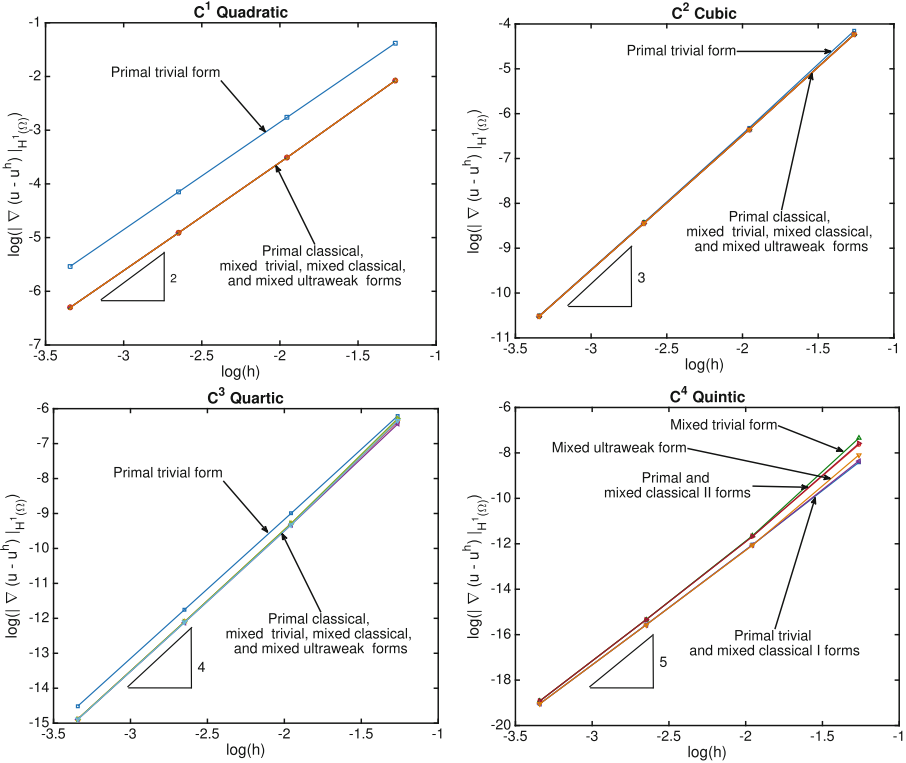
The main result of these numerical tests is that the various formulations we discuss above are equivalent in the sense of resulting the same (optimal) error convergence rates. We focus on 2D and consider the problem (1) with  $\kappa = 1, \gamma = 1, \boldsymbol{\beta} = (1, 1)^T$  and a manufactured solution

$$u(x, y) = \sin(\pi x) \sin(\pi y)(2 - x + 3y). \quad (24)$$

$f$  is the corresponding forcing satisfying (1). The true flux  $\mathbf{q}$  is calculated from (2). We apply all six variational formulations, namely, two primal and four mixed formulations, to solve this problem.

Both the primal and mixed trivial formulations do not involve any integration by parts in their variational formulations and their test functions are in  $L^2$ . The difference is that the primal formulation results from (1) while the mixed formulation results from (2). We use  $C^{-1}$  basis functions for the test spaces for these formulations. Consequently, we obtain a matrix  $G$  in (17) which is a block-diagonal matrix. The independence of each elemental block from the rest allows these methods to be computable efficiently (cheap elemental inversion) and thus relevant for practical purposes. All other formulations involve integration by parts to pass derivatives to the test functions, which in return results in a matrix  $G$  in (17) which is not block-diagonal. This makes these formulations interesting from the theoretical point of view, but untractable in most general meshes, even though splitting schemes make these methods viable on tensor-product meshes [29]. Thus, once we show that all these formulations deliver equivalent results at the discrete level, we focus on these strong/trivial formulations as they are computationally advantageous. We chose not to compare against well established DPG technologies to brake test spaces, as the goal is simply to show the equivalence of the different variational forms rather than derive alternative computable methods. Lastly, to compare the primal trivial formulation with the first-order system least-square (FOSLS) method [4], the difference is that the primal trivial formulation does not lead to first order system and it introduces a Gramm matrix to solve for the residual errors simultaneously.

Figure 1 shows the  $H^1$  semi-norm errors when using all six formulations for  $p = 2, 3, 4, 5$ . For the mixed formulations, we approximate the flux  $\mathbf{q}^h$  using basis functions of the same polynomial order as for the solution  $u^h$ . The parameters of the bilinear form  $g(\cdot, \cdot)$  are set to be the default values. The mesh configurations are  $5 \times 5, 10 \times 10, 20 \times 20, 40 \times 40$ . Herein, we plot the errors in natural logarithmic scale. As predicted, in all scenarios, the  $H^1$  semi-norm errors converge in optimal rates, that is, order  $p$  for all the variational formulations. This confirms the

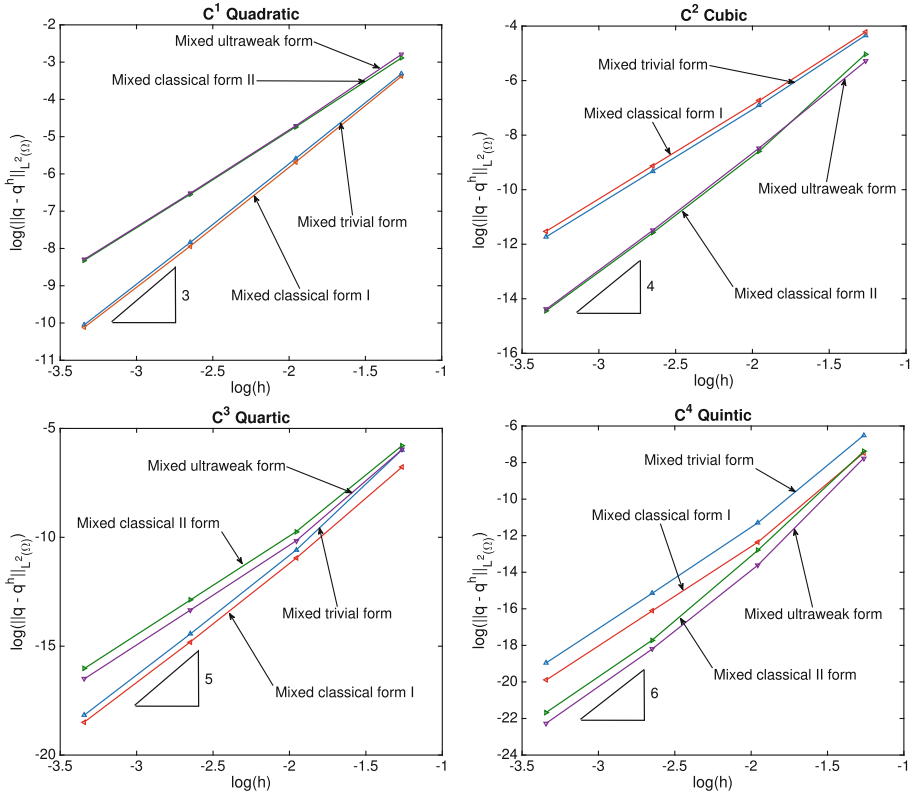


**Fig. 1.**  $H^1$  semi-norm errors when using six formulations with isogeometric elements.

theoretical result (21). Therefore, all formulations are equivalent in the quality of the approximation they deliver. Interestingly, all primal trivial forms for even orders have optimal convergence rates, but the constants seem to be worse for this discretization for even orders than all the other weak forms we compared against. Nevertheless, for odd order polynomials both the rate of convergence and constant are comparable to those observed for all the other variational forms.

The mixed formulations introduce auxiliary variables for the fluxes. Thus, for the same mesh configuration, the resulting matrix systems of the mixed formulations have dimensions which are three times larger than the primal ones. However, the mixed formulations deliver more accurate fluxes approximations.

Figure 2 shows the flux errors in the  $L^2$  norm when using the mixed forms for  $p = 2, 3, 4, 5$ . The fluxes are of optimal convergence rates  $p + 1$ , which confirms the error estimate (23). For  $p = 2, 4$ , we observe super-convergence rates approximately  $p + 2$ . Once again, these results numerically verify that these four formulations are equivalent. Therefore, for the following numerical results, we only show the results for the strong primal and mixed formulations. Finally, we note that the formulations with different choices of the auxiliary bilinear forms



**Fig. 2.**  $L^2$  errors of  $q^h$  when using mixed formulations with isogeometric elements.

$g(\cdot, \cdot)$  in (14) and (16) also lead to optimal error convergence rates. Other inner products are also possible. The convergence rates in the  $H^1$  semi-norm of all these scenarios are optimal.

## 5 Concluding Remarks

We introduce a residual minimization based mixed formulation for solving partial differential equations. A key feature of this method is the framework that uses highly-continuous B-splines for the trial spaces and basis functions with minimal regularity and lower order polynomials for the test spaces. The method shares with the discontinuous Petrov-Galerkin methodology the idea of stabilizing the formulation considering an adequate norm for the test space and unifies the interpretation of several methods such as the classical finite element method, isogeometric analysis, and isogeometric collocation methods. Under the standard assumption, the proposed variational formulations are stable and result in optimal approximation properties.

**Acknowledgement.** This publication was made possible in part by the CSIRO Professorial Chair in Computational Geoscience at Curtin University and the Deep Earth Imaging Enterprise Future Science Platforms of the Commonwealth Scientific Industrial Research Organisation, CSIRO, of Australia. The J. Tinsley Oden Faculty Fellowship Research Program at the Institute for Computational Engineering and Sciences (ICES) of the University of Texas at Austin has partially supported the visits of VMC to ICES. This work was partially supported by the European Union’s Horizon 2020 Research and Innovation Program of the Marie Skłodowska-Curie grant agreement No. 777778, the Mega-grant of the Russian Federation Government (N 14.Y26.31.0013), the Institute for Geoscience Research (TIGeR), and the Curtin Institute for Computation.

## References

1. Auricchio, F., Da Veiga, L.B., Hughes, T., Reali, A., Sangalli, G.: Isogeometric collocation methods. *Math. Models Methods Appl. Sci.* **20**, 2075–2107 (2010)
2. Bazilevs, Y., Beirão da Veiga, L., Cottrell, J.A., Hughes, T.J.R., Sangalli, G.: Isogeometric analysis: approximation, stability and error estimates for h-refined meshes. *Math. Models Methods Appl. Sci.* **16**, 1031–1090 (2006)
3. Buffa, A., De Falco, C., Sangalli, G.: Isogeometric analysis: new stable elements for the stokes equation. *Int. J. Numer. Methods Fluids* (2010)
4. Cai, Z., Lazarov, R., Mantueffel, T.A., McCormick, S.F.: First-order system least squares for second-order partial differential equations: Part I. *SIAM J. Numer. Anal.* **31**, 1785–1799 (1994)
5. Calo, V.M., Collier, N.O., Niemi, A.H.: Analysis of the discontinuous petrov-galerkin method with optimal test functions for the reissner-mindlin plate bending model. *Comput. Math. Appl.* **66**, 2570–2586 (2014)
6. Calo, V.M., Deng, Q., Puzyrev, V.: Dispersion optimized quadratures for isogeometric analysis, arXiv preprint [arXiv:1702.04540](https://arxiv.org/abs/1702.04540) (2017)
7. Calo, V.M., Romkes, A., Valseth, E.: Automatic variationally stable analysis for FE computations: an introduction, arXiv preprint [arXiv:1808.01888](https://arxiv.org/abs/1808.01888) (2018)
8. Cohen, A., Dahmen, W., Welper, G.: Adaptivity and variational stabilization for convection-diffusion equations. *ESAIM Math. Model. Numer. Anal.* **46**, 1247–1273 (2012)
9. Collier, N., Dalcin, L., Calo, V.M.: On the computational efficiency of isogeometric methods for smooth elliptic problems using direct solvers. *Int. J. Numer. Methods Eng.* **100**, 620–632 (2014)
10. Collier, N., Dalcin, L., Pardo, D., Calo, V.M.: The cost of continuity: performance of iterative solvers on isogeometric finite elements. *SIAM J. Sci. Comput.* **35**, A767–A784 (2013)
11. Collier, N., Pardo, D., Dalcin, L., Paszynski, M., Calo, V.M.: The cost of continuity: a study of the performance of isogeometric finite elements using direct solvers. *Comput. Methods Appl. Mech. Eng.* **213**, 353–361 (2012)
12. Cottrell, J.A., Hughes, T.J.R., Bazilevs, Y.: *Isogeometric Analysis: Toward Integration of CAD and FEA*. Wiley, New York (2009)
13. Cottrell, J.A., Reali, A., Bazilevs, Y., Hughes, T.J.R.: Isogeometric analysis of structural vibrations. *Comput. Methods Appl. Mech. Eng.* **195**, 5257–5296 (2006)
14. Dahmen, W., Huang, C., Schwab, C., Welper, G.: Adaptive Petrov-Galerkin methods for first order transport equations. *SIAM J. Numer. Anal.* **50**, 2420–2445 (2012)
15. De Boor, C.: *A Practical Guide to Splines*, vol. 27. Springer, New York (1978)

16. Demkowicz, L.: Various variational formulations and closed range theorem. ICES Report **15** (2015)
17. Demkowicz, L., Gopalakrishnan, J.: A class of discontinuous Petrov-Galerkin methods. Part I: the transport equation. *Comput. Methods Appl. Mech. Eng.* **199**, 1558–1572 (2010)
18. Demkowicz, L., Gopalakrishnan, J.: A class of discontinuous Petrov-Galerkin methods. II. Optimal test functions. *Numer. Methods Partial Differ. Eqn.* **27**, 70–105 (2011)
19. Demkowicz, L., Gopalakrishnan, J.: Discontinuous Petrov-Galerkin (DPG) method, *Encyclopedia of Computational Mechanics Second Edition*, pp. 1–15 (2017)
20. Demkowicz, L., Gopalakrishnan, J., Niemi, A.H.: A class of discontinuous Petrov-Galerkin methods. Part III: adaptivity. *Appl. Numer. Math.* **62**, 396–427 (2012)
21. Deng, Q., Bartoň, M., Puzyrev, V., Calo, V.: Dispersion-minimizing quadrature rules for C1 quadratic isogeometric analysis. *Comput. Methods Appl. Mech. Eng.* **328**, 554–564 (2018)
22. Deng, Q., Calo, V.: Dispersion-minimized mass for isogeometric analysis. *Comput. Methods Appl. Mech. Eng.* **341**, 71–92 (2018)
23. Evans, J.A., Hughes, T.J.: Isogeometric divergence-conforming B-splines for the Darcy-Stokes-Brinkman equations. *Mathe. Models Methods Appl. Sci.* **23**, 671–741 (2013)
24. Garcia, D., Pardo, D., Dalcin, L., Calo, V.M.: Refined isogeometric analysis for a preconditioned conjugate gradient solver. *Comput. Methods Appl. Mech. Eng.* **335**, 490–509 (2018)
25. Garcia, D., Pardo, D., Dalcin, L., Paszyński, M., Collier, N., Calo, V.M.: The value of continuity: refined isogeometric analysis and fast direct solvers. *Comput. Methods Appl. Mech. Eng.* **316**, 586–605 (2017)
26. Hughes, T.J.R., Cottrell, J.A., Bazilevs, Y.: Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Comput. Methods Appl. Mech. Eng.* **194**, 4135–4195 (2005)
27. Hughes, T.J.R., Evans, J.A., Reali, A.: Finite element and NURBS approximations of eigenvalue, boundary-value, and initial-value problems. *Comput. Methods Appl. Mech. Eng.* **272**, 290–320 (2014)
28. Hundsdorfer, W., Verwer, J.: *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations. Computational Mathematics*, vol. 33. Springer, Heidelberg (2003)
29. Łoś, M., Deng, Q., Muga, I., Calo, V., Paszyński, M.: Isogeometric residual minimization method (iGRM) with direction splitting for implicit problems (2018, in preparation)
30. Niemi, A., Collier, N., Calo, V.M.: Discontinuous Petrov-Galerkin method based on the optimal test space norm for one-dimensional transport problems. *Procedia Comput. Sci.* (2011)
31. Niemi, A.H., Collier, N.O., Calo, V.M.: Automatically stable discontinuous Petrov-Galerkin methods for stationary transport problems: quasi-optimal test space norm. *Comput. Math. Appl.* **66**, 2096–2113 (2013)
32. Niemi, A.H., Collier, N.O., Calo, V.M.: Discontinuous Petrov-Galerkin method based on the optimal test space norm for steady transport problems in one space dimension. *J. Comput. Sci.* **4**, 157–163 (2013)
33. Pardo, D., Paszynski, M., Collier, N., Alvarez, J., Dalcin, L., Calo, V.M.: A survey on direct solvers for Galerkin methods. *SeMA J.* **57**, 107–134 (2012)

34. Piegl, L., Tiller, W.: The NURBS Book. Springer, Heidelberg (1997). <https://doi.org/10.1007/978-3-642-97385-7>
35. Puzyrev, V., Deng, Q., Calo, V.M.: Dispersion-optimized quadrature rules for isogeometric analysis: modified inner products, their dispersion properties, and optimally blended schemes. *Comput. Methods Appl. Mech. Eng.* **320**, 421–443 (2017)
36. Raviart, P.A., Thomas, J.M.: A mixed finite element method for 2-nd order elliptic problems. In: Galligani, I., Magenes, E. (eds.) *Mathematical Aspects of Finite Element Methods*. LNM, vol. 606, pp. 292–315. Springer, Heidelberg (1977). <https://doi.org/10.1007/BFb0064470>
37. Schillinger, D., Evans, J.A., Reali, A., Scott, M.A., Hughes, T.J.: Isogeometric collocation: cost comparison with Galerkin methods and extension to adaptive hierarchical NURBS discretizations. *Comput. Methods Appl. Mech. Eng.* **267**, 170–232 (2013)
38. Zitelli, J., Muga, I., Demkowicz, L., Gopalakrishnan, J., Pardo, D., Calo, V.M.: A class of discontinuous Petrov-Galerkin methods. Part IV: the optimal test norm and time-harmonic wave propagation in 1D. *J. Comput. Phys.* **230**, 2406–2432 (2011)