Journal of Scientific Computing (2023) 96:15
https://doi.org/10.1007/s10915-023-02246-x

Least-Squares Methods with Nonconforming Finite Elements for General Second-Order Elliptic Equations

Yuxiang Liang[1] · **Shun Zhang**[1]

Received: 22 May 2022 / Revised: 13 February 2023 / Accepted: 11 May 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
In this paper, we study least-squares finite element methods (LSFEM) for general second-order elliptic equations with nonconforming finite element approximations. The equation may be indefinite. For the two-field potential-flux div LSFEM with Crouzeix–Raviart (CR) element approximation, we present three proofs of the discrete solvability under the condition that mesh size is small enough. One of the proof is based on the coerciveness of the original bilinear form. The other two are based on the minimal assumption of the uniqueness of the solution of the second-order elliptic equation. A counterexample shows that div least-squares functional does not have norm equivalence in the sum space of H^1 and CR finite element spaces. Thus it cannot be used as an a posteriori error estimator. Several versions of reliable and efficient error estimators are proposed for the method. We also propose a three-filed potential-flux-intensity div-curl least-squares method with general nonconforming finite element approximations. The norm equivalence in the abstract nonconforming piecewise H^1-space is established for the three-filed formulation on the minimal assumption of the uniqueness of the solution of the second-order elliptic equation. The three-filed div-curl nonconforming formulation thus has no restriction on the mesh size, and the least-squares functional can be used as the built-in a posteriori error estimator. Under some restrictive conditions, we also discuss a potential-flux div-curl least-squares method.

1 Introduction

The least-squares variational principle and the corresponding least-squares finite element methods based on a first-order system reformulation have been widely used in numerical solutions of partial differential equations; see for example [1–10]. Compared to the standard variational formulation and the related finite element methods, the first-order system least-squares finite element methods have several known advantages, such as the discrete problem

✉ Shun Zhang
shun.zhang@cityu.edu.hk

Yuxiang Liang
yuxiliang7-c@my.cityu.edu.hk

[1] Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong SAR, China

Published online: 24 May 2023

is stable without the inf-sup condition of the discrete spaces and mesh size restriction, and the least-squares functional itself is a good built-in a posteriori error estimator.

Since the introduction in the classic 1973 paper of Crouzeix and Raviart [11], the nonconforming finite elements, including the Crouzeix-Raviart (CR) element [12], and various elements introduced in [13–17] are very useful for numerical computation of many physical problems. In [18], least-squares methods with various nonconforming finite element approximations are introduced for the diffusion problem without the lower order terms. In [18], contradicting least-squares methods with conforming approximations, the coerciveness of the discrete problem needs to be proved independently. The discrete problem is coercive for the CR element for the equation without the lower-order terms. For some other elements, the assumption that the mesh size is sufficiently small is needed to ensure the coerciveness in [18]. The a posteriori error estimator is not discussed in [18].

Besides the traditional nonconforming finite elements such as those developed in [11, 13, 15], there are other widely used specially designed nonconforming finite elements in various applications, for example, the multiscale finite elements [19, 20] and the immersed finite elements [21]. This paper is our first attempt to study the least-squares finite element methods in a nonconforming setting. It is hoped that the methods and analysis developed in the paper will be useful in studying the least-squares methods for these non-traditional nonconforming elements.

1.1 Stability Results for Various Finite Element Approximations to Second-Order Linear Elliptic PDEs

In this paper, we plan to extend the least-squares methods with nonconforming finite element approximations to general second-order elliptic equations. We first review the existence and uniqueness result of the general second-order elliptic equation and its various finite element approximations.

For the general linear second-order elliptic Eq. (2.10), the solution's existence and uniqueness can be discussed in two cases. The first case is a simpler coercive case. Assuming that the coefficients satisfy some assumption, the standard bilinear form is coercive in the H^1-norm; see also Section 6.1.2 of [22]. The existence, uniqueness, and stability of the solution can be obtained from the Lax-Milgram Lemma. For its conforming finite element approximation, the method is also coercive.

The second case is the general case, where the bilinear form associated with the PDE can be genuinely indefinite. In the general case, we only need to assume a very mild assumption of the domain and the coefficients; see Assumption 2.1. Furthermore, due to the compactness of the operator, the uniqueness, existence, and well-posedness are equivalent; see discussion in Sect. 2.2. We thus only assume a *minimal assumption: the homogeneous equation has a unique zero solution* to ensure the well-posedness of the problem. The coercive case is a special case of the general case.

The analysis of finite element approximations to the possible indefinite general second-order elliptic equations is not simple and straightforward. In [23, 24], conforming finite element approximations of general elliptic equations are discussed. Recently, in [25, 26], nonconforming and mixed finite element approximations are discussed. The results require some regularity assumption and that the mesh size of the discretization is small enough. On the other hand, once the coerciveness of the least-squares method is proved, see [1, 7, 27–30], the LSFEM with H^1-$H(\mathrm{div})$ conforming finite element approximations is automatic coercive, without the restriction on the mesh size.

The discussion for closed-related discontinuous least-squares finite element methods can be found in [31] for a Div-Curl problem, and [32] for a simple symmetric equation with the convection term.

1.2 Contributions of this Paper

For the CR-LSFEM for general second-order elliptic equations, we first present a two-field potential-flux div least-squares method, which is a direct generalization of the nonconforming method suggested in [18] to general second-order elliptic equations and a generalization of the conforming least-squares method in [1] to the CR approximation.

In the paper, we present a negative result on the norm equivalence of the nonconforming div least-squares functional in the sum space of H^1 and CR finite element spaces; see Lemma 7.1. Thus, the two-field potential-flux div CR-LSFEM does not have the two most important properties of a standard LSFEM: automatic discrete stability without the assumption of the mesh size and a built-in least-squares functional a posteriori error estimator. We need to prove the discrete solvability and develop a posteriori error estimator for the method.

To ensure the solvability of the discrete problem from the two-field CR-LSFEM, we discuss three proofs. The first proof is for the coercive case only. The second and the third proofs are for the general cases under the minimal uniqueness assumption. Our second proof is based on Schatz's argument. We present an L^2-error estimate of the potential-flux div CR-LSFEM based on Cai-Ku's paper [33]. The proof also corrects a small error in the original [33] when handling mixed boundary conditions. In the third proof, we present a very short proof by combining the new proof presented in [30] and the discrete stability result proved in [25] for CR approximation of the general elliptic equation. Different from the conforming LSFEM in [1], in the second and third proofs, the div CR-LSFEM requires the same regularity and sufficient small mesh size assumptions as the non-least-squares (conforming, nonconforming, and mixed) methods [23–26]. In the first proof (coercive case), the restriction of the mesh size is local, and no regularity assumption is needed. The mesh size restriction in the second and third proofs is global since the regularity assumption is global.

For a posteriori error estimator of two-field potential-flux div CR-LSFEM, we suggest several a posteriori error estimators for the potential-flux div CR-LSFEM by adding different terms measuring the nonconforming error. Reliability and efficiency results are proved.

To overcome the shortcoming of the two-field div CR-LSFEM due to the lack of norm equivalence, we suggest three-field formulations, potential-flux-density div-curl least-squares methods. A curl term of the intensity (gradient of the solution) is added to the least-squares formulation. In this new formulation, we prove norm equivalence for the abstract nonconforming piecewise H^1-space. Since the nonconforming space is mesh-dependent, we use Helmholtz decomposition to avoid the coerciveness constant depending on the mesh. The coerciveness of potential-flux-density div-curl least-squares methods in the abstract nonconforming space setting is then proved in the same minimal uniqueness assumption as the standard least-squares formulation without any requirement on the mesh size. Since the norm equivalence in the abstract nonconforming piecewise H^1-space is true, we automatically have the standard built-in least-squares a posteriori error estimator. The only other ingredient needed for the proof is a discrete Poincaré-Friedrichs inequality in the abstract nonconforming piecewise H^1-space. Thus the proof is not only true for the CR element but also true for nonconforming elements introduced in [13–17]. In a sense, we recover the good properties of the LSFEM (non-restriction on mesh size and built-in a posteriori estimator)

over the conforming, nonconforming, and mixed methods with the three-filed nonconforming LSFEM.

The failure and success of the norm equivalence with and without the curl term can be connected to the theory of a posteriori error estimates of nonconforming finite elements. For the nonconforming finite element approximation, besides the standard residual and the fact the numerical flux from the potential u_h is not in the $H(\mathrm{div})$ space (the so-called conforming error), we also need to measure one extra error: the numerical potential u_h is not in the H^1 space or its consequence that the numerical intensity (gradient) is not in the $H(\mathrm{curl})$ space (the so-called nonconforming error). This explains that we need the intensity (gradient) as an independent unknown and add its curl in the potential-flux-density div-curl least-squares method. Previous discussions of a posteriori error estimates of the nonconforming finite element approximation can be found in [34–41], where the different contributions of the error are discussed.

We also discuss the application and restriction of the original potential-flux div-curl least-squares method [2]. When the domain is nice and the coefficient is sufficiently smooth, the original formulation introduced in [2] can be used in the nonconforming case. The norm equivalence can also be established similarly. However, the two-field div-curl formulation can cause serious problems when the conditions on the domain and coefficients are not satisfied.

1.3 Structure of the Paper

The paper is organized as follows. In Sect. 2, we present preliminaries about the abstract and discrete spaces. Properties of the CR space, the Helmholtz decomposition, and the discrete Poincaré-Friedrichs inequality are discussed. The solution theory of a general second-order elliptic equation is also discussed. Sections 3 to 7 are about the two-field potential-flux div CR-LSFEM. In Sect. 3, we introduce the formulation. A discrete coerciveness is proved with the assumption of the coefficients ensuring the coerciveness of the bilinear form of the original variational problem in Sect. 4. We present two proofs of the discrete solvability with minimal uniqueness assumption and a regularity assumption in Sects. 5 and 6. In Sect. 7, we present a counterexample to show that div least-squares functional does not have norm equivalence in the sum space of H^1 and CR finite element spaces and we propose several versions of reliable and efficient error estimators. We propose the potential-flux-intensity div-curl least-squares method with general nonconforming finite element approximations in Sect. 8. The norm equivalence in the abstract nonconforming piecewise H^1-space is established. Under some restrictive conditions, we also discuss the potential-flux div-curl least-squares method in Sect. 9. In Sect. 10, a numerical counterexample to show that div least-squares functional does not have norm equivalence. We also present adaptive numerical tests for methods presented in the paper for a non-self-adjoint and non-coercive problem. Several concluding remarks are made in Sect. 11.

2 Preliminaries

2.1 Notations and the Function Spaces

Let Ω be a bounded, open, connected subset of $\mathbb{R}^d (d = 2 \text{ or } 3)$ with a Lipschitz continuous boundary $\partial\Omega$. We partition the boundary of the domain Ω into two open subsets Γ_D and Γ_N,

such that $\partial\Omega = \overline{\Gamma_D} \cup \overline{\Gamma_N}$ and $\Gamma_D \cap \Gamma_N = \emptyset$. For simplicity, we assume that Γ_D is not empty (i.e., meas(Γ_D) $\neq 0$) and is connected.

We use the standard notations and definitions for the Sobolev spaces $H^s(\Omega)$ for $s \geq 0$. The standard associated inner product is denoted by $(\cdot,\ \cdot)_{s,\Omega}$, and its norm is denoted by $\|\cdot\|_{s,\Omega}$. The notation $|\cdot|_{s,\Omega}$ is used for the semi-norm. (We suppress the superscript d because the dependence on dimension will be clear by context. We also omit the subscript Ω from the inner product and norm designation when there is no risk of confusion.) For $s = 0$, $H^s(\Omega)$ coincides with $L^2(\Omega)$. The symbols $\nabla\cdot$ and ∇ stand for the divergence and gradient operators, respectively. Set $H^1_D(\Omega) := \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\}$, $H^1_N(\Omega) := \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_N\}$, and $H^1_0(\Omega) := \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}$.

In two dimensions, for a vector-valued function $\boldsymbol{\tau} = (\tau_1,\ \tau_2)^t$, define the curl operator by $\nabla\times\boldsymbol{\tau} := \dfrac{\partial\tau_2}{\partial x_1} - \dfrac{\partial\tau_1}{\partial x_2}$. For a scalar-valued function v, define the operator $\nabla^\perp$ by $\nabla^\perp v = (-\dfrac{\partial v}{\partial x_2},\ \dfrac{\partial v}{\partial x_1})$. In three dimensions, $\nabla\times\boldsymbol{\tau}$ is defined standardly for a vector valued function $\boldsymbol{\tau}$. To unify the notation in both dimensions, we define the vector curl as following. Let $k := 1$ if $d = 2$ and $k := 3$ if $d = 3$. The Curl of a function $\mathtt{v} \in \mathbb{R}^k$ is defined by

$$\mathtt{Curl\ v} := \nabla^\perp \mathtt{v} \text{ if } d = 2 \quad \text{and} \quad \mathtt{Curl\ v} := \nabla\times\mathtt{v} \text{ if } d = 3. \tag{2.1}$$

We use a special font for $\mathtt{v}$ that it is a scalar function when $d = 2$ and it is a vector function when $d = 3$. Given a unit normal $\mathbf{n}$ we define the tangential component of a vector $\mathbf{v} \in \mathbb{R}^d$ with respect to $\mathbf{n}$ by

$$\gamma_t(\mathbf{v}) := \begin{cases} \mathbf{v}\cdot\mathbf{t} & \text{if } d = 2, \text{ where } \mathbf{t} = (-n_2, n_1) \text{ if } \mathbf{n} = (n_1, n_2), \\ \mathbf{v}\times\mathbf{n} & \text{if } d = 3. \end{cases} \tag{2.2}$$

We use the standard $H(\text{div};\ \Omega)$ and $H(\text{curl};\ \Omega)$ spaces, equipped with the norms

$$\|\boldsymbol{\tau}\|_{H(\text{div};\ \Omega)} = \left(\|\boldsymbol{\tau}\|^2_{0,\Omega} + \|\nabla\cdot\boldsymbol{\tau}\|^2_{0,\Omega}\right)^{\frac{1}{2}} \quad \text{and} \quad \|\boldsymbol{\tau}\|_{H(\text{curl};\ \Omega)} = \left(\|\boldsymbol{\tau}\|^2_{0,\Omega} + \|\nabla\times\boldsymbol{\tau}\|^2_{0,\Omega}\right)^{\frac{1}{2}},$$

respectively. Denote their subspaces by

$$\begin{aligned} H_N(\text{div};\ \Omega) &= \{\boldsymbol{\tau} \in H(\text{div};\ \Omega) : \boldsymbol{\tau}\cdot\mathbf{n}|_{\Gamma_N} = 0\} \quad \text{and} \quad H_D(\text{curl};\ \Omega) \\ &= \{\boldsymbol{\tau} \in H(\text{curl};\ \Omega) : \gamma_t(\boldsymbol{\tau})|_{\Gamma_D} = 0\}, \end{aligned}$$

where $\mathbf{n}$ is the unit vectors outward normal to the boundary $\partial\Omega$.

Let $\mathcal{T} = \{K\}$ be a triangulation of Ω using simplicial elements. The mesh $\mathcal{T}$ is assumed to be regular. Denote the set of all nodes of the triangulation by $\mathcal{N} := \mathcal{N}_{int} \cup \mathcal{N}_D \cup \mathcal{N}_N$, where $\mathcal{N}_{int}$ is the set of all interior nodes, and $\mathcal{N}_D$ and $\mathcal{N}_N$ are the sets of all boundary nodes belonging to the respective $\overline{\Gamma}_D$ and Γ_N. Denote the set of all faces(3D)/edges(2D) of the triangulation by $\mathcal{E} := \mathcal{E}_{int} \cup \mathcal{E}_D \cup \mathcal{E}_N$, where $\mathcal{E}_{int}$ is the set of all interior element faces/edges and $\mathcal{E}_D$ and $\mathcal{E}_N$ are the sets of all boundary faces/edges belonging to the respective Γ_D and Γ_N. For each $F \in \mathcal{E}$, denote by h_F the diameter of the face/edge F; denote by $\mathbf{n}_F$ a unit vector normal to F. When $F \in \mathcal{E}_D \cup \mathcal{E}_N$, assume that $\mathbf{n}_F$ is the unit outward normal vector. For each interior face/edge $F \in \mathcal{E}_{int}$, let K_F^+ and K_F^- be the two elements sharing the common edge F such that the unit outward normal vector of K_F^- coincides with $\mathbf{n}_F$.

Define jumps and averages over faces/edges by

$$[\![v]\!]_F := \begin{cases} v|_F^- - v|_F^+ & F \in \mathcal{E}_{int}, \\ v|_F & F \in \mathcal{E}_D \\ 0 & F \in \mathcal{E}_N, \end{cases} \quad \text{and} \quad \{v(x)\}_F = \begin{cases} (v_F^- + v_F^+)/2 & F \in \mathcal{E}_{int}, \\ v|_F & F \in \mathcal{E}_D \cup \mathcal{E}_N \end{cases}$$

for all $F \in \mathcal{E}$. A simple calculation leads to the following identity: $[\![uv]\!]_F = \{v\}_F [\![u]\!]_F + \{u\}_F [\![v]\!]_F$, for all $F \in \mathcal{E}_{int}$. Let $P_k(K)$ be the space of polynomials of degree k on element K. Denote the Crouzeix-Raviart nonconforming piecewise linear finite element spac [11, 12] associated with the triangulation $\mathcal{T}$ by

$$V^{cr} = \{v \in L^2(\Omega) : v|_K \in P_1(K) \ \forall \ K \in \mathcal{T} \text{ and } \int_F [\![v]\!] ds = 0 \ \forall \ F \in \mathcal{E}_{int}\}$$

and its subspace by $V_D^{cr} = \{v \in V^{cr} : \int_F v ds = 0 \ \forall \ F \in \mathcal{E}_D\}$. Let

$$W_D^{1+cr} := H_D^1(\Omega) + V_D^{cr} = \{v = v_1 + v_2 : v_1 \in H_D^1(\Omega), v_2 \in V_D^{cr}\}, \tag{2.3}$$

$$W_D(\mathcal{T}) := \{v \in L^2(\Omega) : v|_K \in H^1(K), \forall K \in \mathcal{T} \text{ and } \int_F [\![v]\!] ds = 0 \ \forall \ F \in \mathcal{E}_{int} \cup \mathcal{E}_D\}. \tag{2.4}$$

It is easy to see that $W_D^{1+cr} \subset W_D(\mathcal{T})$. Many other classic nonconforming elements [13–17] belong to $W_D(\mathcal{T})$.

Denote the local lowest-order Raviart-Thomas (RT) [42] on element $K \in \mathcal{T}$ by $RT_0(K) = P_0(K)^d + \mathbf{x} P_0(K)$. Then the standard lowest-order $H(\text{div}; \Omega)$ conforming RT space is defined by

$$RT_{0,N} = \{\boldsymbol{\tau} \in H_N(\text{div}; \Omega) : \boldsymbol{\tau}|_K \in RT_0(K) \ \ \forall \ K \in \mathcal{T}\}.$$

Also, let

$$P_0 = \{v \in L^2(\Omega) : v|_K \in P_0(K) \ \forall \ K \in \mathcal{T}\}.$$

Denote the first type of local lowest order Nédélec space [43] on element $K \in \mathcal{T}$ by

$$N_0(K) := \begin{cases} P_0(K)^2 + (x_2, -x_1) P_0(K) & d = 2, \\ P_0(K)^3 + (0, -x_3, x_2) P_0(K) + (x_3, 0, -x_1) P_0(K) + (-x_2, x_1, 0) P_0(K) & d = 3. \end{cases}$$

Then the lowest order $H(\text{curl}\,; \Omega)$ conforming Nédélec spaces are defined by

$$N_{0,D} = \{\boldsymbol{\tau} \in H_D(\text{curl}; \Omega) : \boldsymbol{\tau}|_K \in N_0(K) \ \ \forall \ K \in \mathcal{T}\}.$$

We define the discrete gradient operator as $(\nabla_h v)|_K := \nabla(v|_K)$, for all $K \in \mathcal{T}$.

For a fixed $r > 0$, denote by $I_{rt} : H(\text{div}; \Omega) \cap [H^r(\Omega)]^d \mapsto RT_0$ the standard RT interpolation operator. We have the following local approximation property: for $\boldsymbol{\tau} \in H^{\ell_K}(K)$, $0 < \ell_K \leq 1$,

$$\|\boldsymbol{\tau} - I_{rt}\boldsymbol{\tau}\|_{0,K} \leq C h_K^{\ell_K} |\boldsymbol{\tau}|_{\ell_K, K} \ \ \forall \ K \in \mathcal{T}, \tag{2.5}$$

$$\|\nabla \cdot (\boldsymbol{\tau} - I_{rt}\boldsymbol{\tau})\|_{0,K} \leq C h_K^{\ell_K} |\nabla \cdot \boldsymbol{\tau}|_{\ell_K, K} \ \ \forall \ K \in \mathcal{T}. \tag{2.6}$$

The estimates in (2.5) and (2.6) are standard for $\ell_K = 1$ and can be proved by the average Taylor series developed in [44] and the standard reference element technique with Piola transformation for $0 < \ell_K < 1$. The interpolations and approximation properties are entirely local.

Similarly, denote by $I_n : H(\text{curl}\,; \Omega) \mapsto N_0$ the standard lowest-order Nédélec interpolation operator. We have the following approximation property: for $\boldsymbol{\tau} \in H^1(\Omega)$ and $\nabla \times \boldsymbol{\tau} \in H^1(\Omega)^d$,

$$\|\boldsymbol{\tau} - I_n\boldsymbol{\tau}\|_0 \leq C h |\boldsymbol{\tau}|_1 \tag{2.7}$$

$$\|\nabla \times (\boldsymbol{\tau} - I_n\boldsymbol{\tau})\|_0 \leq C h_K |\nabla \times \boldsymbol{\tau}|_1. \tag{2.8}$$

Denote by $\theta_F(\mathbf{x})$ the nodal basis function of V_D^{cr} associated with the face $F \in \mathcal{E}$. For $v \in L^1(F)$, define $\Pi_F^0 v = (v, 1)_F/|F|$, the average value of v on F. The local Crouzeix-Raviart interpolant is defined by $I_K^{cr} v = \sum_{F\in\mathcal{E}\cap\partial K}(\Pi_F^0 v)\theta_F(\mathbf{x})$ for $v \in W^{1,1}(K)$. It was shown (see [39, 45]) that for $v \in H^{1+\ell_K}(K)$ with $0 \le \ell_K \le 1$

$$\|v - I_K^{cr} v\|_{0,K} \le C\, h_K^{1+\ell_K} |\nabla v|_{\ell_K,K} \quad \forall\, K \in \mathcal{T}. \tag{2.9}$$

2.2 General Second-Order Elliptic Equations

Consider the general second-order elliptic equation in divergence form

$$-\nabla\cdot(A\nabla u) + Xu = f \text{ in } \Omega, \quad u = 0 \text{ on } \Gamma_D, \quad A\nabla u\cdot\mathbf{n} = 0 \text{ on } \Gamma_N. \tag{2.10}$$

where

$$Xv := \mathbf{b}\cdot\nabla v + cv, \quad \forall v \in H^1(\Omega).$$

The following very mild conditions on the domain and coefficients are assumed.

Assumption 2.1 The domain Ω is a bounded, open, connected subset of $\mathbb{R}^d (d = 2 \text{ or } 3)$ with a Lipschitz continuous boundary $\partial\Omega$. The diffusion coefficient matrix $A \in L^\infty(\Omega)^{d\times d}$ is a given $d \times d$ tensor-valued function; the matrix A is uniformly symmetric positive definite: there exist positive constants $0 < \Lambda_0 \le \Lambda_1$ such that

$$\Lambda_0 \mathbf{y}^T\mathbf{y} \le \mathbf{y}^T A\mathbf{y} \le \Lambda_1 \mathbf{y}^T\mathbf{y} \tag{2.11}$$

for all $\mathbf{y} \in \mathbb{R}^d$ and almost all $x \in \Omega$. The coefficients $\mathbf{b} \in L^\infty(\Omega)^d$ and $c \in L^\infty(\Omega)$ are given vector- and scalar-valued bounded functions, respectively.

The variational problem of (2.10) is: Find $u \in H_D^1(\Omega)$, such that

$$a(u, v) = (f, v) \quad \forall v \in H_D^1(\Omega), \tag{2.12}$$

where the bilinear form a is defined as

$$a(w, v) := (A\nabla w, \nabla v) + (Xw, v) \quad \text{for } w, v \in H_D^1(\Omega). \tag{2.13}$$

It is easy to check that the bilinear form is continuous due to Assumption 2.1:

$$a(w, v) \le C\|\nabla v\|_0\|\nabla w\|_0 \quad v \in H_D^1(\Omega) \text{ and } w \in H_D^1(\Omega). \tag{2.14}$$

We also consider the adjoint problem of (2.10), a PDE in physical form,

$$-\nabla\cdot(A\nabla z + \mathbf{b}z) + cz = g \text{ in } \Omega, \quad z = 0 \text{ on } \Gamma_D, \quad (A\nabla z + \mathbf{b}z)\cdot\mathbf{n} = 0 \text{ on } \Gamma_N. \tag{2.15}$$

With the help of integrations by parts, it is easy to check that the variational problem of (2.15) is: Find $z \in H_D^1(\Omega)$, such that

$$a(v, z) = (g, v) \quad \forall v \in H_D^1(\Omega). \tag{2.16}$$

Thus (2.15) is the adjoint problem of (2.10).

Remark 2.2 When $\mathbf{b} = 0$ and $c = -k^2$ for some $k > 0$, the equation is a real Helmholtz equation. The equation is indefinite. However, as long as k^2 square is not an eigenvalue of $(A\nabla v, \nabla w)$, it still has a unique solution, see for example, discussions in [30].

To discuss the solution theory of linear second-order elliptic equations, we consider two slightly more general problems with righthand sides in $(H_D^1(\Omega))'$, the dual space of $H_D^1(\Omega)$.

For a $\phi \in (H_D^1(\Omega))'$, assume that $u \in H_D^1(\Omega)$ is the solution of the weak problem of the equation in divergence form:

$$\text{Find } u \in H_D^1(\Omega), \text{ such that } a(u, v) = \langle \phi, v \rangle_{(H_D^1(\Omega))' \times H_D^1(\Omega)} \quad \forall v \in H_D^1(\Omega), \tag{2.17}$$

or the adjoint weak problem of the equation in physical form, a $\psi \in (H_D^1(\Omega))'$

$$\text{Find } z \in H_D^1(\Omega), \text{ such that } a(v, z) = \langle \psi, v \rangle_{(H_D^1(\Omega))' \times H_D^1(\Omega)} \quad \forall v \in H_D^1(\Omega). \tag{2.18}$$

For a possible indefinite linear second-order elliptic equation, the solution's existence and uniqueness theory is based on the Fredholm alternative. Since we assume the ellipticity of the PDEs (conditions (2.11) on A), the operators associated with the divergence form and adjoint physical form problems are Fredholm operators of index zero. The uniqueness, existence, and well-posedness are equivalent for the linear second-order elliptic equation, for example, see discussions in various standard PDE books [46–48]. A detailed discussion of the following theorem with two proofs can also be found in Theorem 2.3 of [30]. Some more discussion can also be found in the introductions of [25, 26].

Theorem 2.3 *Assume Assumption 2.1 is true. The following assumptions are equivalent:*

(1) The homogeneous equation $a(u, v) = 0$, for all $v \in H_D^1(\Omega)$, has $u = 0$ as its unique solution.

(2) The weak problem (2.17) has the following stability bound:

$$\|\nabla u\|_0 \le C \|\phi\|_{(H_D^1(\Omega))'} \quad \forall \phi \in (H_D^1(\Omega))'. \tag{2.19}$$

(3) The weak problem (2.17) has a unique solution $u \in H_D^1(\Omega)$ for any $\phi \in (H_D^1(\Omega))'$.

(4) The homogeneous equation associated to (2.18), i.e., $a(v, z) = 0$, for all $v \in H_D^1(\Omega)$, has $z = 0$ as its unique solution.

(5) The adjoint weak problem (2.18) has the following stability bound:

$$\|\nabla z\|_0 \le C \|\psi\|_{(H_D^1(\Omega))'} \quad \forall \psi \in (H_D^1(\Omega))'. \tag{2.20}$$

(6) The adjoint weak problem (2.18) has a unique solution $z \in H_D^1(\Omega)$ for any $\psi \in (H_D^1(\Omega))'$.

In this paper, except for the case in Sect. 4, we only assume that one of the conditions in Theorem 2.3 is true.

2.3 First-Order Systems of General Second-Order Elliptic Equation

Following the notations of [4], we call the solution u as potential. Let the flux $\boldsymbol{\sigma} = -A\nabla u$. We have the following two-field potential-flux first-order system:

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma} + Xu = f \ \text{ in } \Omega \\ A\nabla u + \boldsymbol{\sigma} \ \ = 0 \ \text{ in } \Omega, \end{cases} \tag{2.21}$$

with boundary conditions $u = 0$ on Γ_D and $\mathbf{n} \cdot \boldsymbol{\sigma} = 0$ on Γ_N. We have $\boldsymbol{\sigma} \in H_N(\text{div } \Omega)$ and $u \in H_D^1(\Omega)$.

Introduce the intensity $\boldsymbol{\phi} = -\nabla u$. Then we have the following three-field potential-flux-intensity $(u, \boldsymbol{\sigma}, \boldsymbol{\phi})$ div-curl first-order (redundant) system as seen in Chapter 5 of [4]:

$$\begin{cases} \nabla \cdot \boldsymbol{\sigma} + Xu = f & \text{in } \Omega \\ \nabla \times \boldsymbol{\phi} = 0 & \text{in } \Omega, \\ A\nabla u + \boldsymbol{\sigma} = 0 & \text{in } \Omega, \\ A\boldsymbol{\phi} - \boldsymbol{\sigma} = 0 & \text{in } \Omega, \\ \nabla u + \boldsymbol{\phi} = 0 & \text{in } \Omega, \end{cases} \tag{2.22}$$

with boundary conditions $u = 0$ on Γ_D, $\gamma_t(\boldsymbol{\phi}) = 0$ on Γ_D, and $\mathbf{n} \cdot \boldsymbol{\sigma} = 0$ on Γ_N. We have the flux $\boldsymbol{\sigma} \in H_N(\mathrm{div}; \Omega)$, the intensity $\boldsymbol{\phi} \in H_D(\mathrm{curl}; \Omega)$, and the potential $u \in H_D^1(\Omega)$.

2.4 Some Results for Crouzeix–Raviart Elements

Let $S_{2,D} \subset H_D^1(\Omega)$ be the conforming P_2 Lagrange finite element space associated with the mesh $\mathcal{T}$. Define the following enriching operator $E_h : V_D^{cr} \to S_{2,D}$ by averaging:

$$(E_h v)(z) = \frac{1}{|\mathcal{T}_z|} \sum_{K \in \mathcal{T}_z} v|_K(z) \quad \forall z \in \mathcal{N}_{int} \cup \mathcal{N}_N, \tag{2.23}$$

where $\mathcal{T}_z$ is the set of the elements in $\mathcal{T}$ that share z as a common vertex and $\mathcal{T}_z$ is the number of the elements in $\mathcal{T}_z$. We have the following estimate (see (2.27), (2.37), and (2.28) of [12]):

$$\sum_{K \in \mathcal{T}} h_K^{-2} \|v_{cr} - E_h v_{cr}\|_{0,K}^2 \le C \sum_{F \in \mathcal{E}} \frac{1}{h_F} \|[\![v_{cr}]\!]\|_{0,F}^2 \quad \forall v_{cr} \in V_D^{cr}, \tag{2.24}$$

$$\|\nabla_h (v_{cr} - E_h v_{cr})\|_0^2 \le C \sum_{F \in \mathcal{E}} \frac{1}{h_F} \|[\![v_{cr}]\!]\|_{0,F}^2 \quad \forall v_{cr} \in V_D^{cr}, \tag{2.25}$$

$$\|\nabla (E_h v_{cr})\|_{0,\Omega} \le C \|\nabla_h v_{cr}\|_{0,\Omega} \quad \forall v_{cr} \in V_D^{cr}. \tag{2.26}$$

Although the above estimates are proved in [12] for pure Dirichlet boundary condition only, it is not hard to see they are also true for mixed boundary conditions with a non-empty Γ_D.

Following the argument in (2.41) of [12], for $v_{cr} \in V_D^{cr}$ and $v \in H_D^1(\Omega)$, we also have

$$\frac{1}{h_F} \|[\![v_{cr}]\!]\|_{0,F}^2 = \frac{1}{h_F} \|[\![v_{cr} - v]\!] - \Pi_F^0 [\![v_{cr} - v]\!]\|_{0,F}^2 \le C \|\nabla_h (v - v_{cr})\|_{0,K_F^+ \cup K_F^-}^2. \tag{2.27}$$

In other words,

$$\sum_{F \in \mathcal{E}} \frac{1}{h_F} \|[\![v_{cr}]\!]\|_{0,F}^2 \le C \inf_{v \in H_D^1(\Omega)} \|\nabla_h (v - v_{cr})\|_0^2 \quad \forall v_{cr} \in V_D^{cr}. \tag{2.28}$$

Specifically, let $v = 0$ in (2.28),

$$\sum_{F \in \mathcal{E}} \frac{1}{h_F} \|[\![v_{cr}]\!]\|_{0,F}^2 \le C \|\nabla_h v_{cr}\|_0^2 \quad \forall v_{cr} \in V_D^{cr}. \tag{2.29}$$

Choosing $v = u$, the solution of (2.10), in (2.28), we have

$$\sum_{F \in \mathcal{E}} \frac{1}{h_F} \|[\![v_{cr}]\!]\|_{0,F}^2 \le C \|\nabla_h (u - v_{cr})\|_0^2 \quad \forall v_{cr} \in V_D^{cr}. \tag{2.30}$$

We also have the following equivalence between the function jump and tangential jump of the discrete gradient: For $F \in \mathcal{E}$ and $v_{cr} \in V_D^{cr}$, we have

$$Ch_F \|[\![\gamma_{t_F}(\nabla v_{cr})]\!]\|_{0,F}^2 \leq \frac{1}{h_F}\|[\![v_{cr}]\!]\|_{0,F}^2 \leq Ch_F \|[\![\gamma_{t_F}(\nabla v_{cr})]\!]\|_{0,F}^2. \tag{2.31}$$

The equivalence in two dimensions is proved by a direct calculation in [38]. The upper bound in three dimension in proved in Lemma 3.3 of [39]. To show that $Ch_F \|[\![\nabla v_{cr} \times \mathbf{n}_F]\!]\|_{0,F}^2 \leq \frac{1}{h_F}\|[\![v_{cr}]\!]\|_{0,F}^2$ in a discrete setting, we let $\|[\![v_{cr}]\!]\|_{0,F} = 0$, then v_{cr} is continuous on F. Thus v_{cr} is of H^1 in the two elements assocaited with F and ∇v_{cr} is in $H(\mathrm{curl})$. We get $\|[\![\nabla v_{cr} \times \mathbf{n}_F]\!]\|_{0,F} = 0$. By the discrete norm equivalence, we get $Ch_F \|[\![\gamma_{t_F}(\nabla v_{cr})]\!]\|_{0,F}^2 \leq \frac{1}{h_F}\|[\![v_{cr}]\!]\|_{0,F}^2$. The weight can be obtained by the standard reference argument.

For $v_{cr} \in V_D^{cr}$ and $\boldsymbol{\tau}_{rt} \in RT_{0,N}$, with the property of the CR elements, we have

$$\begin{aligned}(\nabla_h v_{cr}, \boldsymbol{\tau}_{rt}) &= -(v_{cr}, \nabla \cdot \boldsymbol{\tau}_{rt}) + \sum_{K\in\mathcal{T}}\sum_{F\in\partial K}(\boldsymbol{\tau}_{rt}\cdot\mathbf{n}_F, v_{cr})_F \\ &= -(v_{cr}, \nabla \cdot \boldsymbol{\tau}_{rt}) + \sum_{F\in\mathcal{E}_{int}\cup\mathcal{E}_D}(\boldsymbol{\tau}_{rt}\cdot\mathbf{n}_F, [\![v_{cr}]\!])_F = -(v_{cr}, \nabla\cdot\boldsymbol{\tau}_{rt}).\end{aligned}$$

Thus we have the following discrete version of integration by parts,

$$(\nabla_h v_{cr}, \boldsymbol{\tau}_{rt}) + (v_{cr}, \nabla\cdot\boldsymbol{\tau}_{rt}) = 0 \quad \forall\, v_{cr} \in V_D^{cr} \text{ and } \boldsymbol{\tau}_{rt} \in RT_{0,N} \tag{2.32}$$

2.5 Helmholtz Decompositions

For simplicity, we assume both Γ_D and Γ_N are not empty.

The two-dimensional version of the Helmholtz decomposition can be found in [37, 49].

Lemma 2.4 *For a vector-valued function $\boldsymbol{\tau} \in L^2(\Omega)^2$, there exists $\alpha \in H_D^1(\Omega)$ and $\beta \in H_N^1(\Omega)$ such that*

$$\boldsymbol{\tau} = A\nabla\alpha + \nabla^\perp\beta.$$

We have the following Helmholtz decomposition in three dimensions (Theorem 2.6 [2] and Theorem 2.1 of [35]).

Lemma 2.5 *Assume that Ω is a simply connected, bounded, and open domain in $\mathbb{R}^3$. The boundary $\partial\Omega$ is Lipschitz continuous. Then, for a vector-valued function $\boldsymbol{\tau} \in L^2(\Omega)^3$, there exists a unique $\alpha \in H_D^1(\Omega)$ and a unique $\beta \in H^1(\Omega)^3$ with $\nabla\cdot\boldsymbol{\beta} = 0$, $\boldsymbol{\beta}\cdot\mathbf{n} = 0$ on Γ_D, $(\nabla\times\boldsymbol{\beta})\cdot\mathbf{n} = 0$ on Γ_N and $\boldsymbol{\beta}\times\mathbf{n} = 0$ on Γ_N, such that*

$$\boldsymbol{\tau} = A\nabla\alpha + \nabla\times\boldsymbol{\beta} \quad \text{and} \quad \|\boldsymbol{\beta}\|_1 \leq C\|\nabla\times\boldsymbol{\beta}\|_0.$$

Remark 2.6 A discussion of the more complicated multiply-connected domain can also be found in [2]; we omit it in this paper for simplicity.

Let

$$\mathbb{Q} = \begin{cases} H_N^1(\Omega) & d = 2, \\ \{\boldsymbol{\beta} \in H^1(\Omega)^3 : \nabla\cdot\boldsymbol{\beta} = 0, \boldsymbol{\beta}\cdot\mathbf{n} = 0 \text{ on } \Gamma_D, (\nabla\times\boldsymbol{\beta})\cdot\mathbf{n} = 0 \text{ and } \boldsymbol{\beta}\times\mathbf{n} = 0 \text{ on } \Gamma_N\} & d = 3. \end{cases} \tag{2.33}$$

We have the following Poincaré-Friedrichs inequality,

$$\|\mathtt{q}\|_0 \le C\,\|\mathtt{Curl}\,\mathtt{q}\|_0 \quad \forall \mathtt{q} \in \mathbb{Q}. \tag{2.34}$$

Thus $\|\mathtt{Curl}\,\mathtt{q}\|_0$ is a norm on $\mathbb{Q}$.

The following integrations by parts formulas also hold:

$$(\nabla p, \mathtt{Curl}\,\mathtt{q}) = 0 \quad \forall (p, \mathtt{q}) \in H_D^1(\Omega) \times \mathbb{Q}, \tag{2.35}$$

$$(\boldsymbol{\tau}, \mathtt{Curl}\,\mathtt{q}) - (\nabla \times \boldsymbol{\tau}, \mathtt{q}) = 0 \quad \forall (\boldsymbol{\tau}, \mathtt{q}) \in H_D(\mathrm{curl};\Omega) \times \mathbb{Q}. \tag{2.36}$$

The identities in (2.35) and the two-dimensional case in (2.36) are easy to prove. We only give proof of the second identity in three dimensions. First, we have the following integration by parts formula for $H(\mathrm{curl})$ functions (see Theorem 3.31 of [50]),

$$(\nabla \times \boldsymbol{\tau}, \mathtt{q}) - (\boldsymbol{\tau}, \nabla \times \mathtt{q}) = (\gamma_t(\boldsymbol{\tau}), \gamma_T(\mathtt{q}))_{\partial\Omega}, \quad \forall (\boldsymbol{\tau}, \mathtt{q}) \in H(\mathrm{curl};\Omega)^2,$$

where $\gamma_T(\mathtt{q}) = (\mathtt{q} \times \mathbf{n}) \times \mathbf{n}$. Thus, for $(\boldsymbol{\tau}, \mathtt{q}) \in H_D(\mathrm{curl};\Omega) \times \mathbb{Q}$, we have

$$\begin{aligned}(\nabla \times \boldsymbol{\tau}, \mathtt{q}) - (\boldsymbol{\tau}, \nabla \times \mathtt{q}) &= (\gamma_t(\boldsymbol{\tau}), \gamma_T(\mathtt{q}))_{\partial\Omega} = (\boldsymbol{\tau} \times \mathbf{n}, (\mathtt{q} \times \mathbf{n}) \times \mathbf{n}))_{\partial\Omega} \\ &= (\boldsymbol{\tau} \times \mathbf{n}, (\mathtt{q} \times \mathbf{n}) \times \mathbf{n})_{\Gamma_D} + (\boldsymbol{\tau} \times \mathbf{n}, (\mathtt{q} \times \mathbf{n}) \times \mathbf{n}))_{\Gamma_N} = 0.\end{aligned}$$

In summary, we have the following Helmholtz decomposition theorem

Theorem 2.7 *Assuming the conditions of Lemmas 2.4 and 2.5 hold, we have: For* $\boldsymbol{\tau} \in L^2(\Omega)^d$, *there exists* $\alpha \in H_D^1(\Omega)$ *and* $\beta \in \mathbb{Q}$ *such that*

$$\boldsymbol{\tau} = A\nabla\alpha + \mathtt{Curl}\,\mathtt{fi} \text{ and } \|A^{-1/2}\boldsymbol{\tau}\|_0^2 = \|A^{1/2}\nabla\alpha\|_0^2 + \|A^{-1/2}\mathtt{Curl}\,\mathtt{fi}\|_0^2. \tag{2.37}$$

The identity $\|A^{-1/2}\boldsymbol{\tau}\|_0^2 = \|A^{1/2}\nabla\alpha\|_0^2 + \|A^{-1/2}\mathtt{Curl}\,\mathtt{fi}\|_0^2$ is a simple consequence of the orthogonality relation (2.35).

2.6 Discrete Poincaré-Friedrichs Inequality

The proof of the following discrete Poincaré-Friedrichs inequality can be found in papers [51, 52].

Theorem 2.8 *There exists a positive constant* C *depending on* Ω *such that*

$$\|v\|_0 \le C\|\nabla_h v\|_0 \quad v \in W_D^1(\mathcal{T}). \tag{2.38}$$

With the help of the discrete Poincaré-Friedrichs inequality, we can define the following combined norm for the space $H_N(\mathrm{div};\Omega) \times W_D^1(\mathcal{T})$:

$$|||(\boldsymbol{\tau}, v)|||^2 := \|\nabla_h v\|_0^2 + \|\boldsymbol{\tau}\|_0^2 + \|\nabla\cdot\boldsymbol{\tau}\|_0^2 \quad (\boldsymbol{\tau}, v) \in H_N(\mathrm{div};\Omega) \times W_D^1(\mathcal{T}).$$

3 Two-Field Potential-Flux Div CR-LSFEM

This section introduces the two-field potential-flux div least-squares method for general second-order elliptic equations with Crouzeix-Raviart elements.

For $(\boldsymbol{\tau}, v) \in H_N(\mathrm{div};\Omega) \times W_D^{1+cr}$, define the two-field potential-flux div least-squares functional for (2.10) as:

$$\mathcal{J}_h^{div}(\boldsymbol{\tau}, v; f) := \|A^{1/2}\nabla_h v + A^{-1/2}\boldsymbol{\tau}\|_0^2 + \|\nabla\cdot\boldsymbol{\tau} + X_h v - f\|_0^2, \tag{3.1}$$

where $X_h v = \mathbf{b} \cdot \nabla_h v + cv$.

For $(\boldsymbol{\chi}, w) \in H_N(\mathrm{div}; \Omega) \times W_D^{1+cr}$ and $(\boldsymbol{\tau}, v) \in H_N(\mathrm{div}; \Omega) \times W_D^{1+cr}$, define the following bilinear form b_h:

$$b_h((\boldsymbol{\chi}, w), (\boldsymbol{\tau}, v)) := (A\nabla_h w + \boldsymbol{\chi}, \nabla_h v + A^{-1}\boldsymbol{\tau}) + (\nabla \cdot \boldsymbol{\chi} + X_h w, \nabla \cdot \boldsymbol{\tau} + X_h v). \tag{3.2}$$

When $v \in H_D^1(\Omega)$, we can remove the subscript h in ∇_h of the above definitions, and they are the standard potential-flux div least-squares functional and bilinear forms. We have,

$$\mathcal{J}_h^{div}(\boldsymbol{\tau}, v; 0) = b_h((\boldsymbol{\tau}, v), (\boldsymbol{\tau}, v)) \quad (\boldsymbol{\tau}, v) \in H_N(\mathrm{div}; \Omega) \times W_D^{1+cr}.$$

The solution $(\boldsymbol{\sigma}, u)$ satisfies the least-squares minimization problem: Find $(\boldsymbol{\sigma}, u) \in H_N(\mathrm{div}; \Omega) \times H_D^1(\Omega)$ such that

$$\mathcal{J}_h^{div}(\boldsymbol{\sigma}, u; f) = \inf_{(\boldsymbol{\tau}, v) \in H_N(\mathrm{div}; \Omega) \times H_D^1(\Omega)} \mathcal{J}_h^{div}(\boldsymbol{\tau}, v; f). \tag{3.3}$$

Or equivalently, a weak problem: Find $(\boldsymbol{\sigma}, u) \in H_N(\mathrm{div}; \Omega) \times H_D^1(\Omega)$ such that

$$b_h((\boldsymbol{\sigma}, u), (\boldsymbol{\tau}, v)) = (f, \nabla \cdot \boldsymbol{\tau} + X_h v) \quad \forall (\boldsymbol{\tau}, v) \in H_N(\mathrm{div}; \Omega) \times H_D^1(\Omega). \tag{3.4}$$

The following norm equivalence is standard. Different proofs can be found in [1, 7, 28–30].

Theorem 3.1 *Assuming one of the conditions in Theorem 2.3 is true, then for all* $(\boldsymbol{\tau}, v) \in H_N(\mathrm{div}; \Omega) \times H_D^1(\Omega)$, *we have*

$$C_1 |\!|\!| (\boldsymbol{\tau}, v) |\!|\!|^2 \le \mathcal{J}_h^{div}(\boldsymbol{\tau}, v; 0) = \|A^{1/2}\nabla v + A^{-1/2}\boldsymbol{\tau}\|_0^2 + \|\nabla \cdot \boldsymbol{\tau} + Xv\|_0^2 \le C_2 |\!|\!| (\boldsymbol{\tau}, v) |\!|\!|^2. \tag{3.5}$$

Consider the two-field potential-flux div CR-LSFEM: Find $(\boldsymbol{\sigma}_{rt}, u_{cr}) \in RT_{0,N} \times V_D^{cr}$ such that

$$\mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_{cr}; f) = \inf_{(\boldsymbol{\tau}_{rt}, v_{cr}) \in RT_{0,N} \times V_D^{cr}} \mathcal{J}_h^{div}(\boldsymbol{\tau}_{rt}, v_{cr}; f). \tag{3.6}$$

Or equivalently: Find $(\boldsymbol{\sigma}_{rt}, u_{cr}) \in RT_{0,N} \times V_D^{cr}$ such that

$$b_h((\boldsymbol{\sigma}_{rt}, u_{cr}), (\boldsymbol{\tau}_{rt}, v_{cr})) = F_h(\boldsymbol{\tau}_{rt}, v_{cr}) \quad \forall (\boldsymbol{\tau}_{rt}, v_{cr}) \in RT_{0,N} \times V_D^{cr}. \tag{3.7}$$

Unlike the standard CR finite element method [11, 12], as discussed in [18], we have the error equation for the least-squares method.

Lemma 3.2 *Assume that* $(\boldsymbol{\sigma}, u) \in H_N(\mathrm{div}; \Omega) \times H_D^1(\Omega)$ *is the solution of (3.4) and* $(\boldsymbol{\sigma}_{rt}, u_{cr}) \in RT_{0,N} \times V_D^{cr}$ *is the numerical solution of (3.7). The following error equation is true,*

$$b_h((\boldsymbol{\sigma} - \boldsymbol{\sigma}_{rt}, u - u_{cr}), (\boldsymbol{\tau}_{rt}, v_{cr})) = 0 \quad \forall (\boldsymbol{\tau}_{rt}, v_{cr}) \in RT_{0,N} \times V_D^{cr}. \tag{3.8}$$

Proof Since $\boldsymbol{\sigma}$ and u satisfy the first-order system (2.21) in the L^2 sense, then for any $(\boldsymbol{\tau}_{rt}, v_{cr}) \in RT_{0,N} \times V_D^{cr}$, we have

$$\begin{aligned} b_h((\boldsymbol{\sigma}, u), (\boldsymbol{\tau}_{rt}, v_{cr})) &= (A\nabla u + \boldsymbol{\sigma}, \nabla_h v_{cr} + A^{-1}\boldsymbol{\tau}_{rt}) + (\nabla \cdot \boldsymbol{\sigma} + Xu, \nabla \cdot \boldsymbol{\tau}_{rt} + X_h v_{cr}) \\ &= (f, \nabla \cdot \boldsymbol{\tau}_{rt} + X_h v_{cr}) = b_h((\boldsymbol{\sigma}_{rt}, u_{cr}), (\boldsymbol{\tau}_{rt}, v_{cr})). \end{aligned}$$

Thus the error equation is proved. □

4 Discrete Existence and Uniqueness of the Two-Field Div CR-LSFEM: I. Coerciveness Assumption

We first present a coerciveness result of the two-field potential-flux div CR- LSFEM with some assumptions of the coefficients ensuring the coerciveness of the original bilinear form $a(\cdot,\cdot)$ (2.13).

Assumption 4.1 Assume that Assumption 2.1 is true. It is further assumed that $\nabla\cdot\mathbf{b}\in L^\infty(\Omega)$ and $c\in L^\infty(\Omega)$. Thus we have $\mathbf{b}$ is an $H(\mathrm{div})$ function and $\mathbf{b}\cdot\mathbf{n}$ is continuous for $F\in\mathcal{E}_{int}$. We also assume that $c-\frac{1}{2}\nabla\cdot\mathbf{b}\geq 0$ in Ω, $\mathbf{b}\cdot\mathbf{n}\geq 0$ on Γ_N, and $\mathbf{b}\cdot\mathbf{n}|_F$ has a fixed sign on each $F\in\mathcal{E}_{int}$.

Theorem 4.2 *Assuming that Assumption (4.1) is true and the mesh size of $\mathcal{T}$ is small enough, there exists constants $C_1>0$ and $C_2>0$ independent of mesh size, such that*

$$C_1|\!|\!|(\boldsymbol{\tau}_{rt},v_{cr})|\!|\!|^2\leq\mathcal{J}_h^{div}(\boldsymbol{\tau}_{rt},v_{cr};0)\leq C_2|\!|\!|(\boldsymbol{\tau}_{rt},v_{cr})|\!|\!|^2\quad\forall(\boldsymbol{\tau}_{rt},v_{cr})\in RT_{0,N}\times V_D^{cr},\tag{4.1}$$

Proof The upper bound is easy to prove by the triangle and the discrete Poincaré inequalities. Our main task is to prove the coerciveness.

Using integration by parts, we can easily derive that $(\mathbf{b}\cdot\nabla_h v_{cr},v_{cr})=-\frac{1}{2}(\nabla\cdot\mathbf{b},v_{cr}^2)+\frac{1}{2}\sum_{K\in\mathcal{T}}(\mathbf{n}\cdot\mathbf{b},v_{cr}^2)_{\partial K}$. Then due to fact that $\mathbf{n}\cdot\mathbf{b}\geq 0$ on Γ_N and the uniform bound of $c-\frac{1}{2}\nabla\cdot\mathbf{b}\geq 0$ in Ω,

$$\begin{aligned}-(X_h v_{cr},v_{cr})&=-(c-\frac{1}{2}\nabla\cdot\mathbf{b},v_{cr}^2)-\frac{1}{2}\sum_{K\in\mathcal{T}}(\mathbf{n}\cdot\mathbf{b},v_{cr}^2)_{\partial K}\leq-\frac{1}{2}\sum_{K\in\mathcal{T}}(\mathbf{n}\cdot\mathbf{b},v_{cr}^2)_{\partial K}\\&\leq-\frac{1}{2}\sum_{F\in\mathcal{E}_{int}}(\mathbf{n}\cdot\mathbf{b},[\![v_{cr}^2]\!])_F-\frac{1}{2}\sum_{F\in\mathcal{E}_D}(\mathbf{n}\cdot\mathbf{b},v_{cr}^2)_F.\end{aligned}$$

For any $v_{cr}\in V^{cr}$ and any $F\in\mathcal{E}_{int}$, by the fact that $\int_F[\![v_{cr}]\!]ds=0$, the mean value of v_{cr} over F is single-valued constant, $\overline{v}_{cr,F}=\Pi_F^0 v_{cr}|_{K_F^+}=\Pi_F^0 v_{cr}|_{K_F^-}$, where K_F^+ and K_F^- are two elements sharing the common F. We also have $\overline{v}_{cr,F}=0$ for $F\in\mathcal{E}_D$. Then for an $F\in\mathcal{E}_{int}$, we have

$$\begin{aligned}\frac{1}{2}(\mathbf{b}\cdot\mathbf{n},[\![v_{cr}^2]\!])_F&=(\mathbf{b}\cdot\mathbf{n}[\![v_{cr}]\!],\{v_{cr}\})_F\leq|\mathbf{b}\cdot\mathbf{n}|_{\infty,F}|([\![v_{cr}]\!],\{v_{cr}\})_F|\\&=|\mathbf{b}\cdot\mathbf{n}|_{\infty,F}|([\![v_{cr}]\!],\{v_{cr}-\overline{v}_{cr,F}\})_F|\\&\leq|\mathbf{b}\cdot\mathbf{n}|_{\infty,F}\|[\![v_{cr}]\!]\|_F(\|v_{cr}|_{K_F^-}-\overline{v}_{cr,F}\|_{0,F}+\|v_{cr}|_{K_F^+}-\overline{v}_{cr,F}\|_{0,F})/2.\end{aligned}$$

We have (see e.g. p.110 of [53])

$$\|v_{cr}|_{K_F}-\overline{v}_{cr,F}\|_{0,F}\leq Ch_F^{1/2}\|\nabla v_{cr}\|_{0,K_F}.$$

Combing the above estimate and (2.27), we get

$$(\mathbf{b}\cdot\mathbf{n},[\![v_{cr}^2]\!])_F\leq C_F h_F|\mathbf{b}\cdot\mathbf{n}|_{\infty,F}\|\nabla v_{cr}\|^2_{0,K_F^-\cup K_F^+}\quad\text{for } F\in\mathcal{E}_{int},$$

where C_F only depends on the shape of $K\in K_F^-\cup K_F^+$. Similarly, we have

$$(\mathbf{b}\cdot\mathbf{n},v_{cr}^2)_F\leq C_F h_F|\mathbf{b}\cdot\mathbf{n}|_{\infty,F}\|\nabla v_{cr}\|^2_{0,K_F}\quad\text{for } F\in\mathcal{E}_D.$$

Thus, we have

$$-(X_h v_{cr}, v_{cr}) \le C \sum_{K\in\mathcal{T}} \Big(\sup_{F\in\partial K\cap(\mathcal{E}_{int}\cup\mathcal{E}_D)} |\mathbf{b}\cdot\mathbf{n}|_{\infty,F} h_F \Big) \|\nabla v_{cr}\|_{0,K}^2. \tag{4.2}$$

By (2.32), the discrete Poincaré inequality, and (4.2),

$$\begin{aligned}
\|A^{1/2}\nabla_h v_{cr}\|_0^2 &= (A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}, A^{1/2}\nabla_h v_{cr}) + (\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}, v_{cr}) - (X_h v_{cr}, v_{cr}) \\
&\le C\|A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}\|_0 \|\nabla_h v_{cr}\|_0 \\
&\quad + C\|\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0 \|\nabla_h v_{cr}\|_0 \\
&\quad + C_0 \sum_{K\in\mathcal{T}} \Big(\sup_{F\in\partial K\cap(\mathcal{E}_{int}\cup\mathcal{E}_D)} |\mathbf{b}\cdot\mathbf{n}|_{\infty,F} h_F \Big) \|\nabla v_{cr}\|_{0,K}^2.
\end{aligned}$$

where C_0 only depends on the shape regularity of the mesh $\mathcal{T}$. Choosing the mesh size small enough such that on each element K,

$$C_0 \sup_{F\in\partial K\cap(\mathcal{E}_{int}\cup\mathcal{E}_D)} |\mathbf{n}_F\cdot\mathbf{b}|_{\infty,F} h_F < \frac{1}{2}\Lambda_0|_K, \tag{4.3}$$

we have the following bound for a $C > 0$ independent of the mesh size,

$$C\|\nabla_h v_{cr}\|_0^2 \le \mathcal{J}_h^{div}(\boldsymbol{\tau}_{rt}, v_{cr}; 0). \tag{4.4}$$

An application of the triangle inequality shows that

$$C\|\boldsymbol{\tau}_{rt}\|_0 \le \|A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}\|_0 + C\|\nabla_h v_{cr}\|_0, \tag{4.5}$$

$$\text{and } C\|\nabla\cdot\boldsymbol{\tau}_{rt}\|_0 \le \|\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0 + \|X_h v_{cr}\|_0 \le \|\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0 + C\|\nabla_h v_{cr}\|_0. \tag{4.6}$$

The lemma is then proved. □

Remark 4.3 We know from (4.3) that the mesh size needs not to be extremely small. It depends on the shape of the element, the coefficient A on the element, and $\mathbf{b}\cdot\mathbf{n}_F$ on its faces/edges. Also, (4.3) is a local result suitable for an adaptively refined non-uniform mesh.

No regularity assumption is needed in this proof.

Assumption 4.1 is not the most general one to ensure the coerciveness of the bilinear form a. One can find in [22] for a more general setting, and the analysis presented in this section can also be extended to the more general setting. The proof in this section shows that an explicit assumption on the coefficient helps determine the local mesh size to ensure the coerciveness of the discrete problem.

We also notice that even with Assumption 4.1, the original bilinear form for the pure CR finite element is only coercive under the assumption that the mesh is fine enough.

With the coerciveness result (4.1), the error equation (3.8), local approximation properties (2.5), (2.6), and (2.9), we immediately have the following a priori error estimate.

Theorem 4.4 *For piecewise constant function s on mesh $\mathcal{T}$ with $s_K = s|_K$ satisfying $0 < s_K \le 1$, assume that $u|_K \in H^{1+s_K}(K)$, $\boldsymbol{\sigma}|_K \in H^{s_K}(K)$, and $\nabla\cdot\boldsymbol{\sigma}|_K \in H^{s_K}(K)$, for $K \in \mathcal{T}$. We also assume that Assumption (4.1) is true and the mesh size of $\mathcal{T}$ is small enough, then the following a priori error estimate is true:*

$$\begin{aligned}
|||(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{rt}, u - u_{cr})||| &\le C \inf_{(\boldsymbol{\tau}_{rt}, v_{rt})\in RT_{0,N}\times V_D^{cr}} |||(\boldsymbol{\sigma} - \boldsymbol{\tau}_{rt}, u - v_{cr})||| \\
&\le C \sum_{K\in\mathcal{T}} h_K^{s_K} (|u|_{1+s_K,K} + |\boldsymbol{\sigma}|_{s_K,K} + |\nabla\cdot\boldsymbol{\sigma}|_{s_K,K}).
\end{aligned}$$

5 Discrete Existence and Uniqueness of the Two-Field Div CR-LSFEM: II. Schatz's Argument

In this section, we present a discrete existence and uniqueness proof based on the argument of Schatz [23] without assuming Assumption 4.1 to ensure the coerciveness of the bilinear form a. Thus, the original variational problem (2.12) can be indefinite.

In this proof, we only assume one of the conditions of Theorem 2.3 and a minimal regularity of the original and adjoint elliptic equations. First, we prove an L^2-error estimate of the two-field potential-flux div method, and then a Garding-like inequality is proved. With these tools, we can get an a priori estimate, then the existence and uniqueness of the discrete problem follow.

5.1 An L^2 Estimate of the Potential-Flux Div CR-LSFEM

In this subsection, we present an L^2-error estimate of the two-field potential-flux div CR-LSFEM based on the argument of [33]. We also correct a minor mistake in the original proof of [33].

We assume the following very mild regularity assumption.

Assumption 5.1 Assume one of the conditions of Theorem 2.3 is true, and the following H^{1+s} regularity estimates are true for (2.12) and (2.16), respectively:

$$\|u\|_{1+s} \le C\|f\|_0 \quad \text{and} \quad \|z\|_{1+s} \le C\|g\|_0, \quad \text{for some } 0 < s \le 1. \tag{5.1}$$

To discuss an L^2-error estimate, we introduce a first-order system:

$$\begin{cases} A\nabla w + \boldsymbol{\gamma} &= A\nabla z \text{ in } \Omega \\ \nabla\cdot\boldsymbol{\gamma} + Xw &= z \quad \text{in } \Omega, \end{cases} \tag{5.2}$$

with boundary conditions $w = 0$ on Γ_D and $\mathbf{n}\cdot\boldsymbol{\gamma} = 0$ on Γ_N.

Lemma 5.2 *Let z be the solution of the adjoint problem (2.16) and $(\boldsymbol{\gamma}, w)$ be the functions defined in (5.2), we have the following estimates under Assumption 5.1:*

$$\|w\|_{1+s} + \|\boldsymbol{\gamma}\|_s + \|\nabla\cdot\boldsymbol{\gamma}\|_s \le C\|g\|_0. \tag{5.3}$$

Proof The first-order system (5.2) can be understood in the PDE form as:

$$-\nabla\cdot(A\nabla w) + \mathbf{b}\cdot\nabla w + cw = z - \nabla\cdot(A\nabla z) \text{ in } \Omega,$$

with boundary conditions

$$w = 0 \text{ on } \Gamma_D \quad \text{and} \quad A\nabla w\cdot\mathbf{n} = A\nabla z\cdot\mathbf{n} = (\mathbf{b}\cdot\mathbf{n})z \text{ on } \Gamma_N.$$

Note that the boundary condition of w on Γ_N is not standard. Thus the claim in p.1729 (below (5.7)) of the paper [33] is not accurate. The boundary condition shows that w is not a solution of $w \in H^1_D(\Omega)$ with $a(w, v) = (z - \nabla\cdot(A\nabla z), v)$, for all $v \in H^1_D(\Omega)$. We can not use the regularity result of (2.12) directly. To fix this, introduce $y = w - z \in H^1_D(\Omega)$. Then

$$\begin{cases} A\nabla y + \boldsymbol{\gamma} &= 0 \quad \text{in } \Omega \\ \nabla\cdot\boldsymbol{\gamma} + Xy &= z - Xz \text{ in } \Omega. \end{cases}$$

Then we have $y \in H_D^1(\Omega)$ satisfying

$$-\nabla \cdot (A\nabla y) + Xy = z - Xz \quad \text{in } \Omega, \quad y = 0 \text{ on } \Gamma_D, \quad A\nabla y \cdot \mathbf{n} = 0 \text{ on } \Gamma_N.$$

Or, equivalently,

$$a(y, v) = (z - Xz, v) \quad \forall v \in H_D^1(\Omega).$$

Then by the regularity assumption (5.1) (first for y, then for z), we have

$$\|y\|_{1+s} \le C\|z - Xz\|_0 \le C\|z\|_1 \le C\|g\|_0.$$

Then, the following estimate for w is true:

$$\|w\|_{1+s} = \|y + z\|_{1+s} \le \|y\|_{1+s} + \|z\|_{1+s} \le C\|g\|_0.$$

We also have

$$\begin{aligned} \|\boldsymbol{\gamma}\|_s &= \|A\nabla y\|_s \le C\|y\|_{1+s} \le C\|g\|_0, \\ \|\nabla \cdot \boldsymbol{\gamma}\|_s &= \|z - Xw\|_s \le C(\|z\|_1 + \|w\|_{1+s}) \le C\|g\|_0. \end{aligned}$$

This completes the proof of the lemma. □

Lemma 5.3 *Assume h is the maximum mesh size of the mesh $\mathcal{T}$. Let z be the solution of (2.16), u be the solution of (2.10), and u_{cr} be the solution of (3.7), then the following inequality holds under Assumption 5.1:*

$$\sum_{F \in \mathcal{E}_{int} \cup \mathcal{E}_D} ((A\nabla z + \mathbf{b}z) \cdot \mathbf{n}, [\![u_{cr}]\!])_F \le Ch^s \|\nabla_h(u - u_{cr})\|_0 \|g\|_0. \tag{5.4}$$

Proof Let $\boldsymbol{\xi} = A\nabla z + \mathbf{b}z$, then $\nabla \cdot \boldsymbol{\xi} = cz - g \in L^2(\Omega)$, thus $\boldsymbol{\xi} \in H_N(\text{div}; \Omega)$. By the regularity assumption on z, we have $\boldsymbol{\xi} \in H^s(\Omega)$. Define $\boldsymbol{\xi}_{rt0}$ to be $\boldsymbol{\xi}$'s interpolation in $RT_{0,N}$. We have

$$\|\boldsymbol{\xi} - \boldsymbol{\xi}_{rt0}\|_0 \le Ch^s \|\boldsymbol{\xi}\|_s \quad \text{and} \quad \|\nabla \cdot \boldsymbol{\xi} - \nabla \cdot \boldsymbol{\xi}_{rt0}\|_0 \le \|\nabla \cdot \boldsymbol{\xi}\|_0 = \|g - cz\|_0 \le C\|g\|_0.$$

Summing up all elements and using the regularity assumption and the fact $0 < s \le 1$, we get

$$\sum_{K \in \mathcal{T}} (\|\boldsymbol{\xi} - \boldsymbol{\xi}_{rt0}\|_{0,K} + h_K \|\nabla \cdot (\boldsymbol{\xi} - \boldsymbol{\xi}_{rt0})\|_{0,K}) \le h^s \|g\|_0. \tag{5.5}$$

By the property of $([\![u_{cr}]\!], 1)_F = 0$ and the fact $\boldsymbol{\xi}_{rt0} \cdot \mathbf{n}|_F$ is a constant, we have

$$(\boldsymbol{\xi}_{rt0} \cdot \mathbf{n}, [\![u_{cr}]\!])_F = 0 \quad \forall F \in \mathcal{E}_{int} \cup \mathcal{E}_D. \tag{5.6}$$

By the above result, the trace theorem of Lemma 2.4 and Remark 2.5 of [39], (2.30), and (5.5), we have

$$\begin{aligned} &\sum_{F \in \mathcal{E}_{int} \cup \mathcal{E}_D} (\boldsymbol{\xi} \cdot \mathbf{n}, [\![u_{cr}]\!])_F = \sum_{F \in \mathcal{E}_{int} \cup \mathcal{E}_D} ((\boldsymbol{\xi} - \boldsymbol{\xi}_{rt0}) \cdot \mathbf{n}, [\![u_{cr}]\!])_F \\ &\le C \sum_{F \in \mathcal{E}_{int} \cup \mathcal{E}_D} h_F^{-1/2} \|[\![u_{cr}]\!]\|_{0,F} (\|\boldsymbol{\xi} - \boldsymbol{\xi}_{rt0}\|_{0,K_F^- \cup K_F^+} + h_K \|\nabla \cdot (\boldsymbol{\xi} - \boldsymbol{\xi}_{rt0})\|_{0,K_F^- \cup K_F^+}) \\ &\le Ch^s \|\nabla_h(u - u_{cr})\|_0 \|g\|_0. \end{aligned}$$

□

We also have the following approximation property by (2.9), (2.5), (2.6), and (5.3),

$$\inf_{(\boldsymbol{\gamma}_h, w_h)\in RT_{0,N}\times V_D^{cr}} |\!|\!|(\boldsymbol{\gamma}-\boldsymbol{\gamma}_h, w-w_h)|\!|\!| \le Ch^s(\|\boldsymbol{\gamma}\|_s + \|\nabla\cdot\boldsymbol{\sigma}\|_s + \|w\|_{1+s}) \le Ch^s\|g\|_0. \tag{5.7}$$

Theorem 5.4 *Assume h is the maximum mesh size of the mesh $\mathcal{T}$. Let $(\boldsymbol{\sigma}, u)$ be the solution of (2.21) and $(\boldsymbol{\sigma}_{rt}, u_{cr})$ be the solution of (3.7), then the following L^2-error estimate holds under Assumption 5.1:*

$$\|u-u_{cr}\|_0 \le Ch^s |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_{cr})|\!|\!|. \tag{5.8}$$

Proof Let $e = u - u_{cr} \in W_D^{1+cr}$ and $\mathbf{E} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_{rt} \in H_N(\mathrm{div};\Omega)$. Multiplying both sides of (2.15) by e and integrating by parts, we have

$$\begin{aligned}(e,g) &= -(\nabla\cdot(A\nabla z + \mathbf{b}z), e) + (ce, z)\\ &= (A\nabla z + \mathbf{b}z, \nabla_h e) + (cz, e) - \sum_{K\in\mathcal{T}} ((A\nabla z + \mathbf{b}z)\cdot\mathbf{n}_{\partial K}, e)_{\partial K}\\ &= (A\nabla_h e, \nabla z) + (X_h e, z) + \sum_{F\in\mathcal{E}_{int}\cup\mathcal{E}_D} ((A\nabla z + \mathbf{b}z)\cdot\mathbf{n}, [\![u_{cr}]\!])_F.\end{aligned}$$

Using the fact that $(\mathbf{E}, \nabla z) + (\nabla\cdot\mathbf{E}, z) = 0$, the system (5.2), the error Eq. (3.8), and the approximation property (5.7), the first two terms can be bounded by the following estimate,

$$\begin{aligned}(A\nabla_h e, \nabla z) + (X_h e, z) &= (A\nabla_h e + \mathbf{E}, \nabla z) + (\nabla\cdot\mathbf{E} + X_h e, z)\\ &= (A\nabla_h e + \mathbf{E}, \nabla w + A^{-1}\boldsymbol{\gamma}) + (\nabla\cdot\mathbf{E} + X_h e, \nabla\cdot\boldsymbol{\gamma} + Xw)\\ &= b_h(\mathbf{E}, e; \boldsymbol{\gamma}, w) = b_h(\mathbf{E}, e; \boldsymbol{\gamma}-\boldsymbol{\gamma}_h, w-w_h)\\ &\le |\!|\!|(\mathbf{E}, e)|\!|\!| |\!|\!|(\boldsymbol{\gamma}-\boldsymbol{\gamma}_h, w-w_h)|\!|\!| \le ch^s\|g\|_0 |\!|\!|(\mathbf{E}, e)|\!|\!|\end{aligned}$$

Combining with Lemma 5.3, we have the result of the theorem. □

5.2 Discrete Existence and Uniqueness Based on Schatz's Argument

We first prove a discrete Garding-like inequality, then derive a priori error estimate, and show the existence and uniqueness of the discrete potential-flux div CR method by using Schatz's argument [23].

Lemma 5.5 *The following inequality is true for $(\boldsymbol{\tau}_{rt}, v_{cr}) \in RT_{0,N}\times V_D^{cr}$:*

$$C|\!|\!|(\boldsymbol{\tau}_{rt}, v_{cr})|\!|\!|^2 \le \mathcal{J}_h^{div}(\boldsymbol{\tau}_{rt}, v_{cr}; 0) + \|v_{cr}\|_0^2 = b_h((\boldsymbol{\tau}_{rt}, v_{cr}), (\boldsymbol{\tau}_{rt}, v_{cr})) + \|v_{cr}\|_0^2. \tag{5.9}$$

Proof By (2.32) and the discrete Poincaré inequality,

$$\begin{aligned}\|A^{1/2}\nabla_h v_{cr}\|_0^2 &= (A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}, A^{1/2}\nabla_h v_{cr})\\ &\quad +(\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}, v_{cr}) - (X_h v_{cr}, v_{cr})\\ &\le C(\|A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}\|_0\\ &\quad +\|\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0)\|\nabla_h v_{cr}\|_0 + C\|\nabla_h v_{cr}\|_0\|v_{cr}\|_0\\ &\le C(\|A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}\|_0\\ &\quad +\|\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0 + \|v_{cr}\|_0)\|A^{1/2}\nabla_h v_{cr}\|_0.\end{aligned}$$

Thus,

$$\|\nabla_h v_{cr}\|_0 \le C(\|A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}\|_0 + \|\nabla\cdot\boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0 + \|v_{cr}\|_0).$$

The lemma is proved by combining the above result, (4.5), and (4.6). □

Theorem 5.6 *Assume Assumption 5.1 is true. There exists an $h_0 > 0$, such that when the maximum mesh size h of $\mathcal{T}$ is smaller than h_0, the discrete problem of the two-field potential-flux div CR-LSFEM (3.3) or (3.4) has a unique solution. For piecewise constant function s on mesh $\mathcal{T}$ with $s_K = s|_K$ satisfying $0 < s_K \le 1$, assume that $u|_K \in H^{1+s_K}(K)$, $\boldsymbol{\sigma}|_K \in H^{s_K}(K)$, and $\nabla\cdot\boldsymbol{\sigma}|_K \in H^{s_K}(K)$, for $K \in \mathcal{T}$. The following a priori error estimate is true,*

$$\begin{aligned}|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_{cr})|\!|\!| &\le C \inf_{(\boldsymbol{\tau}_{rt}, v_{rt})\in RT_{0,N}\times V_D^{cr}} |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{rt}, u-v_{cr})|\!|\!| \\ &\le C\sum_{K\in\mathcal{T}} h_K^{s_K}(|u|_{1+s_K,K} + |\boldsymbol{\sigma}|_{s_K,K} + |\nabla\cdot\boldsymbol{\sigma}|_{s_K,K}).\end{aligned}$$

Proof Let v_{cr} be an arbitrary function in V_D^{cr} and $\boldsymbol{\tau}_{cr}$ be an arbitrary function in $RT_{0,N}$. Denote $e_h = u_{cr} - v_{cr}$ and $\mathbf{E}_h = \boldsymbol{\sigma}_{cr} - \boldsymbol{\tau}_{cr}$. By the result of Lemma 5.5, the error Eq. (3.8), we have

$$\begin{aligned}|\!|\!|(\mathbf{E}_h, e_h)|\!|\!|^2 &\le C(b_h(\mathbf{E}_h, e_h; \mathbf{E}_h, e_h) + \|e_h\|_0^2) \\ &\le C(b_h(\boldsymbol{\sigma}-\boldsymbol{\tau}_{rt}, u-v_{cr}; \mathbf{E}_h, e_h) + \|u-v_{cr}\|_0^2 + \|e\|_0^2) \\ &\le C(|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{rt}, u-v_{cr})|\!|\!||\!|\!|(\mathbf{E}_h, e_h)|\!|\!| + |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{rt}, u-v_{cr})|\!|\!|^2 + \|e\|_0^2)\end{aligned}$$

An application of Young's inequality with ϵ shows

$$C|\!|\!|(\mathbf{E}_h, e_h)|\!|\!| \le |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{rt}, u-v_{cr})|\!|\!| + \|e\|_0. \tag{5.10}$$

Then by the triangle inequality, (5.10), and the L^2-error estimates (5.8),

$$\begin{aligned}|\!|\!|(\mathbf{E}, e)|\!|\!| &\le |\!|\!|(\mathbf{E}_h, e_h)|\!|\!| + |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{cr}, u-v_{cr})|\!|\!| \le C(|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{cr}, u-v_{cr})|\!|\!| + \|e\|_0) \\ &\le C(|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{cr}, u-v_{cr})|\!|\!| + h^s|\!|\!|(\mathbf{E}, e)|\!|\!|).\end{aligned}$$

Choosing h_0 small enough, we have

$$|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_{cr})|\!|\!| \le C \inf_{(\boldsymbol{\tau}_{rt}, v_{rt})\in RT_{0,N}\times V_D^{cr}} |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\tau}_{rt}, u-v_{cr})|\!|\!|.$$

The other part of the a priori error analysis is from the local approximation properties (2.5), (2.6), and (2.9). The existence and uniqueness of the discrete problem are then a simple consequence of the a priori error estimate. □

6 Discrete Existence and Uniqueness of the Two-Field Div CR-LSFEM: III. Inf-Sup Condition of a Standard CR Method

In this section, we present a third proof of the coerciveness of the two-field div CR-LSFEM by using the stability of the standard CR method for the second-order linear elliptic equation [25]. Define the discrete bilinear form a_h corresponding to (2.13),

$$a_h(w, v) := (A\nabla_h w, \nabla_h v) + (X_h w, v) \quad \text{for } w, v \in W_D^{1+cr}. \tag{6.1}$$

Then the problem of the standard CR element approximation to (2.10) is: Find $u_{cr}^p \in V_D^{cr}$, such that

$$a_h(u_{cr}^p, v_{cr}) = (f, v_{cr}) \quad \forall v_{cr} \in V_D^{cr}. \tag{6.2}$$

And the problem of the standard CR element approximation to the adjoint Eq. (2.15) is: Find $u_{cr}^d \in V_D^{cr}$, such that

$$a_h(v_{cr}, u_{cr}^d) = (f, v_{cr}) \quad \forall v_{cr} \in V_D^{cr}. \tag{6.3}$$

In [25], for the problem (6.3) with $\Gamma_D = \partial\Omega$ and a more general righthand side, the following CR approximation of the adjoint problem is considered: For all $f_0 \in L^2(\Omega)$ and $\mathbf{f}_1 \in L^2(\Omega)^d$, find $u_{cr}^d \in V_D^{cr}$, such that

$$a_h(v_{cr}, u_{cr}^d) = (f_0, v_{cr}) + (\mathbf{f}_1, \nabla_h v_{cr}) \quad \forall v_{cr} \in V_D^{cr}. \tag{6.4}$$

It is proved that (6.4) has a unique solution if the mesh size is small enough and Assumption 5.1 is true. The following stability result is also proved:

$$\beta \|\nabla_h u_{cr}^d\|_0 \le \|f_0\|_0 + \|\mathbf{f}_1\|_0 \quad \forall f_0 \in L^2(\Omega) \quad \mathbf{f}_1 \in L^2(\Omega)^d. \tag{6.5}$$

It is easy to derive that the above result still holds for mixed boundary conditions with $\Gamma_D \neq \emptyset$. By the Banach-Necăs-Babuška theory [53–55], it is known that the well-posedness of (6.4) is equivalent to the inf-sup stability of a_h:

$$0 < \beta = \inf_{v_{cr} \in V_D^{cr}} \sup_{w_{cr} \in V_D^{cr}} \frac{a_h(v_{cr}, w_{cr})}{\|\nabla_h v_{cr}\|_0 \|\nabla_h w_{cr}\|_0} = \inf_{w_{cr} \in V_D^{cr}} \sup_{v_{cr} \in V_D^{cr}} \frac{a_h(v_{cr}, w_{cr})}{\|\nabla_h v_{cr}\|_0 \|\nabla_h w_{cr}\|_0}. \tag{6.6}$$

The inf-sup condition (6.6) is also equivalent to

$$\beta \|\nabla_h v_h\|_0 \le \sup_{w_h \in V_D^{cr}} \frac{a_h(v_h, w_h)}{\|\nabla_h w_h\|_0} \quad \text{and} \quad \beta \|\nabla_h v_h\|_0 \le \sup_{w_h \in V_D^{cr}} \frac{a_h(w_h, v_h)}{\|\nabla_h w_h\|_0} \quad \forall v_h \in V_D^{cr}. \tag{6.7}$$

It is then obvious that once we have (6.6) or (6.7), we also have the well-posedness of the Eq. (6.2).

With the help of the inf-sup condition of $a_h(\cdot, \cdot)$ (6.7), we can prove the coerciveness following the argument in [30].

Theorem 6.1 *Assume Assumption 5.1 is true. There exists an $h_0 > 0$, such that when the maximum mesh size h of $\mathcal{T}$ is smaller than h_0, the following inequality is true for $(\boldsymbol{\tau}_{rt}, v_{cr}) \in RT_{0,N} \times V_D^{cr}$:*

$$C |\!|\!| (\boldsymbol{\tau}_{rt}, v_{cr}) |\!|\!|^2 \le \mathcal{J}_h^{div}(\boldsymbol{\tau}_{rt}, v_{cr}; 0). \tag{6.8}$$

Proof By the discrete integration by parts (2.32), for $\boldsymbol{\tau}_{rt} \in RT_{0,N}$ and v_{cr} and w_{cr} in V_D^{cr}, we have

$$a_h(v_{cr}, w_{cr}) = = (A\nabla_h v_{cr} + \boldsymbol{\tau}_{rt}, \nabla_h w_{cr}) + (\nabla \cdot \boldsymbol{\tau}_{rt} + X_h v_{cr}, w_{cr}). \tag{6.9}$$

It follows from (6.7), (6.9), the Cauchy-Schwarz and discrete Poincaré inequalities, and the assumption on A, for any $(\boldsymbol{\tau}_{rt}, v_{cr}) \in RT_{0,N} \times V_D^{cr}$,

$$\begin{aligned}
\beta \|\nabla_h v_{cr}\|_0 &\le \sup_{w_{cr} \in V_D^{cr}} \frac{a_h(v_{cr}, w_{cr})}{\|\nabla_h w_{cr}\|_0} = \sup_{w_{cr} \in V_D^{cr}} \frac{(A\nabla_h v_{cr} + \boldsymbol{\tau}_{rt}, \nabla_h w_{cr}) + (\nabla \cdot \boldsymbol{\tau}_{rt} + X_h v_{cr}, w_{cr})}{\|\nabla_h w_{cr}\|_0} \\
&\le \sup_{w_{cr} \in V_D^{cr}} \frac{\|A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}\|_0 \|A^{1/2}\nabla_h w_{cr}\|_0 + \|\nabla \cdot \boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0 \|w_{cr}\|_0}{\|\nabla_h w_{cr}\|_0} \\
&\le C(\|A^{1/2}\nabla_h v_{cr} + A^{-1/2}\boldsymbol{\tau}_{rt}\|_0 + \|\nabla \cdot \boldsymbol{\tau}_{rt} + X_h v_{cr}\|_0).
\end{aligned}$$

The theorem is proved by combining the above result, (4.5), and (4.6). □

Remark 6.2 We present three proofs of the existence and uniqueness of the two-field potential-flux div CR-LSFEM (3.4). The first proof is based on Assumption 4.1, ensuring the coerciveness of the original variational problem. The restriction on the mesh size in the first proof is local and explicit. The regularity assumption is not needed for the first proof. However, the first proof can not be applied to the indefinite problems. On the other hand, the second and the third proofs are based on Assumption 5.1. It can be applied to more general indefinite cases once the uniqueness and a minimal regularity are assumed. Another restriction of the second and third proofs is that the regularity assumption is global, so the global mesh size h instead of the local mesh size is small enough is needed.

7 A Posteriori Estimates of Two-Field Div CR-LSFEM

7.1 A Negative Result on Norm Equivalence

First, we present a negative result on the norm equivalence of the least-squares functional $\mathcal{J}_h^{div}$ and the $|||(\boldsymbol{\tau}, v)|||$-norm for $v \in W_D^{1+cr}$ and $\boldsymbol{\tau} \in H(\mathrm{div}; \Omega)$. For simplicity, we only discuss the simple case that $A = I$, $\mathbf{b} = \mathbf{0}$, and $c = 0$.

Lemma 7.1 *The following inequality is not true,*

$$C|||(\boldsymbol{\tau}, v)|||^2 \le \|\nabla_h v + \boldsymbol{\tau}\|_0^2 + \|\nabla \cdot \boldsymbol{\tau}\|_0^2 \quad \forall (\boldsymbol{\tau}, v) \in H(\mathrm{div}; \Omega) \times W_D^{1+cr}. \tag{7.1}$$

Proof A counterexample in the two-dimensional case is presented. Let $0 \neq w_{cr} \in V_D^{cr}$, but $w_{cr} \notin H_D^1(\Omega)$. For example, choose $w_{cr}(m_F) = 1$, where m_F is the mid-point of an interior edge of a finite element mesh and let all other degrees of freedom of w_{cr} be zero. We have the following Helmholtz decomposition,

$$\nabla_h w_{cr} = \nabla\alpha + \nabla^\perp\beta \text{ and } \|\nabla_h w_{cr}\|_0^2 = \|\nabla\alpha\|_0^2 + \|\nabla^\perp\beta\|_0^2, \tag{7.2}$$

with $\alpha \in H_D^1(\Omega)$ and $\beta \in H_N^1(\Omega)$. Since $0 \neq w_{cr} \in V_D^{cr}$, but $w_{cr} \notin H_D^1(\Omega)$, we get $\nabla^\perp\beta \neq 0$.

Let $v = -\alpha + w_{cr} \in W_D^{1+cr}$ and $\boldsymbol{\tau} = \nabla_h w_{cr} - \nabla\alpha = -\nabla_h v = \nabla^\perp\beta$. Then $\nabla \cdot \boldsymbol{\tau} = \nabla \cdot (\nabla^\perp\beta) = 0$. Thus, we get

$$\begin{aligned} &\|\nabla_h v + \boldsymbol{\tau}\|_0^2 + \|\nabla \cdot \boldsymbol{\tau}\|_0^2 = 0 \quad \text{but} \\ &|||(\boldsymbol{\tau}, v)|||^2 = \|\nabla_h v\|_0^2 + \|\boldsymbol{\tau}\|_0^2 + \|\nabla \cdot \boldsymbol{\tau}\|_0^2 = 2\|\nabla^\perp\beta\|_0^2 > 0. \end{aligned}$$

We get a contradiction. □

Remark 7.2 A numerical counter-example with a non-zero Dirichlet boundary condition is given in Section 10.1.

7.2 A Posteriori Error Estimates

Lemma 7.1 provides a counterexample that $\mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_{cr}; f)$ itself cannot be used as a reliable a posteriori error estimator. From the proof of Lemma 7.1, we need to design an error estimator by adding the missing part.

Let E_h be the enriching operator defined in (2.23). Define

$$u_c = E_h u_{cr} \in S_{2,D}. \tag{7.3}$$

We introduce the following six a posteriori error estimators:

$$\eta_1^2 := \|A^{1/2}\nabla_h(u_{cr}-u_c)\|_0^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_c; f),$$
$$\eta_2^2 := \|A^{1/2}\nabla_h(u_{cr}-u_c)\|_0^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_{cr}; f),$$
$$\eta_3^2 := \sum_{F\in\mathcal{E}} \frac{1}{h_F}\|[\![u_{cr}]\!]\|_{0,F}^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_c; f),$$
$$\eta_4^2 := \sum_{F\in\mathcal{E}} \frac{1}{h_F}\|[\![u_{cr}]\!]\|_{0,F}^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_{cr}; f),$$
$$\eta_5^2 := \sum_{F\in\mathcal{E}} h_F\|\gamma_{t_F}(u_{cr})\|_{0,F}^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_c; f),$$
$$\eta_6^2 := \sum_{F\in\mathcal{E}} h_F\|\gamma_{t_F}(u_{cr})\|_{0,F}^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_{cr}; f).$$

We have the following reliability and efficiency results.

Theorem 7.3 *Assume one of the conditions of Theorem 2.3 is true. Let* $(\boldsymbol{\sigma}_{rt}, u_{cr}) \in RT_{0,N}\times V_D^{cr}$ *be the finite element solution to the problem (3.7). For* $i = 1, \cdots, 6$, *there exist positive constants* C_1 *and* C_2 *such that the following inequalities hold:*

$$C_1\eta_i \le |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_{cr})|\!|\!| \le C_2\eta_i. \tag{7.4}$$

Proof Due to the norm equivalence (3.5) on $H_N(\mathrm{div};\Omega)\times H_D^1(\Omega)$ and the fact that $u_c = E_h u_{cr} \in S_{2,D} \subset H_D^1(\Omega)$, we have

$$C_1|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_c)|\!|\!|^2 \le \mathcal{J}_h^{div}(\boldsymbol{\sigma}_h, u_c; f) = \mathcal{J}_h^{div}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_c; 0)$$
$$\le C_2|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_c)|\!|\!|^2.$$

Then by the triangle inequality,

$$|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_{cr})|\!|\!|^2 \le |\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_c)|\!|\!|^2 + C\|A^{1/2}\nabla_h(u_{cr}-u_c)\|_0^2 \le C\eta_1^2.$$

The reliability result of η_1 is proved.

From (2.25) and (2.28), we have

$$C\|\nabla_h(u_{cr}-u_c)\|_0^2 \le \sum_{F\in\mathcal{E}}\frac{1}{h_F}\|[\![u_{cr}]\!]\|_{0,F}^2 \le C\|\nabla_h(u_{cr}-u_c)\|_0^2, \tag{7.5}$$

$$\sum_{F\in\mathcal{E}}\frac{1}{h_F}\|[\![u_{cr}]\!]\|_{0,F}^2 \le C\|\nabla_h(u-u_{cr})\|_0^2, \tag{7.6}$$

With the triangle inequality, (7.5), and (7.6),

$$\eta_1^2 = \|A^{1/2}\nabla_h(u_{cr}-u_c)\|_0^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_c; f)$$
$$\le C\|A^{1/2}\nabla_h(u_{cr}-u_c)\|_0^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_{cr}; f)$$
$$\le C\|A^{1/2}\nabla_h(u-u_{cr})\|_0^2 + \mathcal{J}_h^{div}(\boldsymbol{\sigma}_{rt}, u_{cr}; f) \le C|\!|\!|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{rt}, u-u_{cr})|\!|\!|^2$$

The efficiency of η_1 is also proved.

By the triangle inequality, it is easy to see

$$\eta_1^2 \le \eta_2^2 + \|A^{1/2}\nabla_h(u_{cr}-u_c)\|_0^2 \le 2\eta_2^2.$$

Similarly, $\eta_2^2 \le 2\eta_1^2$. So η_2 is both reliable and efficient.

We get from (7.5) that η_3 and η_4 are both reliable and efficient. The reliability and efficiency of η_5 and η_6 can be obtained from (2.31). □

Remark 7.4 In this remark, we explain the terms in the error estimators. The error estimator contains three parts. The major part is the residual, that is, $\|\nabla \cdot \boldsymbol{\sigma}_{rt} + X_h u_{cr} - f\|_0$ or $\|\nabla \cdot \boldsymbol{\sigma}_{rt} + X u_c - f\|_0$. The other terms represent the fact that the numerical solutions are not in the right spaces. As discussed in [37, 56, 57], we have

$$\boldsymbol{\sigma} = -A\nabla u \in H(\mathrm{div};\Omega), \quad u \in H^1(\Omega), \quad \text{and} \quad \nabla u \in H(\mathrm{curl};\Omega). \tag{7.7}$$

The result $\nabla u \in H(\mathrm{curl};\Omega)$ is obtained from $u \in H^1(\Omega)$. While numerically,

$$-A\nabla_h u_{cr} \notin H(\mathrm{div};\Omega), \quad -A\nabla u_c \notin H(\mathrm{div};\Omega), \quad u_{cr} \notin H^1(\Omega), \quad \text{and} \quad \nabla_h u_{cr} \notin H(\mathrm{curl};\Omega). \tag{7.8}$$

The terms $\|A^{-1/2}\boldsymbol{\sigma}_{rt} + A^{1/2}\nabla u_{cr}\|_0$ and $\|A^{-1/2}\boldsymbol{\sigma}_{rt} + A^{1/2}\nabla u_c\|_0$ measure the distance of the numerical flux from u_{cr} $(A\nabla_h u_{cr})$ or u_c $(A\nabla u_c)$ to $H(\mathrm{div};\Omega)$ space. The term $\sum_{F\in\mathcal{E}} \frac{1}{h_F}\|[\![u_{cr}]\!]\|^2_{0,F}$ measures the distance of the numerical solution u_{cr} to $H^1(\Omega)$ space. The term $\sum_{F\in\mathcal{E}} h_F\|\gamma_{t_F}(u_{cr})\|^2_{0,F}$ measures the distance of the numerical gradient $\nabla_h u_{cr}$ to $H(\mathrm{curl};\Omega)$ space. Due to that fact that $\nabla u \in H(\mathrm{curl};\Omega)$ is actually obtained from $u \in H^1(\Omega)$ and the equivalence (2.31), we can use either $\sum_{F\in\mathcal{E}} \frac{1}{h_F}\|[\![u_{cr}]\!]\|^2_{0,F}$ or $\sum_{F\in\mathcal{E}} h_F\|\gamma_{t_F}(u_{cr})\|^2_{0,F}$ to measure this violation.

From the proof of Lemma 7.1, the curl part of the Helmholtz decomposition (7.2) is not controlled by the div least-squares functional. This part is exactly the so-called nonconforming error. The necessity of measuring the violation that $u_{cr} \notin H^1(\Omega)$ or $\nabla_h u_{cr} \notin H(\mathrm{curl};\Omega)$ is discussed in the literature of a posteriori error estimates of nonconforming finite elements, see [35–41].

This also hints at how to design a least-squares functional with the norm equivalence on $W_D^{1+cr} \times H(\mathrm{div};\Omega)$. An equation related to $\nabla u \in H(\mathrm{curl})$ should be added. We will discuss it in the next section.

Remark 7.5 Compared to η_i, $i = 3,\ldots,6$, the estimators η_1 and η_2 use the matrix A, thus they will be more robust with respect to the coefficient.

8 Three-Field Potential-Flux-Intensity Div-Curl LSFEM with Nonconforming Approximation for General Elliptic Equations

Due to the lack of the norm equivalence in the abstract nonconforming piecewise H^1-space $W_D(\mathcal{T})$, the potential-flux div least-squares CR method does not have the automatic discrete stability and a built-in a posteriori error estimator with the least-squares functional. Inspired by the a posteriori error analysis discussed in the previous section, we propose three-field formulations, potential-flux-density div-curl least-squares methods. An intensity field and its curl equation are added to the least-squares formulation. In this new formulation, we can prove norm equivalence for the abstract nonconforming piecewise H^1-space. Thus, we recover the good qualities of the original LSFEM: automatically discrete stability without a mesh size requirement and regularity assumptions, and a built-in least-squares a posteriori error estimator.

8.1 Three-Field Potential-Flux-Intensity Div-Curl Least-Squares Methods

It is easy to see that only two of the last three equations in (2.22) are independent. Thus, we only need two of them to construct a least-squares functional. We will discuss one case first. The results can be easily generalized to the other two cases; see subsection 8.4.

To simplify the notations, let

$$\begin{aligned}\mathbb{Y} &:= H_N(\mathrm{div};\Omega)\times H_D(\mathrm{curl};\Omega)\times W_D(\mathcal{T}) \quad \text{and} \quad \|(\boldsymbol{\tau},\boldsymbol{\psi},v)\|_{\mathbb{Y}}^2\\ &:= \|\boldsymbol{\tau}\|_{H(\mathrm{div})}^2+\|\boldsymbol{\psi}\|_{H(\mathrm{curl})}^2+\|\nabla_h v\|_0^2.\end{aligned} \tag{8.1}$$

Define the following least-squares functional: For $(\boldsymbol{\tau},\boldsymbol{\psi},v)\in\mathbb{Y}$,

$$\begin{aligned}\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau},\boldsymbol{\psi},v;f) &:= \|A^{-1/2}(\boldsymbol{\tau}+A\nabla_h v)\|_0^2+\|A^{-1/2}\boldsymbol{\tau}-A^{1/2}\boldsymbol{\psi}\|_0^2\\ &\quad +\|\nabla\times\boldsymbol{\psi}\|_0^2+\|\nabla\cdot\boldsymbol{\tau}+X_h v-f\|_0^2.\end{aligned} \tag{8.2}$$

Then the three-field potential-flux-intensity div-curl least-squares minimization problem in the abstract space is: Find $(\boldsymbol{\sigma},\boldsymbol{\phi},u)\in\mathbb{Y}$, such that

$$\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\sigma},\boldsymbol{\phi},u;f)=\inf_{(\boldsymbol{\tau},\boldsymbol{\psi},v)\in\mathbb{Y}}\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau},\boldsymbol{\psi},v;f). \tag{8.3}$$

For $(\boldsymbol{\chi},,w)\in\mathbb{Y}$ and $(\boldsymbol{\tau},\boldsymbol{\psi},v)\in\mathbb{Y}$, define the following bilinear forms c_h:

$$\begin{aligned}c_h((\boldsymbol{\chi},,w),(\boldsymbol{\tau},\boldsymbol{\psi},v)) &:= (A\nabla_h w+\boldsymbol{\chi},\nabla_h v+A^{-1}\boldsymbol{\tau})+(\boldsymbol{\chi}-A,A^{-1}\boldsymbol{\tau}-\boldsymbol{\psi})\\ &\quad +(\nabla\times,\nabla\times\boldsymbol{\psi})+(\nabla\cdot\boldsymbol{\chi}+X_h w,\nabla\cdot\boldsymbol{\tau}+X_h v).\end{aligned}$$

Then the least-squares variational problems are: Find $(\boldsymbol{\sigma},\boldsymbol{\phi},u)\in\mathbb{Y}$,

$$c_h((\boldsymbol{\sigma},\boldsymbol{\phi},u),(\boldsymbol{\tau},\boldsymbol{\psi},v))=(f,\nabla\cdot\boldsymbol{\tau}+X_h v)\quad\forall(\boldsymbol{\tau},\boldsymbol{\psi},v)\in\mathbb{Y}. \tag{8.4}$$

It is clear the exact solution is $(\boldsymbol{\sigma},\boldsymbol{\phi},u)$. The following continuity is also easy to verify:

$$c_h((\boldsymbol{\chi},,w),(\boldsymbol{\tau},\boldsymbol{\psi},v))\le C\|(\boldsymbol{\chi},,w)\|_{\mathbb{Y}}\|(\boldsymbol{\tau},\boldsymbol{\psi},v)\|_{\mathbb{Y}}\quad(\boldsymbol{\chi},,w)\in\mathbb{Y},(\boldsymbol{\tau},\boldsymbol{\psi},v)\in\mathbb{Y}. \tag{8.5}$$

8.2 Coerciveness of Three-Field Div-Curl Least-Squares Method

For a regular mesh $\mathcal{T}$ with any mesh size, we want to prove the following norm equivalence under the minimal assumption that one of the conditions of Theorem 2.3 is true:

$$C\|(\boldsymbol{\tau},\boldsymbol{\psi},v)\|_{\mathbb{Y}}^2\le\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau},\boldsymbol{\psi},v;0)\le C\|(\boldsymbol{\tau},\boldsymbol{\psi},v)\|_{\mathbb{Y}}^2\quad\forall(\boldsymbol{\tau},\boldsymbol{\psi},v)\in\mathbb{Y}. \tag{8.6}$$

A careful look into the proofs of the coerciveness of the least-squares methods in [7, 27, 28] for the general elliptic equations will find that the compactness argument plays a central role in these proofs. In a compactness argument, the proof by contradiction is used. However, for a method depending on a discrete mesh $\mathcal{T}$, we cannot use the proof by contradiction to show the coerciveness constant is independent of the mesh. Thus, we must seek a proof without using a mesh-dependent space like $W_D(\mathcal{T})$ or W_D^{1+cr}.

To this end, we use the Helmholtz decomposition to change the formulation into a mesh-independent setting. Suppose the conditions of Lemmas 2.4 and 2.5 hold, by Theorem 2.7, for $v\in W_D(\mathcal{T})$, there exists $p\in H_D^1(\Omega)$ and $\mathtt{q}\in\mathbb{Q}$ such that

$$A\nabla_h v=A\nabla p+\mathtt{Curl}\,\mathtt{q}\ \text{and}\ \|A^{-1/2}\nabla_h v\|_0^2=\|A^{1/2}\nabla p\|_0^2+\|A^{-1/2}\mathtt{Curl}\,\mathtt{q}\|_0^2. \tag{8.7}$$

To simplify notations, let

$$\mathbb{Z} := H_N(\mathrm{div};\Omega)\times H_D(\mathrm{curl};\Omega)\times H_D^1(\Omega)\times\mathbb{Q},$$
$$\|(\boldsymbol{\tau},\boldsymbol{\psi},p,\mathtt{q})\|_{\mathbb{Z}}^2 := \|\boldsymbol{\tau}\|_{H(\mathrm{div})}^2+\|\boldsymbol{\psi}\|_{H(\mathrm{curl})}^2+\|\nabla p\|_0^2+\|\mathtt{Curl}\,\mathtt{q}\|_0^2.$$

For $(\boldsymbol{\tau},\boldsymbol{\psi},p,\mathtt{q})\in\mathbb{Z}$, define

$$\begin{aligned}L(\boldsymbol{\tau},\boldsymbol{\psi},p,\mathtt{q}) = {}& \|A^{-1/2}(\boldsymbol{\tau}+A\nabla p+\mathtt{Curl}\,\mathtt{q})\|_0^2+\|A^{-1/2}\boldsymbol{\tau}-A^{1/2}\boldsymbol{\psi}\|_0^2\\ &+\|\nabla\times\boldsymbol{\psi}\|_0^2+\|\nabla\cdot\boldsymbol{\tau}+Xp+\mathbf{b}\cdot A^{-1}\mathtt{Curl}\,\mathtt{q}\|_0^2.\end{aligned}$$

Lemma 8.1 *Assuming that one of the conditions of Theorem 2.3 is true, we have the following coerciveness:*

$$C\|(\boldsymbol{\tau},\boldsymbol{\psi},p,\mathtt{q})\|_{\mathbb{Z}}^2\le L(\boldsymbol{\tau},\boldsymbol{\psi},p,\mathtt{q})\quad\forall(\boldsymbol{\tau},\boldsymbol{\psi},p,\mathtt{q})\in\mathbb{Z}. \tag{8.8}$$

Proof By the standard coerciveness (3.5) and the triangle inequality, we have

$$\begin{aligned}&C(\|\boldsymbol{\tau}\|_{H(\mathrm{div})}+\|\nabla p\|_0)\le\|A^{-1/2}\boldsymbol{\tau}+A^{1/2}\nabla p\|_0+\|\nabla\cdot\boldsymbol{\tau}+Xp\|_0\\ &\le\|A^{-\frac{1}{2}}(\boldsymbol{\tau}+A\nabla p+\mathtt{Curl}\,\mathtt{q})\|_0+\|\nabla\cdot\boldsymbol{\tau}+Xp+\mathbf{b}\cdot A^{-1}\mathtt{Curl}\,\mathtt{q}\|_0+C\|\mathtt{Curl}\,\mathtt{q}\|_0.\end{aligned} \tag{8.9}$$

By (2.35), (2.36), the Poincaré-Friedrichs inequality (2.34), Cauchy-Schwarz inequality, and the property of the coefficient matrix A, we have

$$\begin{aligned}&\|A^{-1/2}\mathtt{Curl}\,\mathtt{q}\|_0^2=(A^{-1}\mathtt{Curl}\,\mathtt{q}+\nabla p\\ &\quad+A^{-1}\boldsymbol{\tau},\mathtt{Curl}\,\mathtt{q})-(A^{-1}\boldsymbol{\tau}-\boldsymbol{\psi},\mathtt{Curl}\,\mathtt{q})-(\boldsymbol{\psi},\mathtt{Curl}\,\mathtt{q})\\ &=(A^{-1}\mathtt{Curl}\,\mathtt{q}+\nabla p+A^{-1}\boldsymbol{\tau},\mathtt{Curl}\,\mathtt{q})-(A^{-1}\boldsymbol{\tau}-\boldsymbol{\psi},\mathtt{Curl}\,\mathtt{q})+(\nabla\times\boldsymbol{\psi},\mathtt{q})\\ &\le C(\|A^{-1/2}\mathtt{Curl}\,\mathtt{q}+A^{1/2}\nabla p+A^{-1/2}\boldsymbol{\tau}\|_0\\ &\quad+\|A^{-1/2}\boldsymbol{\tau}-A^{1/2}\boldsymbol{\psi}\|_0+\|\nabla\times\boldsymbol{\psi}\|_0)\|A^{-1/2}\mathtt{Curl}\,\mathtt{q}\|_0.\end{aligned}$$

Thus,

$$C\|\mathtt{Curl}\,\mathtt{q}\|_0\le\|A^{-1/2}\mathtt{Curl}\,\mathtt{q}+A^{1/2}\nabla p+A^{-1/2}\boldsymbol{\tau}\|_0+\|A^{-1/2}\boldsymbol{\tau}-A^{1/2}\boldsymbol{\psi}\|_0+\|\nabla\times\boldsymbol{\psi}\|_0. \tag{8.10}$$

We then get the following inequality:

$$\|\mathtt{Curl}\,\mathtt{q}\|_0^2\le CL(\boldsymbol{\tau},\boldsymbol{\psi},p,q).$$

The term $\|\boldsymbol{\psi}\|_0$ can be bounded by the triangle inequality:

$$\|\boldsymbol{\psi}\|_0\le C\|A^{-1/2}\boldsymbol{\tau}-A^{1/2}\boldsymbol{\psi}\|_0+C\|\boldsymbol{\tau}\|_0.$$

Combined with the above results and using the fact $\|\nabla\times\boldsymbol{\psi}\|_0$ is a part of $L(\boldsymbol{\tau},\boldsymbol{\psi},p,q)$, the lemma is proved. □

Theorem 8.2 *Assuming that one of the conditions of Theorem 2.3 is true, and the conditions of Lemmas 2.4 and 2.5 are true, we have the following norm equivalence,*

$$C\|(\boldsymbol{\tau},\boldsymbol{\psi},v)\|_{\mathbb{Y}}^2\le\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau},\boldsymbol{\psi},v;0)\le C\|(\boldsymbol{\tau},\boldsymbol{\psi},v)\|_{\mathbb{Y}}^2\quad\forall(\boldsymbol{\tau},\boldsymbol{\psi},v)\in\mathbb{Y}. \tag{8.11}$$

Proof The upper bound is straightforward with the triangle inequality and the discrete Poincaré-Friedrichs inequality (2.38).

We will focus on proving the coerciveness. By the Helmholtz decomposition (8.7), the result of Lemma 8.1, the triangle inequality, and the nonconforming Poincaré-Friedrichs inequality (2.38).

$$\begin{aligned}
&\|A^{1/2}\nabla_h v\|_0^2 + \|\boldsymbol{\tau}\|_{H(\mathrm{div})}^2 + \|\boldsymbol{\psi}\|_{H(\mathrm{curl})}^2 = \|A^{1/2}\nabla p\|_0^2 + \|A^{-1/2}\mathtt{Curl}\,\mathfrak{q}\|_0^2 \\
&\quad + \|\boldsymbol{\tau}\|_{H(\mathrm{div})}^2 + \|\boldsymbol{\psi}\|_{H(\mathrm{curl})}^2 \\
&\le CL(\boldsymbol{\tau}, \boldsymbol{\psi}, p, \mathfrak{q}) \\
&= C\big(\|A^{-1/2}(\boldsymbol{\tau} + A\nabla p + \mathtt{Curl}\,\mathfrak{q})\|_0^2 + \|A^{-1/2}\boldsymbol{\tau} - A^{1/2}\boldsymbol{\psi}\|_0^2 + \|\nabla\times\boldsymbol{\psi}\|_0^2 + \|\nabla\cdot\boldsymbol{\tau} \\
&\quad + Xp + \mathbf{b}\cdot A^{-1}\mathtt{Curl}\,\mathfrak{q}\|_0^2 \\
&= C(\|A^{-1/2}(\boldsymbol{\tau} + A\nabla_h v)\|_0^2 + \|A^{-1/2}\boldsymbol{\tau} - A^{1/2}\boldsymbol{\psi}\|_0^2 \\
&\quad + \|\nabla\times\boldsymbol{\psi}\|_0^2 + \|\nabla\cdot\boldsymbol{\tau} + X_h v + c(p - v)\|_0^2) \\
&\le C(\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0) + \|p - v\|_0^2) \le C(\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0) + \|\nabla p - \nabla_h v\|_0^2) \\
&\le C(\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0) + \|A^{-1}\mathtt{Curl}\,\mathfrak{q}\|_0^2).
\end{aligned}$$

By (8.10), we also have

$$\begin{aligned}
C\|A^{-1/2}\mathtt{Curl}\,\mathfrak{q}\|_0^2 \le \|A^{-1/2}\mathtt{Curl}\,\mathfrak{q} + A^{1/2}\nabla p + A^{-1/2}\boldsymbol{\tau}\|_0^2 + \|A^{-1/2}\boldsymbol{\tau} \\
- A^{1/2}\boldsymbol{\psi}\|_0^2 + \|\nabla\times\boldsymbol{\psi}\|_0^2 \le \mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0).
\end{aligned}$$

The theorem is then proved. □

8.3 Three-Field Potential-Flux-Intensity Div-Curl Nonconforming Least-Squares Finite Element Methods

Let

$$\mathbb{Y}_h := RT_{0,N} \times N_{0,D} \times V_D^{cr}.$$

We have $\mathbb{Y}_h \subset \mathbb{Y}$. The three-field div-curl nonconforming LSFEM is to find $(\boldsymbol{\sigma}_h, \boldsymbol{\phi}_h, u_h) \in \mathbb{Y}_h$, such that

$$\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\sigma}_h, \boldsymbol{\phi}_h, u_h; f) = \inf_{(\boldsymbol{\tau}_h, \boldsymbol{\psi}_h, v_h)\in\mathbb{Y}_h} \mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau}_h, \boldsymbol{\psi}_h, v_h; f), \tag{8.12}$$

or equivalently in a weak form: Find $(\boldsymbol{\sigma}_h, \boldsymbol{\phi}_h, u_h) \in \mathbb{Y}_h$, such that

$$c_h((\boldsymbol{\sigma}_h, \boldsymbol{\phi}_h, u_h), (\boldsymbol{\tau}_h, \boldsymbol{\psi}_h, v_h)) = (f, \nabla\cdot\boldsymbol{\tau}_h + X_h v_h) \quad \forall (\boldsymbol{\tau}_h, \boldsymbol{\psi}_h, v_h) \in \mathbb{Y}_h. \tag{8.13}$$

It is easy to see we have the error equation,

$$c_h((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\phi} - \boldsymbol{\phi}_h, u - u_h), (\boldsymbol{\tau}_h, \boldsymbol{\psi}_h, v_h)) = 0 \quad \forall (\boldsymbol{\tau}_h, \boldsymbol{\psi}_h, v_h) \in \mathbb{Y}_h,$$

and the a priori error estimate: Assuming $u \in H^2(\Omega)$ (thus $\boldsymbol{\phi} \in H^1(\Omega)$ and we always have $\nabla\times\boldsymbol{\phi} = 0$), $\boldsymbol{\sigma}|_K \in H^1(K)$, and $\nabla\cdot\boldsymbol{\sigma}|_K \in H^1(K)$, for $K \in \mathcal{T}$, then

$$\begin{aligned}
\|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\phi} - \boldsymbol{\phi}_h, u - u_h)\|_{\mathbb{Y}} &\le C \inf_{(\boldsymbol{\tau}_h, \boldsymbol{\psi}_h, v_h)\in\mathbb{Y}_h} \|(\boldsymbol{\sigma} - \boldsymbol{\tau}_h, \boldsymbol{\phi} - \boldsymbol{\psi}_h, u - v_h)\|_{\mathbb{Y}} \\
&\le Ch(\|u\|_2 + \|\boldsymbol{\sigma}\|_{1,h} + \|\nabla\cdot\boldsymbol{\sigma}\|_{1,h}),
\end{aligned}$$

where $\|v\|_{1,h}^2 = \sum_{K\in\mathcal{T}} \|v\|_{1,K}^2$.

Remark 8.3 Using the more refined analysis of local interpolations or local quasi-interpolations of Nédélec elements [58, 59], we can also show a local optimal error estimate.

Let $(\boldsymbol{\sigma}_h, \boldsymbol{\phi}_h, u_h) \in \mathbb{Y}_h$ be the numerical solution of (8.12). Define the following a posteriori error estimator:

$$\zeta^2 := \mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\sigma}_h, \boldsymbol{\phi}_h, u_h; f). \tag{8.14}$$

Due to the norm equivalence (8.11), we have the global reliability and efficiency of ζ:

$$C_1\zeta^2 \le \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\phi} - \boldsymbol{\phi}_h, u - u_h)\|_{\mathbb{Y}}^2 \le C_2\zeta^2. \tag{8.15}$$

Remark 8.4 Since $-\boldsymbol{\phi}$ and $-A^{-1}\boldsymbol{\sigma}$ are both ∇u, the term $X_h v$ in the above definitions of the least-squares functional can also be replaced by

$$G_{h,rst}(\boldsymbol{\psi}, \boldsymbol{\tau}, v) := \mathbf{b} \cdot (r\nabla_h v - s\boldsymbol{\psi} - tA^{-1}\boldsymbol{\tau}) + cv \quad \forall (\boldsymbol{\tau}, \boldsymbol{\psi}, v) \in \mathbb{Y}, \tag{8.16}$$

where $(r, s, t) \in [0, 1]^3$ and $r + s + t = 1$. We can get similar results of norm equivalence and a priori and a posteriori error estimates.

8.4 Two Variants of Three-Field Potential-Flux-Intensity Div-Curl Nonconforming Least-Squares Method

We can also use other combinations in the first-order system (2.22) to define least-squares functionals. Define the following least-squares functionals: For $(\boldsymbol{\tau}, \boldsymbol{\psi}, v) \in \mathbb{Y}$,

$$\begin{aligned}
\mathcal{J}_{h,2}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; f) &:= \|A^{1/2}(\nabla_h v + \boldsymbol{\psi})\|_0^2 + \|A^{-1/2}\boldsymbol{\tau} \\
&\quad - A^{1/2}\boldsymbol{\psi}\|_0^2 + \|\nabla \times \boldsymbol{\psi}\|_0^2 + \|\nabla \cdot \boldsymbol{\tau} + X_h v - f\|_0^2, \\
\mathcal{J}_{h,3}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; f) &:= \|A^{-1/2}(\boldsymbol{\tau} + A\nabla_h v)\|_0^2 + \|A^{1/2}(\nabla_h v + \boldsymbol{\psi})\|_0^2 \\
&\quad + \|\nabla \times \boldsymbol{\psi}\|_0^2 + \|\nabla \cdot \boldsymbol{\tau} + X_h v - f\|_0^2.
\end{aligned}$$

We then prove the equivalence of different $\mathcal{J}_{h,i}$. We use the notation $B \approx D$ to denote there exist constants C_1 and C_2, such that $C_1 D \le B \le C_2 D$.

Lemma 8.5 *The following equivalences are true:*

$$\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0) \approx \mathcal{J}_{h,2}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0) \approx \mathcal{J}_{h,3}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0) \quad \forall (\boldsymbol{\tau}, \boldsymbol{\psi}, v) \in \mathbb{Y}. \tag{8.17}$$

Proof By the triangle inequality, we have

$$\|A^{1/2}(\nabla_h v + \boldsymbol{\psi})\|_0 \le \|A^{-1/2}\boldsymbol{\tau} + A^{1/2}\nabla_h v\|_0 + \|A^{-1/2}\boldsymbol{\tau} - A^{1/2}\boldsymbol{\psi}\|_0.$$

Thus $\mathcal{J}_{h,2}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0) \le C\mathcal{J}_{h,1}^{divcurl}(\boldsymbol{\tau}, \boldsymbol{\psi}, v; 0)$ is true for all $(\boldsymbol{\tau}, \boldsymbol{\psi}, v) \in \mathbb{Y}$. Similar arguments can prove the other results. □

We thus can define potential-flux-intensity div-curl nonconforming least-squares methods based on functionals $\mathcal{J}_{h,2}^{divcurl}$ and $\mathcal{J}_{h,3}^{divcurl}$. Due to the equivalence (8.17), all results of properties related to $\mathcal{J}_{h,1}^{divcurl}$ can be generalized to methods defined by $\mathcal{J}_{h,2}^{divcurl}$ or $\mathcal{J}_{h,3}^{divcurl}$.

9 A Two-Field Potential–Flux Div-Curl LSFEM with Nonconforming Approximation

In this section, we discuss the application and restrictions of the original potential-flux div-curl least-squares method [2] in a nonconforming setting. When the domain is nice, and the coefficient is sufficiently smooth, the original formulation introduced in [2] can be used in the nonconforming case. But this two-field formulation can cause serious problems when the conditions on the domain and coefficients are not satisfied.

9.1 Two-Field Potential-Flux Div-Curl First-Order System

In the original paper [2], the intensity $\boldsymbol{\phi} = -\nabla u$ is not introduced as an independent variable; instead, the following first-order system is discussed.

$$\nabla \cdot \boldsymbol{\sigma} + Xu = f \text{ in } \Omega, \quad A\nabla u + \boldsymbol{\sigma} = 0 \text{ in } \Omega, \quad \nabla \times (A^{-1}\boldsymbol{\sigma}) = 0 \text{ in } \Omega, \tag{9.1}$$

with boundary conditions $u = 0$ on Γ_D, $\gamma_t(A^{-1}\boldsymbol{\sigma}) = 0$ on Γ_D, and $\mathbf{n} \cdot \boldsymbol{\sigma} = 0$ on Γ_N. Let

$$\begin{aligned} \boldsymbol{\Sigma} &:= \{\boldsymbol{\tau} : \boldsymbol{\tau} \in H_N(\mathrm{div}; \Omega),\ A^{-1}\boldsymbol{\tau} \in H_D(\mathrm{curl}; \Omega)\} \quad \text{with } \|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}}^2 \\ &:= \|\boldsymbol{\tau}\|_0^2 + \|\nabla \cdot \boldsymbol{\tau}\|_0^2 + \|\nabla \times (A^{-1}\boldsymbol{\tau})\|_0^2. \end{aligned}$$

Then the exact flux $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$.

In this section, we assume the following assumption is true. Detailed discussion on this norm equivalence can be found in [2], see also discussion in Section A.3.1 of [4].

Assumption 9.1 We assume that $A \in C^{1,1}$ and the domain Ω is nice enough to guarantee that $\boldsymbol{\Sigma}$ is algebraically and topologically included in $H^1(\Omega)^d$, that is, the following norm equivalence holds:

$$C_1\|\boldsymbol{\tau}\|_1 \le \|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}} \le C_2\|\boldsymbol{\tau}\|_1 \quad \forall \boldsymbol{\tau} \in \boldsymbol{\Sigma}. \tag{9.2}$$

Here, the domain is nice enough means one of the following is true (Theorem A.8 of [4]):

- $\partial\Omega$ is of class $C^{1,1}$,
- $\partial\Omega$ is piecewise smooth with no reentrant corners for $d = 2$,
- Ω is a convex polyhedron for $d = 3$.

9.2 A Two-Field Potential-Flux Div-Curl Least-Squares Method

Now, the two-field potential-flux div-curl least-squares functional is given by

$$\begin{aligned} \mathcal{G}_h(\boldsymbol{\tau}, v; f) &:= \|A^{-1/2}\boldsymbol{\tau} + A^{1/2}\nabla_h v\|_0^2 \\ &\quad + \|\nabla \cdot \boldsymbol{\tau} + X_h v - f\|_0^2 + \|\nabla \times (A^{-1}\boldsymbol{\tau})\|_0^2 \quad \forall (\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T}) \end{aligned} \tag{9.3}$$

Then the potential-flux div-curl least-squares minimization problem in the abstract space is:

$$\text{Find } (\boldsymbol{\sigma}, u) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T}), \text{ such that } \mathcal{G}_h(\boldsymbol{\sigma}, u; f) = \inf_{(\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T})} \mathcal{G}_h(\boldsymbol{\tau}, v; f). \tag{9.4}$$

For $(\boldsymbol{\chi}, w) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T})$ and $(\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T})$, define the following bilinear forms d_h:

$$\begin{aligned} d_h((\boldsymbol{\chi}, w), (\boldsymbol{\tau}, v)) &:= (A\nabla_h w + \boldsymbol{\chi}, \nabla_h v + A^{-1}\boldsymbol{\tau}) \\ &\quad + (\nabla \times (A^{-1}\boldsymbol{\chi}), \nabla \times (A^{-1}\boldsymbol{\tau})) + (\nabla \cdot \boldsymbol{\chi} + X_h w, \nabla \cdot \boldsymbol{\tau} + X_h v). \end{aligned}$$

Then the two-field potential-flux div-curl least-squares variational problems is: Find $(\boldsymbol{\sigma}, u) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T})$,

$$d_h((\boldsymbol{\sigma}, u), (\boldsymbol{\tau}, v)) = (f, \nabla \cdot \boldsymbol{\tau} + X_h v) \quad \forall (\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T}). \tag{9.5}$$

We use almost the same argument of the three-field potential-flux-curl least-squares method to prove the coerciveness of $\mathcal{G}_h(\boldsymbol{\tau}, v; 0)$. We assume the same Helmholtz decomposition (8.7) holds. Let

$$\begin{aligned} \mathbb{T} &:= \boldsymbol{\Sigma} \times H_D^1(\Omega) \times \mathbb{Q} \quad \text{and} \quad \|(\boldsymbol{\tau}, p, \mathtt{q})\|_{\mathbb{T}}^2 \\ &:= \|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}}^2 + \|\nabla p\|_0^2 + \|\mathtt{Curl}\, \mathtt{q}\|_0^2. \end{aligned}$$

For $(\boldsymbol{\tau}, p, \mathtt{q}) \in \mathbb{T}$, define

$$\begin{aligned} M(\boldsymbol{\tau}, p, \mathtt{q}) := &\|A^{-1/2}(\boldsymbol{\tau} + A\nabla p + \mathtt{Curl}\, \mathtt{q})\|_0^2 + \|\nabla \times (A^{-1}\boldsymbol{\tau})\|_0^2 \\ &+ \|\nabla \cdot \boldsymbol{\tau} + Xp + \mathbf{b} \cdot A^{-1}\mathtt{Curl}\, \mathtt{q}\|_0^2. \end{aligned}$$

Lemma 9.2 *Assuming that one of the conditions of Theorem 2.3 is true and Assumption 9.1 is true, we have the following coerciveness:*

$$C\|(\boldsymbol{\tau}, p, \mathtt{q})\|_{\mathbb{T}}^2 \le M(\boldsymbol{\tau}, p, \mathtt{q}) \quad \forall (\boldsymbol{\tau}, p, \mathtt{q}) \in \mathbb{T}. \tag{9.6}$$

Proof By (2.35), (2.36), the Poincaré-Friedrichs inequality (2.34), the Cauchy-Schwarz inequality, and the property of the coefficient matrix A, we have

$$\begin{aligned} \|A^{-1/2}\mathtt{Curl}\, \mathtt{q}\|_0^2 &= (A^{-1}\mathtt{Curl}\, \mathtt{q} + \nabla p + A^{-1}\boldsymbol{\tau}, \mathtt{Curl}\, \mathtt{q}) - (A^{-1}\boldsymbol{\tau}, \mathtt{Curl}\, \mathtt{q}) \\ &= (A^{-1}\mathtt{Curl}\, \mathtt{q} + \nabla p + A^{-1}\boldsymbol{\tau}, \mathtt{Curl}\, \mathtt{q}) + (\nabla \times ((A^{-1}\boldsymbol{\tau}), \mathtt{q}) \\ &\le C(\|A^{-1/2}\mathtt{Curl}\, \mathtt{q} + A^{1/2}\nabla p + A^{-1/2}\boldsymbol{\tau}\|_0 \\ &\quad + \|\nabla \times (A^{-1}\boldsymbol{\tau})\|_0)\|A^{-1/2}\mathtt{Curl}\, \mathtt{q}\|_0. \end{aligned}$$

Then, we get $C\|\mathtt{Curl}\, \mathtt{q}\|_0 \le \|A^{-1/2}\mathtt{Curl}\, \mathtt{q} + A^{1/2}\nabla p + A^{-1/2}\boldsymbol{\tau}\|_0 + \|\nabla \times (A^{-1}\boldsymbol{\tau})\|_0$. Thus, $\|\mathtt{Curl}\, \mathtt{q}\|_0^2 \le CM(\boldsymbol{\tau}, p, \mathtt{q})$. By (8.9), the lemma is then proved. □

The following theorem can be easily proved by using the same argument as Theorem 8.2.

Theorem 9.3 *Assuming that one of the conditions of Theorem 2.3 is true, Assumption 9.1 is true, and the conditions of Lemmas 2.4 and 2.5 are true, we have the following norm equivalence,*

$$C_1(\|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}}^2 + \|\nabla_h v\|_0^2) \le \mathcal{G}_h(\boldsymbol{\tau}, v; 0) \le C_2(\|\boldsymbol{\tau}\|_{\boldsymbol{\Sigma}}^2 + \|\nabla_h v\|_0^2) \quad \forall (\boldsymbol{\tau}, v) \in \boldsymbol{\Sigma} \times W_D(\mathcal{T}). \tag{9.7}$$

With Assumption 9.1, we can use standard H^1-conforming finite element to approximate $\boldsymbol{\sigma}$. For simplicity, we use the linear finite element space. Define

$$\boldsymbol{\Sigma}_h = \{\boldsymbol{\tau} \in C^0(\Omega)^d : \boldsymbol{\tau}|_K \in P_1(K)^d \ \forall K \in \mathcal{T}, \boldsymbol{\tau} \in \boldsymbol{\Sigma}\}.$$

Then the potential–flux div-curl nonconforming LSFEM of (9.4) or (9.5) is: Find $(\boldsymbol{\sigma}_h, u_h) \in \boldsymbol{\Sigma}_h \times V_D^{cr}$,

$$d_h((\boldsymbol{\sigma}_h, u_h), (\boldsymbol{\tau}_h, v_h)) = (f, \nabla \cdot \boldsymbol{\tau} + X_h v) \quad \forall (\boldsymbol{\tau}_h, v_h) \in \boldsymbol{\Sigma}_h \times V_D^{cr}. \tag{9.8}$$

We then have the following a priori error estimate as the classic div-curl LSFEM problem (Theorem 3.1 of [2]):

Theorem 9.4 *Let $\alpha \in (0,1]$. Assume that $u \in H^{1+\alpha}(\Omega)$ and $\boldsymbol{\sigma} \in H^{1+\alpha}(\Omega)^d$ and let $(\boldsymbol{\sigma}_h, u_h) \in \boldsymbol{\Sigma}_h \times V_D^{cr}$ be the numerical solution of (9.8). Then*

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_1 + \|\nabla_h(u - u_h)\|_0 \leq Ch^{\alpha}(\|u\|_{1+\alpha} + \|\boldsymbol{\sigma}\|_{1+\alpha}).$$

Define the following a posteriori error estimator:

$$\xi^2 := \mathcal{G}_h(\boldsymbol{\sigma}_h, u_h; f), \tag{9.9}$$

where $(\boldsymbol{\sigma}_h, u_h) \in \boldsymbol{\Sigma}_h \times V_D^{cr}$ is the numerical solution of (9.8). Due to the norm equivalence (9.7), we immediately have the global reliability and efficiency of ξ:

$$C_1\xi \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_1 + \|\nabla_h(u - u_h)\|_0 \leq C_2\xi. \tag{9.10}$$

9.3 Restrictions of the Two-Field Potential-Flux Div-Curl Least-Squares Formulation

We discuss the reactions of the two-field potential-flux div-curl least-squares formulation due to the violation of Assumption 9.1.

The first restriction is about the domain even when A is $C^{1,1}$, for example, $A = I$. We have the following famous negative result [60] that when $d = 3$, $\{\boldsymbol{\tau} \in H(\mathrm{div};\Omega) : \boldsymbol{\tau}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega\} \cap H^1(\Omega)^3$ and $\{\boldsymbol{\tau} \in H(\mathrm{curl};\Omega) : \boldsymbol{\tau}\times\mathbf{n} = 0 \text{ on } \partial\Omega\} \cap H^1(\Omega)^3$ are closed, infinite-codimensional subspaces of $\{\boldsymbol{\tau} \in H(\mathrm{div};\Omega) : \boldsymbol{\tau}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega\} \cap H(\mathrm{curl};\Omega)$ and $\{\boldsymbol{\tau} \in H(\mathrm{curl};\Omega) : \boldsymbol{\tau}\times\mathbf{n} = 0 \text{ on } \partial\Omega\} \cap H(\mathrm{div};\Omega)$, respectively. For such cases, H^1-conforming finite elements can not be used to approximate the flux, see detailed discussion in Section B.2.2 of [4].

The second restriction is about the possible discontinuity of coefficient A. As discussed in [37, 56, 57, 61], for the exact solution, we have

$$\boldsymbol{\sigma} = -A\nabla u \in H(\mathrm{div};\Omega) \quad \text{and} \quad \nabla u \in H(\mathrm{curl};\Omega). \tag{9.11}$$

But, for a discontinuous A, we also have

$$\boldsymbol{\sigma} = -A\nabla u \notin H^1(\Omega)^d. \tag{9.12}$$

Thus for a discontinuous A, $\boldsymbol{\Sigma}_h$ or any H^1-conforming finite element space is a wrong approximation space for $\boldsymbol{\sigma}$. The error estimator (9.9) will never be zero even the numerical solution $u_h = u$ is exact; see examples given at [61].

In conclusion, the two-field potential-flux div-curl least-squares formulation should be used very carefully for both the conforming and nonconforming cases.

10 Numerical Tests

In this section, we present several numerical tests. The first one is a numerical negative example on the norm equivalence of the two-field Div CR-LSFEM. We also perform numerical tests for adaptive nonconforming two-field Div CR-LSFEM and three-field potential-flux-intensity Div-Curl LSFEM.

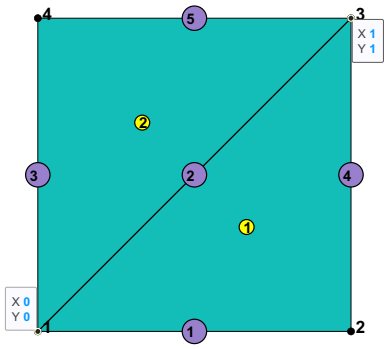


Fig. 1 An initial triangulation and degree of freedoms

10.1 A Numerical Negative Example on the Norm Equivalence of the Two-Field Div CR-LSFEM

Consider the problem

$$-\Delta u = 0 \text{ in } \Omega := (0,1)^2, \tag{10.1}$$

with the pure Dirichlet boundary condition on $\partial\Omega$. The true solution is given by $u(x,y) = 3(x^2 - y^2)$. The example is modified from an example in the appendix of [62].

The purpose of this example is to show that we have the numerical solution $(u_{cr}, \boldsymbol{\sigma}_{rt}) \in V^{cr} \times RT_0$, with u_{cr} satisfying the Dirichlet boundary condition in the interpolation sense, that the numerical solution minimizes the two-field Div CR-LS functional to zero,

$$\|\nabla_h u_{cr} + \boldsymbol{\sigma}_{rt}\|_0^2 + \|\nabla \cdot \boldsymbol{\sigma}_{rt}\|_0^2 = 0 \tag{10.2}$$

but the numerical error is not zero.

Figure 1 shows the triangulation with two elements K_1 and K_2 sharing an edge and the corresponding degrees of freedom.

Let θ_i^{cr} be the standard CR finite element basis function defined on edge F_i, $i = 1, \ldots, 5$, such that the piecewise linear function θ_i is one at the midpoint of F_i and zero at the midpoints of other edges. We assume our numerical solution u_{cr} is $u_{cr} = \sum_{i=1}^{5} \theta_i^{cr}$. On the boundary edges, the value of the CR numerical solution is chosen to be the CR interpolation of the exact solution, i.e., $a_i = \int_{F_i} u ds / |F_i|$, for $i = 1, 3, 4, 5$. We then get

$$u_{cr} = \theta_1^{cr} + a_2\theta_2^{cr} - \theta_3^{cr} + 2\theta_4^{cr} - 2\theta_5^{cr}.$$

We then can solve the rest of u_{cr} and $\boldsymbol{\sigma}_{cr}$ by minimizing the functional $\|\nabla_h u_{cr} + \boldsymbol{\sigma}_{rt}\|_0^2 + \|\nabla \cdot \boldsymbol{\sigma}_{rt}\|_0^2$. In fact, we find for our simple example, the CR finite element interpolations of the exact solution u is the numerical solution u_{cr}. We get $a_2 = 0$, and thus

$$u_{cr} = \theta_1^{cr} - \theta_3^{cr} + 2\theta_4^{cr} - 2\theta_5^{cr}.$$

It is easy to get that

$$\nabla u_{cr}|_{K_1} = (4, -2) \text{ in } K_1 \quad \text{and} \quad \nabla u_{cr}|_{K_2} = (2, -4) \text{ in } K_2.$$

Notice that $\nabla_h u_{cr} \in H(\text{div}; (0,1)^2)$ since the normal jump of $\nabla_h u_{cr}$ is zero across the edge F. In fact, $\nabla_h u_{cr} \in RT_0$. Thus, we let $\boldsymbol{\sigma}_{rt} = -\nabla_h u_{cr}$. It is obvious that $\nabla \cdot \boldsymbol{\sigma}_{rt} = 0$. Thus, we have $\|\nabla_h u_{cr} + \boldsymbol{\sigma}_{rt}\|_0^2 + \|\nabla \cdot \boldsymbol{\sigma}_{rt}\|_0^2 = 0$. That is, $(u_{cr}, \boldsymbol{\sigma}_{rt})$ is the numerical solution of the two-field Div CR-LSFEM. On the other hand, the exact error $|||(\boldsymbol{\sigma} - \boldsymbol{\sigma}_{rt}, u - u_{cr})|||$ is not zero. In fact, we have $\|\nabla_h(u - u_{cr})\|_0 = 4$, see the calculation on Appendix of [62].

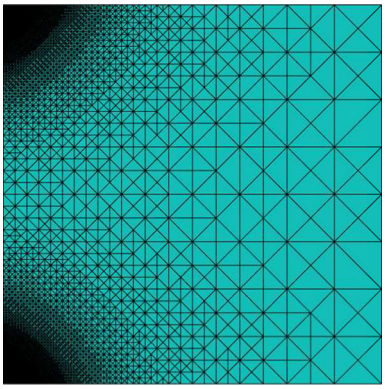


Fig. 2 Meshes generated by η_4 with 17628 elements

10.2 Numerical Tests for Adaptive Nonconforming Two-Field Div CR-LSFEM and Three-Field Potential-Flux-Intensity Div-Curl LSFEM

Next, we consider the general second-order elliptic problem $-\nabla\cdot(A\nabla u)+\mathbf{b}\cdot\nabla u+cu=f$ in $(0,1)^2$ with the pure Dirichlet boundary condition on $\partial\Omega$ and the coefficients $A=I$, $\mathbf{b}=(1,0)^t$ and $c=-10$ in $(0,1)^2$. Note that the problem is indefinite and non-symmetric. But, it has a unique solution and is well-posed. We choose f such that the true solution is

$$u(x,y)=\exp\left([(x+0.2)^2+(y+0.2)^2]^{-1/2}\right)+\exp\left([(x+0.2)^2+(y-1.2)^2]^{-1/2}\right).$$

For the two-field formulation, we use η_4 as the error estimator in the numerical test since $\eta_i, i=1,\cdots,6$ are equivalent to each other. For the three-field formulation, $\mathcal{J}_{h,1}^{dicvurl}$ is used. In the following numerical tests, we set the initial triangulation in Fig. 1 and use the Dofler marking strategy with parameter 0.2. For the three-field formulation. the true error is measure in the $\mathbb{Y}-$norm, see (8.1). The relative error for the two-field formulation is computed by rel-error $=|||(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{cr},u-u_{cr})|||/|||(\boldsymbol{\sigma},u)|||$, whereas for the three-field formulation, the relative error is rel-error $=|||(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{\phi}-\boldsymbol{\phi}_h,u-u_h)|||_{\mathbb{Y}}/|||(\boldsymbol{\sigma},\boldsymbol{\phi},u)|||_{\mathbb{Y}}$. We stop the algorithm when both tests use rel-error ≤ 0.01.

The decay of $\eta_4/|||(\boldsymbol{\sigma},u)|||$ and the relative error are shown in Fig. 3, in which there are 53348 degrees of freedom used in the two-field formulation. The final finite element meshes of η_4 are displayed in Fig. 2, with 17628 elements generated by the two-field method. In Fig. 3, the number of degrees of freedom is 82161 with 18102 elements shown in Fig. 4. Optimal convergences are observed for both cases (Fig. 5).

11 Concluding Remarks

We present two least-squares formulations for the general second-order elliptic equations with nonconforming finite element approximation. One is the two-field potential-flux div formulation, and the other is the three-field potential-flux-intensity div-curl formulation. For the two-field div formulation, the CR-LSFEM does not have the norm equivalence in the abstract setting and thus does not have automatic discrete coerciveness and built-in a posteriori error estimates. The three-field formulation recovers the norm equivalence in the abstract setting and the good properties of the least-squares method. Furthermore, examining the proof carefully, it is easy to find that the three-field potential-flux-intensity div-curl formulation

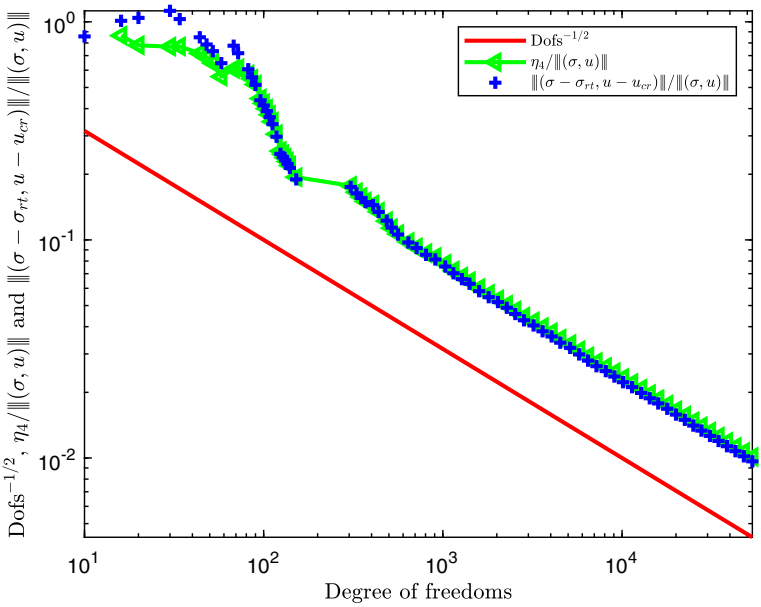


Fig. 3 Relative error versus $\eta_4/|||(\boldsymbol{\sigma}, u)|||$ with reference slope Dofs$^{-1/2}$

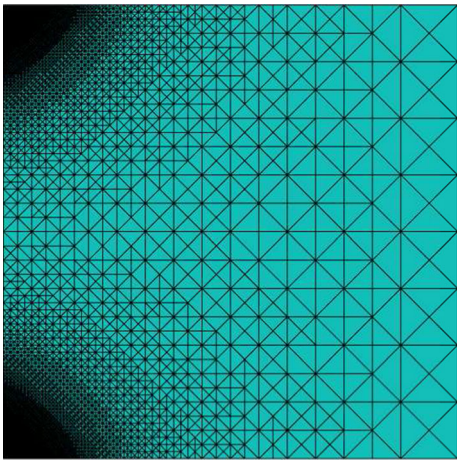


Fig. 4 Meshes generated by $\mathcal{J}_{h,1}^{divcurl}$ with 18102 elements

has the potential to be generalized to other non-traditional nonconforming least-squares approximation, for example, the multiscale finite elements [19, 20] and the immersed finite elements [21]. In this paper, we do not discuss the formulations with jump-stabilizations. In a future paper, we plan to discuss the least-squares methods with jump-stabilizations for nonconforming and discontinuous finite element approximations for the possible indefinite second-order linear elliptic equations.

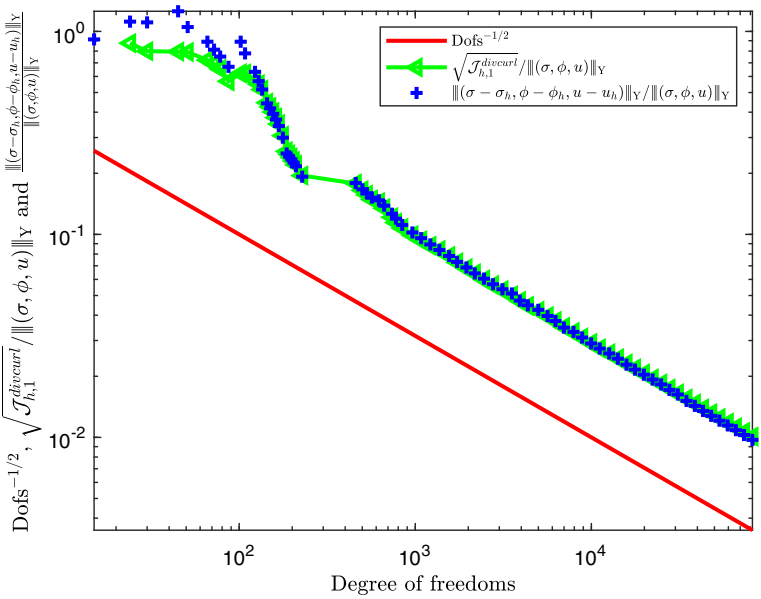


Fig. 5 Relative error versus $\sqrt{\mathcal{J}_{h,1}^{divcurl}}/|||(\boldsymbol{\sigma},\boldsymbol{\phi},u)|||_{\mathbb{Y}}$ with reference slope Dofs$^{-1/2}$

Funding This work was supported in part by Research Grants Council of the Hong Kong SAR, China, under the GRF Grant Projects No. CityU 11302519, CityU 11305319, and CityU 11316222.

Data Availability The paper has no associated data.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

1. Cai, Z., Lazarov, R., Manteuffel, T., McCormick, S.: First order system least-squares for second order partial differential equations: part I. SIAM J. Numer. Anal. **31**, 1785–1799 (1994)
2. Cai, Z., Manteuffel, T., McCormick, S.F.: First-order system least squares for second-order partial differential equations: part ii. SIAM J. Numer. Anal. **34**(2), 425–454 (1997)
3. Jiang, B.: The Least-Squares Finite Element Method Theory and Applications in Computational Fluid Dynamics and Electromagnetics. Springer, Scientifc Computation (1998)
4. Bochev, P.B., Gunzburger, M.D.: Least-Squares Finite Element Methods. Applied Mathematical Sciences, 166. Springer, (2009)
5. Cai, Z., Starke, G.: Least-squares methods for linear elasticity. SIAM J. Numer. Anal. **42**, 826–842 (2004)
6. Cai, Z., Lee, B., Wang, P.: Least-squares methods for incompressible newtonian fluid flow: linear stationary problems. SIAM J. Numer. Anal. **42**, 843–859 (2004)
7. Cai, Z., Falgout, R., Zhang, S.: Div first-order system LL* (FOSLL*) least-squares for second-order elliptic partial differential equations. SIAM J. Numer. Anal. **53**(1), 405–420 (2015)
8. Liu, Q., Zhang, S.: Adaptive least-squares finite element methods for linear transport equations based on an H(div) flux reformulation. Comput. Methods Appl. Mech. Eng. **366**, 113041 (2020)
9. Liu, Q., Zhang, S.: Adaptive flux-only least-squares finite element methods for linear transport equations. J. Sci. Comput. **84**, 26 (2020)

10. Qiu, W., Zhang, S.: Adaptive first-order system least-squares finite element methods for second order elliptic equations in non-divergence form. SIAM J. Numer. Anal. **58**(6), 3286–3308 (2020)
11. Crouzeix, M., Raviart, P.-A.: Conforming and nonconforming finite element methods for solving the stationary stokes equations i. RAIRO Anal Numér **7**, 33–75 (1973)
12. Brenner, S.: Forty years of the Crouzeix-Raviart element. Numer. Methods Partial Differ. Equ. **31**(2), 367–396 (2015)
13. Fortin, M., Soulie, M.: A non-conforming piecewise quadratic finite element on triangles. Internat. J. Numer. Methods Eng. **19**, 505–520 (1983)
14. Han, H.: A finite element approximation of Navier-Stokes equations using nonconforming elements. J. Comput. Math. **2**, 77–88 (1984)
15. Rannacher, R., Turek, S.: Simple nonconforming quadrilateral stokes element. Numer. Methods Partial Differ. Equ. **8**, 97–111 (1992)
16. Santos, J.E., Sheen, D., Ye, X.: Nonconforming Galerkin methods based on quadrilateral elements for second order elliptic problems. Math. Model. Numer. Anal. **33**, 747–770 (1999)
17. Cai, Z., Douglas, J., Ye, X.: A stable nonconforming quadrilateral finite element method for the stationary stokes and Navier-Stokes equations. Calcolo **36**, 215–232 (1999)
18. Duan, H.-Y., Liang, G.-P.: Nonconforming elements in least-squares mixed finite element methods. Math. Comp **73**(245), 1–18 (2003)
19. Hou, T.Y., Xiao-Hui, W., Cai, Z.: Convergence of multi-scale finite element method for elliptic problems with rapidly oscillating coefficients. Math. Comp **68**(227), 913–943 (1999)
20. Hesthaven, J.S., Zhang, S., Zhu, X.: High-order multiscale finite element methods for elliptic problems. Multiscale Model. Simul. **12**(2), 650–666 (2014)
21. Li, Z.: The immersed interface method using a finite element formulation. Appl. Numer. Math. **27**(3), 253–267 (1998)
22. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Different Equations, volume 23. Springer, Springer Series in Computational Mathematics, (1994)
23. Schatz, A.H.: An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. Math. Comp. **28**(128), 959–962 (1974)
24. Schatz, A.H., Wang, J.: Some new error estimates for ritz galerkin methods with minimal regularity assumptions. Math. Comp. **65**(213), 19–27 (1996)
25. Carstensen, C., Dond, A.K., Nataraj, N., Pani, A.K.: Error analysis of nonconforming and mixed fems for second-order linear non-selfadjoint and indefinite elliptic problems. Numer. Math. **133**, 557–597 (2016)
26. Carstensen, C., Nataraj, N., Pani, A.K.: Stability of mixed fems for non-selfadjoint indefinite second-order linear elliptic pdes. Numer. Math. **150**, 975–992 (2022)
27. Bramble, J.H., Lazarov, R.D., Pasciak, J.E.: A least-squares approach based on a discrete minus one inner product for first order systems. Math. Comput. **66**(219), 935–955 (1997)
28. Cai Z.: Least-squares method. Lecture notes of Purdue University, (2004)
29. JaEun, K.: A remark on the coercivity for a first-order least-squares method. Numer. Methods Partial Differ. Equ. **23**(6), 1577–1581 (2007)
30. Zhang, S.: A simple proof of coerciveness of first-order system least-squares methods for general second-order elliptic pdes. Comput. Math. Appl. **130**, 98–104 (2023)
31. Bensow, R., Larson, M.G.: Discontinuous least-squares finite element method for the div-curl problem. Numer. Math. **101**, 601–617 (2005)
32. Ye, X., Zhang, S.: A discontinuous least-squares finite-element method for second-order elliptic equations. Int. J. Comput. Math. **96**(3), 557–567 (2019)
33. Cai, Z., JaEun, K.: The L^2 norm error estimates for the div least-squares methods. SIAM J. Numer. Anal. **44**(4), 1721–1734 (2006)
34. Dari, E., Duran, R., Padra, C., Vampa, V.: A posteriori error estimators for nonconforming finite element methods. ESAIM Math. Modell. Numer. Anal. **30**(4), 385–400 (1996)
35. Carstensen, C., Bartels, S., Jansche, S.: A posteriori error estimates for nonconforming finite element methods. Numer. Math. **92**, 233–256 (2002)
36. Ainsworth, M.: Robust a posteriori error estimation for nonconforming finite element approximation. SIAM J. Numer. Anal. **42**(6), 2320–2340 (2005)
37. Cai, Z., Zhang, S.: Recovery-based error estimators for interface problems: mixed and nonconforming finite elements. SIAM J. Numer. Anal. **48**(1), 30–52 (2010)
38. Cai, Z., He, C., Zhang, S.: Residual-based a posteriori error estimate for interface problems: nonconforming linear elements. Math. Comp. **86**, 617–636 (2017)
39. Cai, Z., He, C., Zhang, S.: Discontinuous finite element methods for interface problems: Robust a priori and a posteriori error estimates. SIAM J. Numer. Anal. **55**, 400–418 (2017)

40. Cai, Z., He, C., Zhang, S.: Generalized Prager-Synge inequality and equilibrated error estimators for discontinuous elements. J. Comput. Appl. Math. **398**, 113673 (2021)
41. Cai, Z., He, C., Zhang, S.: Improved zz a posteriori error estimators for diffusion problems: discontinuous element. Appl. Numer. Math. **159**, 174–189 (2021)
42. Raviart, P.A., Thomas J.M.: A mixed finite element method for second order elliptic problems. In: I. Galligani and E. Magenes (Eds.) Mathematical Aspects of the Finite Element Method, vol 606 of Lectures Notes in Mathematics. Springer, (1977)
43. Nédélec, J.-C.: Mixed finite elements in r3. Numer. Math. **35**, 315–341 (1980)
44. Dupont, T., Scott, R.: Polynomial approximation of functions in sobolev spaces. Math. Comp. **34**, 441–463 (1980)
45. Ern, A., Guermond, J.L.: Theory and Practice of Finite Elements. Springer (2004)
46. Bers, L., John, F., Schechter, M.: Partial Differential Equations, volume 3A of Lecture Notes in Applied MAthematics. American Mathematical Society, (1964)
47. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Classics in Mathematics. Springer, (2001)
48. Evans, L.C.: Partial Differential Equations. American Mathematical Society, 2 edition, (2010)
49. Girault, V., Raviart, P.-A.: Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms. Springer, (1986)
50. Monk, P.: Finite Element Methods for Maxwell's Equations. Oxford University Press, (2006)
51. Brenner, S.: Poincaré-Friedrichs inequalities for piecewise H^1 functions. SIAM J. Numer. Anal. **41**, 306–324 (2003)
52. Vohralik, M.: On the discrete Poincaré-Friedrichs inequalities for nonconforming approximations of the sobolev space H^1. Numer. Funct. Anal. Optim. **26**(7–8), 925–952 (2005)
53. Braess, D.: Finite elements: theory, fast solvers, and applications in solid mechanics. Cambridge University Press, (2007)
54. Babuška, I.: Error bound for finite element method. Numer. Math. **16**, 322–333 (1971)
55. Jinchao, X., Zikatanov, L.: Some observations on Babuška and Brezzi theories. Numer. Math. **94**, 195–202 (2003)
56. Cai, Z., Zhang, S.: Recovery-based error estimator for interface problems: conforming linear elements. SIAM J. Numer. Anal. **47**(3), 2132–2156 (2009)
57. Cai, Z., Ye, X., Zhang, S.: Discontinuous galerkin finite element methods for interface problems: a priori and a posteriori error estimations. SIAM J. Numer. Anal. **49**(5), 1761–1787 (2011)
58. Ern, A., Guermond, J.L.: Finite Elements I: Approximation and Interpolation, volume 72 of Texts in Applied Mathematics. Springer, (2021)
59. Ern, A., Guermond, J.-L.: Finite element quasi-interpolation and best approximation. ESAIM Math. Model. Numer. Anal. **57**, 1367–1385 (2017)
60. Costabel, M.: A coercive bilinear form for maxwell's equations. J. Math. Anal. Appl. **157**(2), 527–541 (1991)
61. Cai, Z., He, C., Zhang, S.: Improved zz a posteriori error estimators for diffusion problems: conforming linear elements. Comput. Methods Appl. Mech. Eng. **313**, 433–449 (2017)
62. Kim, K.Y.: A posteriori error analysis for locally conservative mixed methods. Math. Comp. **76**(257), 43–66 (2007)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.