INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 16, Number 2, Pages 340–356

A DISCONTINUOUS RITZ METHOD FOR A CLASS OF CALCULUS OF VARIATIONS PROBLEMS

XIAOBING FENG AND STEFAN SCHNAKE

Abstract. This paper develops an analogue (or counterpart) to discontinuous Galerkin (DG) methods for approximating a general class of calculus of variations problems. The proposed method, called the discontinuous Ritz (DR) method, constructs a numerical solution by minimizing a discrete energy over DG function spaces. The discrete energy includes standard penalization terms as well as the DG finite element (DG-FE) numerical derivatives developed recently by Feng, Lewis, and Neilan in [7]. It is proved that the proposed DR method converges and that the DG-FE numerical derivatives exhibit a compactness property which is desirable and crucial for applying the proposed DR method to problems with more complex energy functionals. Numerical tests are provided on the classical p-Laplace problem to gauge the performance of the proposed DR method.

Key words. Variational problems, minimizers, discontinuous Galerkin (DG) methods, DG finite element numerical calculus, compactness, convergence.

1. Introduction

In this paper we develop a numerical method using totally discontinuous piecewise polynomial functions for approximating solutions to the following problem from the calculus of variations: Find $u \in W_g^{1,p}(\Omega)$ such that

$$\mathcal{J}(u) \leq \mathcal{J}(v) \qquad \forall v \in W_g^{1,p}(\Omega), \tag{1}$$

where

$$\mathcal{J}(v) = \int_\Omega f(\nabla v, v, x)\,\mathrm{d}x \tag{2}$$

is the energy functional, $f : \mathbb{R}^d \times \mathbb{R} \times \Omega \to \mathbb{R}_+$ is called the energy density, $\Omega \subset \mathbb{R}^d$ is an open bounded domain, and

$$W_g^{1,p}(\Omega) := \{v \in W^{1,p}(\Omega) : u = g \text{ on } \partial\Omega\}.$$

If such a u exists, it is called a minimizer of $\mathcal{J}$ over $W_g^{1,p}(\Omega)$ and is written as

$$u \in \operatorname*{arg\,min}_{v \in W_g^{1,p}(\Omega)} \mathcal{J}(v). \tag{3}$$

Although the calculus of variations is an old field in mathematics, its growth and boundary have kept expanding because new applications arising from physics, differential geometry, image processing, materials science, and optimal control (just to name a few). Those problems are often formulated as calculus of variations problems, among them are the Brachistochrone problem [5], the minimal surface problem [6], and the Erickson energy for nematic liquid crystals [11].

Numerically solving those problems means to approximate the exact minimizer u of $\mathcal{J}$ over $W_g^{1,p}(\Omega)$ via a numerical approximation u_h. As expected, there are many methods for constructing an approximate solution u_h. The existing numerical methods can be divided into two categories: the indirect approach and the direct

Received by the editors XXX.
2000 *Mathematics Subject Classification.* 65N30, 65N12, 35J60.

approach. The indirect approach is based on the fact that the minimizer u must satisfy, in some sense, the following Euler-Lagrange equation:

$$\sum_{i=1}^{d} \frac{\partial}{\partial x_i} \left(f_{\xi_i}(\nabla u, u, x) \right) = f_u(\nabla u, u, x) \qquad \forall x \in \Omega. \tag{4}$$

As equation (4) is a second order PDE in divergence (or conservative) form, it can be discretized using a variety of methods such as finite difference, finite element, discontinuous Galerkin and spectral method for constructing an approximate solution u_h. This indirect approach is often the preferred approach because of the wealthy amount of numerical methods available for discretizing PDEs. However, this approach does have two drawbacks. First, the Euler-Lagrange equation is only a necessary condition for a minimizer and it may not be a sufficient one. More information must be known about $\mathcal{J}$ in order to determine if the solution of the Euler-Lagrange equation indeed globally minimizes $\mathcal{J}$. Second, a discretization of the PDE may lose some important properties of the original energy functional, such as conservation or dissipation laws. On the other hand, the direct approach seeks an approximate solution u_h by first constructing a discrete energy functional $\mathcal{J}_h$ and then setting

$$u_h \in \arg\min_{v_h \in X_h} \mathcal{J}_h(v_h), \tag{5}$$

where X_h is a finite-dimensional space which approximates $W_g^{1,p}(\Omega)$. Since problem (5) is equivalent to a minimization problem in $\mathbb{R}^N$, a variety of algorithms (or solvers) can be employed to compute u_h. For example, we may minimize $\mathcal{J}_h$ by using a quasi-Newton algorithm or by first deriving the (discrete) Euler-Lagrange equation to $\mathcal{J}_h$ and then solving for u_h. The key issue of this approach is how to construct a "good" discrete energy functional $\mathcal{J}_h$ which can ensure the convergence of u_h to u. One important advantage of the direct approach is that a "good" discrete energy functional $\mathcal{J}_h$ will automatically preserve key properties of the original energy functional $\mathcal{J}$. For example, the discrete variational derivative method by Furihata and Matsuo for the KdV equation, nonlinear Schrödinger equations, and the Cahn-Hillard equation [9]; the Variational DGFEM method by Buffa and Ortner [2] for calculus of variations problems, and the finite element method by Nochetto *et al.* [11] for nematic liquid crystals all have such a trait.

Our goal in this paper is to develop a discontinuous Ritz (DR) framework for a class of variational problems described by (1). Our numerical method belongs to the direct approach and takes $X_h = V_h$ - the discontinuous Galerkin (DG) space consisting of totally discontinuous piecewise polynomial functions on a mesh $\mathcal{T}_h$ of Ω. We call our method a *discontinuous Ritz* method because it directly approximates problem (1). In the special case when

$$\mathcal{J}(v) = \frac{1}{2} a(v, v) - F(v),$$

and $a(\cdot, \cdot)$ is a symmetric and coercive bilinear form, problem (1) is known as the Ritz formulation of the following Galerkin (or weak) formulation: find $u \in V$ (which is assumed to be a Hilbert space) such that

$$a(u, v) = F(v) \qquad \forall v \in V.$$

As mentioned earlier, the key issue we face is to construct a "good" discrete energy functional $\mathcal{J}_h$. Since DG functions are discontinuous across element edges, two roadblocks arise when creating a discrete energy functional $\mathcal{J}_h$ that makes sense

on V_h. First, $\mathcal{J}_h$ must weakly enforce continuity and the Dirichlet boundary data. The standard way to cope with this issue in the DG framework is to use interior penalty terms, and indeed including interior penalty terms in $\mathcal{J}_h$ is sufficient to obtain these properties in the limit as $h \to 0$. Second and more importantly, these discontinuities also make DG functions not globally differentiable in general, and one has to determine how to approximate the gradient operator ∇ in the energy functional $\mathcal{J}$. An obvious choice is to approximate it by a piecewisely defined gradient operator over $\mathcal{T}_h$. However, such a naive choice of a discrete gradient may lead to divergent numerical method [2]. To overcome this difficulty, our idea is to use the newly developed discontinuous Galerkin finite element (DG-FE) numerical derivatives (and gradient) by Feng, Lewis, and Neilan in [7] as our discrete derivatives (and gradient). The bulk of this paper will devote to demonstrating the discrete energy functional so-constructed is a "good" one, in the sense that the resulting discontinuous Ritz method converges for a general class of energy functionals $\mathcal{J}$. On the other hand, no error estimate (or rates of convergence) will be provided for the general framework, such a result may only be feasible for specific problems and will be reported in a future work.

The rest of the paper is organized as follows. In Section 2, we give the notation used for the paper as well as the assumptions on the density function f in order to guarantee the well-posedness of problem (1) and the convergence of the proposed discontinuous Ritz method. In Section 3, we illustrate the need for a proper discretization of the gradient operator by showing some failed choices of discrete gradients tested on Poisson problem with homogeneous Dirichlet boundary data. In Section 4, we present the definition of the DG-FE numerical derivatives, the motivation for using it, and then define our discontinuous Ritz method. Section 5 is devoted to the convergence analysis of the proposed DR method. We prove that the proposed DR method and the variational DGFEM by Buffa and Ortner [2] are actually equivalent schemes and verify the convergence of the proposed DR method for a class of densities f. In addition, we present a compactness result using our DG-FE numerical gradient, which is of independent interests. In Section 6, we show a few numerical tests for the proposed DR method on the p-Laplace problem.

2. Preliminaries

2.1. Notation. Let Ω be a bounded polygonal domain in $\mathbb{R}^d$ $(d = 1, 2, 3)$. For $1 \le p < \infty$, let $L^p(\Omega)$ and $W^{1,p}(\Omega)$ denote the usual L^p space and Sobolev space on Ω with their standard norms. $(\cdot, \cdot)$ stands for the standard $L^2(\Omega)$ inner product. We use $p^* > 1$ to denote the Sobolev conjugate of p, that is,

$$p^* = \begin{cases} \frac{dp}{d-p} & \text{if } p < d, \\ \infty & \text{if } p \ge d, \end{cases} \tag{6}$$

and $q^* > 1$ such that

$$q^* = \begin{cases} \frac{(d-1)p}{d-p} & \text{if } p < d, \\ \infty & \text{if } p \ge d. \end{cases} \tag{7}$$

Let $\mathcal{T}_h$ be a quasi-uniform and shape regular mesh of Ω, and let $\mathcal{E}_h^I, \mathcal{E}_h^B$ be the interior and boundary edges of $\mathcal{T}_h$, and $\mathcal{E}_h = \mathcal{E}_h^I \cup \mathcal{E}_h^B$. For any $e \in \mathcal{E}_h$, let $\gamma_e > 0$, the penalty parameter, be a constant on e and denote $\gamma^* = \min_{e \in \mathcal{E}_h^I} \gamma_e$.

For any interior edge/face $e = \partial T^+ \cap \partial T^- \in \mathcal{E}_h^I$, we define the jump and average of a scalar or vector valued function v as

$$[v]\big|_e := v^+ - v^-, \qquad \{v\}\big|_e := \frac{1}{2}\left(v^+ + v^-\right),$$

where $v^\pm = v|_{T^\pm}$. On a boundary edge/face $e \in \mathcal{E}_h^B$ with $e = \partial T^+ \cap \partial\Omega$, we set $[v]\big|_e = \{v\}\big|_e = v^+$. For any $e \in \mathcal{E}_h^I$ we use ν_e to denote the unit outward normal vector pointing in the direction of the element with the smaller global index. For $e \in \mathcal{E}_h^B$ we set ν_e to be the outward normal to $\partial\Omega$ restricted to e.

We also define the broken Sobolev space

$$W^{1,p}(\mathcal{T}_h) := \prod_{T\in\mathcal{T}_h} W^{1,p}(T)$$

endowed with the following semi-norm and norm:

$$|v|_{W^{1,p}(\mathcal{T}_h)} = \|\nabla v\|_{L^p(\mathcal{T}_h)} + \Big(\sum_{e\in\mathcal{E}_h^I}\int_e \gamma_e h_e^{1-p}\big|[v]\big|^p \,\mathrm{d}S\Big)^{1/p},$$

$$\|v\|_{W^{1,p}(\mathcal{T}_h)} = |v|_{W^{1,p}(\mathcal{T}_h)} + \Big(\sum_{e\in\mathcal{E}_h^B}\int_e \gamma_e h_e^{1-p}|v-g|^p \,\mathrm{d}S\Big)^{1/p},$$

where

$$\|\nabla v\|_{L^p(\mathcal{T}_h)} := \Big(\sum_{T\in\mathcal{T}_h}\|\nabla v\|^p_{L^p(T)}\Big)^{\frac{1}{p}}.$$

We define the standard discontinuous Galerkin space V_h by

$$V_h = V_h^k = \Big\{v_h \in L^2(\Omega);\, v_h\big|_T \in \mathbb{P}_k(T) \quad \forall T \in \mathcal{T}_h\Big\},$$

where $k \geq 0$ denotes the polynomial degree. Obviously, we have $V_h \subset W^{1,p}(\mathcal{T}_h)$.

2.2. Well-posedness of calculus of variations problems. As previously mentioned, the calculus of variations is an old field in mathematics, which gives us a solid well-posedness theory for problem (1) with a general density function f. To be precise for the remaining presentation, we shall only consider the following class of density functions f:

(1) f is a Carathédory function, that is,
   (a) $x \to f(\xi, v, x)$ is measurable for every $(\xi, v) \in \mathbb{R}^d \times \mathbb{R}$, and
   (b) $(\xi, v) \to f(\xi, v, x)$ is continuous for every $x \in \Omega$.
(2) $\xi \to f(\xi, v, x)$ is convex for every $(v, x) \in \mathbb{R} \times \Omega$.
(3) For fixed $1 < p < \infty$, there exists constants $\alpha_0, \alpha_1 > 0$, $a_0, a_1 \in L^1(\Omega)$, and r and q satisfying $r < p$ and $r \leq q < p^*$ such that the following growth condition holds:

$$\alpha_0\big(|\xi|^p - |v|^r + a_0(x)\big) \leq f(\xi, v, x) \leq \alpha_1\big(|\xi|^p + |v|^q + a_1(x)\big).$$

Under above three assumptions, the direct method of calculus of variations (see [4]) shows that there exists a $u \in W_g^{1,p}(\Omega)$ satisfying (1). Moreover, if the map $(\xi, v) \to f(\xi, v, x)$ is strictly convex for a.e. $x \in \Omega$, then the minimizer u is unique. We refer the interested reader to [4] for the detailed proofs of these results.

We also note that the above structure assumptions exclude the case $p = 1$ and $p = \infty$. Since the spaces $W^{1,1}$ and $W^{1,\infty}$ are non-reflexive, these two cases are expected to be difficult to deal with and must be considered separately. On the other hand, we would like to mention that Stamm and Wihler in [14] developed a

DG method for the total variation energy, which is a special problem for the case $p = 1$. They directly discretize the TV-energy as

$$\mathcal{J}_h(v_h) = \frac{\alpha h}{2} \int_\Omega \sqrt{|\nabla_h v_h|^2 + \beta}\, dx + \frac{1}{2}\|f - v_h\|^2_{L^2(\Omega)},$$

where f is the given noisy image function and ∇_h is the DG-FE numerical derivative introduced in Section 4. As we will see later, this work is in the same spirit as ours, and the numerical tests given in [14] are quite promising.

3. The choice of discrete derivatives

Since functions in the discontinuous Galerkin space V_h are discontinuous across element edges, the energy functional $\mathcal{J}$ is not defined on V_h. To extend its domain to V_h, we define the following discrete energy functional:

$$\begin{aligned}\mathcal{J}_h^*(v_h) = &\sum_{T\in\mathcal{T}_h} \int_T f(\nabla^* v_h, v_h, x)\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h^I} \int_e \gamma_e h_e^{1-p} \big|[v_h]\big|^p \,\mathrm{d}S \\ &+ \sum_{e\in\mathcal{E}_h^B} \int_e \gamma_e h_e^{1-p} |v_h - g|^p \,\mathrm{d}S,\end{aligned} \tag{8}$$

we note that the last two terms, which are called penalty terms, are used to weakly enforce the continuity and the Dirichlet boundary data. Here ∇^* denotes an underdetermined discrete gradient defined on V_h or more generally on $W^{1,p}(\mathcal{T}_h)$.

Below we shall show that the construction of this discrete gradient ∇^* is crucial to the convergence of the numerical method, even when penalty terms are added, it must be defined judiciously to ensure the convergence. We note that the discontinuous nature of DG functions allows us to have flexibility in choosing the discrete gradient and to take into consideration of the properties such as simplicity and ease of implementation.

The simplest approach is to define ∇^* to be the piecewise gradient, that is, $(\nabla^* v)|_T = \nabla(v|_T)$ for any $T \in \mathcal{T}_h$. Obviously, such a discrete gradient is very easy and cheap to compute. This gives us the following discrete energy functional:

$$\begin{aligned}\mathcal{J}_h^{pw}(v_h) = &\sum_{T\in\mathcal{T}_h} \int_T f(\nabla^* v_h, v_h, x)\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h^I} \int_e \gamma_e h_e^{1-p} |[v_h]|^p \,\mathrm{d}S \\ &+ \sum_{e\in\mathcal{E}_h^B} \int_e \gamma_e h_e^{1-p} |v_h - g|^p \,\mathrm{d}S.\end{aligned} \tag{9}$$

However, it is mentioned in [2] that the above approach does not always give a convergent scheme. Indeed, this is true even for nice f. To see why it is so, let $p = 2$, $g = 0$, and $f(\xi, v, x) = \frac{1}{2}|\xi|^2 - F(x)v$, it is easy to check that the Euler-Lagrange equation of (1) is the following Poisson problem:

$$-\Delta u = F \quad \text{in } \Omega, \tag{10a}$$

$$u = 0 \quad \text{on } \partial\Omega. \tag{10b}$$

Requiring the Gâteaux derivative of $\mathcal{J}_h^{pw}$ to vanish at a potential minimizer $u_h \in V_h$, that is, for every $v_h \in V_h$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}_h^{pw}(u_h + tv_h)\Big|_{t=0} = 0 \qquad \forall v_h \in V_h,$$

we arrive at the following problem: find $u_h \in V_h$ such that

$$a_h^{pw}(u_h, v_h) = (F, v_h) \qquad \forall v_h \in V_h, \tag{11}$$

where

$$a_h^{pw}(u_h, v_h) := \sum_{T\in\mathcal{T}_h} \int_T \nabla u_h \cdot \nabla v_h \,\mathrm{d}x + \sum_{e\in\mathcal{E}_h^I} \int_e \frac{\gamma_e}{h_e}[u_h][v_h]\,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^B} \int_e \frac{\gamma_e}{h_e} u_h v_h \,\mathrm{d}S.$$

It is easy to verify that the bilinear form $a_h^{pw}(\cdot,\cdot)$ is coercive and continuous on V_h for any $\gamma_e > 0$, which immediately implies the existence and uniqueness of a solution u_h to problem (11). However, it is not hard to prove that scheme (11) is not consistent to the PDE problem (10), because if u is the weak solution to (10), there is a $v_h \in V_h$ such that

$$a_h^{pw}(u, v_h) \neq (F, v_h).$$

Instead we have

$$a_h^{pw}(u, v_h) = (F, v_h) + \sum_{e\in\mathcal{E}_h^I} \int_e \{\nabla u \cdot \nu_e\}[v_h]\,\mathrm{d}S \qquad \forall v_h \in V_h.$$

We emphasize that the penalty terms are not the cause for the inconsistency, since the regularity and boundary data of u forces them to vanish. It is in fact the discretization of the gradient that causes the inconsistency. The inconsistency in this example, being $\mathcal{O}(\gamma_e^{-1})$, leads to a non-convergent method. To show this, we let $d = 2$, $\Omega = (-1/2, 1/2)^2$ and choose F such that the solution $u(x, y) = (1/4 - x^2)(1/4 - y^2)$. Table 1 shows the piecewise H^1 errors and rates for varying values of γ_e. As we can see, the method is not converging to u as $h \to 0$.

TABLE 1. The piecewise H^1 errors and rates of convergence with various γ_e for the piecewise gradient discretization. Here the polynomial degree $k = 2$ is used in the test.

	$\gamma_e = 10$		$\gamma_e = 100$		$\gamma_e = 1000$	
$1/h$	H^1 Error	Rate	H^1 Error	Rate	H^1 Error	Rate
2	1.69e-02	-	1.16e-02	-	1.53e-02	-
4	1.18e-02	0.52	2.75e-03	2.08	3.59e-03	2.09
8	1.11e-02	0.09	1.31e-03	1.07	8.60e-04	2.06
16	1.11e-02	-0.00	1.31e-03	-0.00	2.21e-04	1.96
32	1.12e-02	-0.01	1.34e-03	-0.04	1.32e-04	0.75
64	1.12e-02	-0.01	1.36e-03	-0.01	1.35e-04	-0.04
128	1.12e-02	-0.00	1.36e-03	-0.00	1.38e-04	-0.03
256	1.12e-02	-0.00	1.36e-03	-0.00	1.39e-04	-0.01

We also note that the piecewise gradient discretization has the ability to produce a consistent scheme if we include additional terms to the discrete energy functional. For example, for the Poisson problem, the standard symmetric interior penalty DG

bilinear form is

$$\begin{aligned} a_h^{SIPDG}(u_h, v_h) = &\sum_{T\in\mathcal{T}_h} \int_T \nabla_h u_h \cdot \nabla_h v_h \,\mathrm{d}x \\ &- \sum_{e\in\mathcal{E}_h^I} \int_e [u_h]\{\nabla v_h \cdot \nu_e\} \,\mathrm{d}S - \sum_{e\in\mathcal{E}_h^I} \int_e [v_h]\{\nabla u_h \cdot \nu_e\} \,\mathrm{d}S \\ &+ \sum_{e\in\mathcal{E}_h^I} \int_e \frac{\gamma_e}{h_e}[u_h][v_h] \,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^B} \int_e \frac{\gamma_e}{h_e} u_h v_h \,\mathrm{d}S. \end{aligned}$$

It can be shown that $a_h^{SIPDG}(\cdot,\cdot)$, being symmetric, is induced by the following discrete energy functional (cf. [8]):

$$\begin{aligned} \mathcal{J}_h^{SIPDG}(v_h) = &\sum_{T\in\mathcal{T}_h} \frac{1}{2}\int_T |\nabla v_h|^2 \,\mathrm{d}x - \sum_{e\in\mathcal{E}_h^I} \int_e [v_h]\{\nabla v_h \cdot \nu_e\} \,\mathrm{d}S \\ &+ \sum_{e\in\mathcal{E}_h^I} \frac{1}{2}\int_e \frac{\gamma_e}{h_e}|[v_h]|^2 \,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^B} \frac{1}{2}\int_e \frac{\gamma_e}{h_e}|v_h - g|^2 \,\mathrm{d}S. \end{aligned}$$

However, this energy is specific to the Poisson problem and cannot be extended to the class of density functions f discussed in this paper.

While defining the numerical gradient as the piecewise gradient does not give a convergent method, there are examples of successful discrete gradients. In [2], Buffa and Ortner introduced a *variational DGFEM*. This method provided a consistent discretization of the gradient that produces a convergent method for a class of convex and coercive densities. Their discrete gradient is defined using the piecewise gradient with help of the following lifting operator $R : W^{1,p}(\mathcal{T}_h) \to [V_h]^d$:

$$\int_\Omega R(v)\cdot\varphi_h = -\sum_{e\in\mathcal{E}_h^I} \int_e [v]\{\varphi_h \cdot \nu_e\} \,\mathrm{d}S \quad \forall \varphi_h \in [V_h]^d. \tag{12}$$

The motivation of using this lifting operator arises from accounting for the contribution of the jumps of a discontinuous function to its distributional derivative. They then defined the following discrete energy functional:

$$\begin{aligned} \mathcal{J}_h^{BO}(v_h) = &\sum_{T\in\mathcal{T}_h} \int_T f(\nabla v_h + R(v_h), v_h, x) \,\mathrm{d}x \\ &+ \sum_{e\in\mathcal{E}_h^I} \int_e \gamma_e h_e^{1-p}|[v_h]|^p \,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^B} \int_e \gamma_e h_e^{1-p}|v_h - g|^p \,\mathrm{d}S. \end{aligned} \tag{13}$$

The bilinear form induced from this energy functional for the the Poisson problem is

$$\begin{aligned} a_h^{BO}(u_h, v_h) = &\sum_{T\in\mathcal{T}_h} \int_T \nabla u_h \cdot \nabla v_h \,\mathrm{d}x + \int_\Omega R(u_h)\cdot R(v_h) \,\mathrm{d}x \\ &- \sum_{e\in\mathcal{E}_h^I} \int_e [u_h]\{\nabla v_h \cdot \nu_e\} \,\mathrm{d}S - \sum_{e\in\mathcal{E}_h^I} \int_e [v_h]\{\nabla u_h \cdot \nu_e\} \,\mathrm{d}S \\ &+ \sum_{e\in\mathcal{E}_h^I} \int_e \frac{2\gamma_e}{h_e}[u_h][v_h] \,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^B} \int_e \frac{2\gamma_e}{h_e} u_h v_h \,\mathrm{d}S, \end{aligned}$$

which is continuous and coercive on V_h for sufficiently large $\gamma_e > 0$. Moreover, $a_h^{BO}(\cdot,\cdot)$ is consistent to the PDE problem since

$$\int_\Omega R(u)\cdot R(v_h)\,\mathrm{d}x = \sum_{e\in\mathcal{E}_h^I}\int_e [u]\{\nabla R(v_h)\cdot\nu_e\}\,\mathrm{d}S = 0 \qquad \forall v_h \in V_h,$$

which contributes to the convergence of the method for the Poisson problem.

Furthermore, it was proved in [2] that the lifting operator ensures compactness of the discrete minimizers u_h. Since the minimizer of $\mathcal{J}_h^{BO}$ is sought in V_h, which is not a subspace of $W^{1,p}(\Omega)$, the reflexive property of $W^{1,p}(\Omega)$ cannot be used to obtain a weakly convergent subsequence. However, V_h is a subset of $\mathrm{BV}(\Omega)$, the space of functions with bounded variations, which does have a compactness property in the weak$*$ topology. This compactness alone only shows that a subsequence u_{h_j} converges to a $u \in BV(\Omega)$, but Buffa and Ortner were able to prove a stronger result: if the sequence of discrete minimizers u_h is bounded in $W^{1,p}(\mathcal{T}_h)$, then a subsequence u_{h_j} converges to $u \in W^{1,p}(\Omega)$. Moreover, there holds the weak convergence

$$\nabla^* u_{h_j} + R(u_{h_j}) \rightharpoonup \nabla^* u \text{ in } L^p(\Omega) \quad \text{as } h \to 0,$$

where $\nabla^* u_{h_j}$ denotes the piecewise gradient of u_{h_j}. This compactness requires the lifting operator to be present in the discretization in order to pass the week limit and prove convergence of the method.

4. The DG-FE numerical derivatives and the discontinuous Ritz framework

4.1. The DG-FE numerical derivatives. To define the DG-FE numerical derivatives, we first introduce some notation used in [7]. Let $i = 1,\dots,d$. We define the following trace operators $\mathcal{Q}_i^+, \mathcal{Q}_i^-, \mathcal{Q}_i$ on every $e \in \mathcal{E}_h^I$:

$$\mathcal{Q}_i^\pm(v) = \{v\} \pm \frac{1}{2}\operatorname{sgn}(\nu_e^i)[v], \qquad \mathcal{Q}_i(v) = \frac{1}{2}\left(\mathcal{Q}_i^+(v) + \mathcal{Q}_i^-(v)\right), \tag{14}$$

where ν_e^i denotes the i^{th} component of the normal vector ν_e to $e \in \mathcal{E}_h$, and

$$\operatorname{sgn}(\xi) = \begin{cases} 1 & \text{if } \xi \geq 0, \\ -1 & \text{if } \xi < 0. \end{cases}$$

For $e \in \mathcal{E}_h^B$, we define $\mathcal{Q}_i^+ v = \mathcal{Q}_i^- v = \mathcal{Q}_i v = v$. Using these trace operators, three numerical partial derivative operators corresponding the left, right, and central traces of v were defined in [7] as follows.

Definition 1. Let $v \in W^{1,p}(\mathcal{T}_h)$ and $i = 1,\dots,d$. Define the numerical partial derivative operators in the x_i coordinate $\partial_{h,x_i}^+, \partial_{h,x_i}^-, \partial_{h,x_i} : W^{1,p}(\mathcal{T}_h) \to V_h$ by

$$\int_\Omega \partial_{h,x_i}^\pm(v)\varphi_h\,\mathrm{d}x = \sum_{e\in\mathcal{E}_h}\int_e \mathcal{Q}_i^\pm(v)\nu_e^i[\varphi_h]\,\mathrm{d}S - \sum_{T\in\mathcal{T}_h}\int_T v\,\partial_{x_i}\varphi_h\,\mathrm{d}x \quad \forall\varphi_h \in V_h, \tag{15}$$

$$\partial_{h,x_i}(v) = \frac{1}{2}\left(\partial_{h,x_i}^+(v) + \partial_{h,x_i}^-(v)\right). \tag{16}$$

We call $\partial_{h,x_i}(v)$ the central numerical partial derivative in the x_i coordinate. The motivation for these numerical derivatives is to require the standard integration by parts formula to hold when tested against any discrete function $\varphi_h \in V_h$. This allows many of the properties of the classical derivatives to hold for the numerical derivatives; among them are the product rule, chain rule, and integration by parts (cf. [7]). Because of this, a discrete energy built using the DG-FE derivatives

should be consistent. In addition, the discrete gradient operators $\nabla_h^+, \nabla_h^-, \nabla_h : W^{1,p}(\mathcal{T}_h) \to [V_h]^d$ were also naturally defined in [7] by

$$\nabla_h^{\pm} v = [\partial_{h,x_1}^{\pm}(v), \partial_{h,x_2}^{\pm}(v), \dots, \partial_{h,x_d}^{\pm}(v)], \tag{17}$$

$$\nabla_h v = [\partial_{h,x_1}(v), \partial_{h,x_2}(v), \dots, \partial_{h,x_d}(v)]. \tag{18}$$

We describe two convergent methods which were developed in [7] with the help of the DG-FE derivatives. Both methods were formulated for problem (10). To introduce these methods, we first define a jump operator $j_h : W^{1,p}(\mathcal{T}_h) \to V_h$ as follows:

$$\sum_{T\in\mathcal{T}_h} \int_T j_h(v)\varphi_h \,\mathrm{d}x = \sum_{e\in\mathcal{E}_h^I} \int_e \frac{\gamma_e}{h_e}[v][\varphi_h] \,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^B} \int_e \frac{\gamma_e}{h_e} v\varphi_h \,\mathrm{d}S \qquad \forall \varphi_h \in V_h.$$

The first method seeks a function $u_h \in V_h$ such that

$$\int_\Omega \nabla_h u_h \cdot \nabla_h \varphi_h \,\mathrm{d}x - \sum_{e\in\mathcal{E}_h^B} \int_e \nabla_h u_h \cdot \nu_e \varphi_h \,\mathrm{d}S + \int_\Omega j_h(u_h)\varphi_h \,\mathrm{d}x = \int_\Omega f\varphi_h \,\mathrm{d}x \tag{19}$$

for all $\varphi_h \in V_h$. This method is equivalent to the well-known local DG method for the model problem [3] and converges provided $\gamma_e > 0$.

The second method, the symmetric dual-wind discontinuous Galerkin (DWDG) method [10], is constructed from the ground up using the DG-FE gradients. The DWDG method seeks $u_h \in V_h$ such that

$$\frac{1}{2}\int_\Omega \left(\nabla_h^+ u_h \cdot \nabla_h^+ \varphi_h + \nabla_h^- u_h \cdot \nabla_h^- \varphi_h\right) \mathrm{d}x + \int_\Omega j_h(u_h)\varphi_h \,\mathrm{d}x = \int_\Omega f\varphi_h \,\mathrm{d}x \tag{20}$$

for all $\varphi_h \in V_h$. Note that the sided gradients ∇_h^+ and ∇_h^-, instead of the central gradient, are used in the formulation. If $\gamma_e > 0$, then the method was proved to be well-posed and convergent. Moreover, if $\mathcal{T}_h$ is quasi-uniform and if each element $T \in \mathcal{T}_h$ has at most one boundary edge, then the method is well-posed and converges provided $\gamma_e > -C_*$ for some h-independent constant $C_* > 0$. Thus one could set $\gamma_e \equiv 0$, that is, ignoring the penalty terms, and still achieve convergence.

We also note that besides their applications in solving PDEs, a complete DG-FE numerical calculus was developed in [7], which is of independent interests as it provides an alternative approach for computing weak (and distributional) derivatives of non-smooth functions. A Matlab Toolbox was recently developed in [13, 12] for implementation of this DG-FE numerical calculus in one and two dimensions. The toolbox provides a convenient software package for both teaching and research related to numerical derivatives.

4.2. Formulation of the discontinuous Ritz method. With the DG-FE gradients in hand, we are ready to introduce our discontinuous Ritz (DR) method.

Definition 2. The discontinuous Ritz method for problem (10) is defined by seeking $u_h \in V_h$ such that

$$u_h \in \arg\min_{v_h\in V_h} \mathcal{J}_h(v_h), \tag{21}$$

where

$$\mathcal{J}_h(v) = \int_\Omega f(\nabla_h v, v, x) \,\mathrm{d}x + \sum_{e\in\mathcal{E}_h^I} \int_e \gamma_e h_e^{1-p} |[v]|^p \,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^B} \int_e \gamma_e h_e^{1-p} |v-g|^p \,\mathrm{d}S, \tag{22}$$

where ∇_h is defined by (18).

To compute the numerical derivative $\partial_{h,x_i} v$, we note that the mass matrix induced by the left-hand side of (15) is actually a block diagonal matrix which means the computation of the derivatives can be done locally and in parallel. Moreover, when determining the DG-FE partial derivatives of a discrete function, the linearity of $\partial^{\pm}_{h,x_i}$ and ∂_{h,x_i} allows the action of taking the DG-FE partial derivatives to be written as a matrix which can be computed off-line (cf. [12]).

5. Convergence analysis of the discontinuous Ritz method

Clearly, the definition of our DR method is quite simple, we simply replace the differential gradient operator ∇ by the DG-FE (central) numerical gradient ∇_h in the the energy functional $\mathcal{J}$ to obtain our discrete energy functional $\mathcal{J}_h$. On the other hand, the convergence analysis of the proposed DR method is much less straightforward. At a glance, it is not clear why this method would work. Thus, the goal of this section is the show the convergence of the method. This will be done indirectly by showing that the proposed DR method as defined in Definition 2 is actually equivalent to the variational DGFEM developed by Buffa and Ortner in [2]. Specifically, we shall prove $\mathcal{J}_h \equiv \mathcal{J}_h^{BO}$ on V_h, thus giving equivalence of these two methods when minimizing over V_h, the equivalence allows us to borrow many technical results from [2]. We also present conditions to give the equivalence of $\|\nabla_h v_h\|$ and $|v_h|_{W^{1,p}(\mathcal{T}_h)}$ as well as a compactness result for the DG-FE derivatives.

First, we show the equivalence of $\mathcal{J}_h^{BO}$ and $\mathcal{J}_h$ on V_h.

Lemma 5.1. *Let $\mathcal{J}_h^{BO}$ and $\mathcal{J}_h$ be defined by* (13) *and* (22) *respectively, then for any $v_h \in V_h$ we have $\mathcal{J}_h(v_h) = \mathcal{J}_h^{BO}(v_h)$.*

Proof. Let $v_h \in V_h$, if we can show that $\nabla_h v_h = \nabla v_h + R(v_h)$, where ∇v_h is the piecewise gradient, then the equivalence of the two methods follows. This property was already proved in Proposition 4.2 of [7], but below we include the whole proof for completeness.

We first state the DG integration by parts formula:

$$\sum_{T\in\mathcal{T}_h}\int_T \tau\cdot\nabla v\,\mathrm{d}x = -\sum_{T\in\mathcal{T}_h}\int_T v\,\mathrm{div}\,\tau\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h^I}\int_e [\tau\cdot\nu_e]\{v\}\,\mathrm{d}S + \sum_{e\in\mathcal{E}_h}\int_e \{\tau\cdot\nu_e\}[v]\,\mathrm{d}S, \tag{23}$$

which holds for any $v \in W^{1,p}(\mathcal{T}_h)$ and $\tau \in [W^{1,p}(\mathcal{T}_h)]^d$.

For any $v \in W^{1,p}(\mathcal{T}_h)$, by the definition of $\nabla_h v_h$ and (23), we have

$$\begin{aligned}\int_\Omega \nabla_h v\cdot\varphi_h &= \sum_{e\in\mathcal{E}_h}\int_e \{v\}[\varphi_h\cdot\nu_e]\,\mathrm{d}S - \sum_{T\in\mathcal{T}_h}\int_T v\,\mathrm{div}\,\varphi_h\,\mathrm{d}x\\ &= -\sum_{e\in\mathcal{E}_h^I}\int_e [v]\{\varphi_h\cdot\nu_e\}\,\mathrm{d}S + \sum_{T\in\mathcal{T}_h}\int_T \nabla v\cdot\varphi_h\,\mathrm{d}x\\ &= \sum_{T\in\mathcal{T}_h}\int_T (\nabla v + R(v))\cdot\varphi_h\,\mathrm{d}x\\ &= \int_\Omega (\nabla v + R(v))\cdot\varphi_h\,\mathrm{d}x. \qquad \forall\varphi_h\in[V_h]^d.\end{aligned} \tag{24}$$

Thus we have

$$\int_\Omega \big(\nabla_h v_h - (\nabla v_h + R(v_h))\big) \cdot \varphi_h \,\mathrm{d}x = 0 \qquad \forall \varphi_h \in [V_h]^d,$$

by (24). Since $\nabla_h v_h, \nabla v_h, R(v_h) \in [V_h]^d$, setting $\varphi_h = \nabla_h v_h - (\nabla v_h + R(v_h))$ we obtain $\nabla_h v_h = \nabla v_h + R(v_h)$ in Ω. Thus $\mathcal{J}_h(v_h) = \mathcal{J}_h^{BO}(v_h)$. The proof is complete. □

With the equivalence we can borrow and take advantage of the convergence result from Theorem 6.1 of [2].

Theorem 5.2. *For $h > 0$, let $u_h \in V_h$ satisfy (21). Then there exists a sequence $h_j \searrow 0$ and a function $u \in W_g^{1,p}(\Omega)$ such that the following hold:*

$$\begin{aligned}
&u_{h_j} \to u \text{ in } L^q(\Omega) \quad \forall q < p^*,\\
&\nabla_{h_j} u_{h_j} \rightharpoonup \nabla u \text{ in } [L^p(\Omega)]^d,\\
&\mathcal{J}_{h_j}(u_{h_j}) \to \mathcal{J}(u),\\
&\sum_{e\in\mathcal{E}_h^B} \int_e h_e^{1-p} |u_{h_j} - g|^p \,\mathrm{d}S + \sum_{e\in\mathcal{E}_h^I} \int_e h_e^{1-p} |[u_{h_j}]|^p \,\mathrm{d}S \to 0
\end{aligned}$$

as $j \to \infty$. Moreover, any accumulation point of the set $\{u_h\}_{h>0}$ is a minimizer of $\mathcal{J}$ over $W_g^{1,p}(\Omega)$. If $\xi \to f(\xi, v, x)$ is strictly convex for all $(v, x) \in \mathbb{R} \times \Omega$, then we have

$$\|u - u_{h_j}\|_{W^{1,p}(\mathcal{T}_h)} \to 0 \qquad \text{as } j \to \infty.$$

If the minimizer u is unique, then the whole sequence $\{u_h\}_{h>0}$ converges.

The following results will be quite useful in later use of the DF-FE derivatives. First, we sate conditions to guarantee equivalence of the semi-norms $\|\nabla_h \cdot\|$ and $|\cdot|_{W^{1,p}(\Omega)}$ on V_h. To this end, we need to quote a discrete inf-sup condition from Buffa and Ortner [2].

Lemma 5.3 (Lemma A.2 of [2])**.** *Let $1 \le p < \infty$ and q be its Hölder conjugate. Then there exists a constant $C > 0$ independent of h such that*

$$\inf_{v_h\in V_h} \sup_{\varphi_h\in V_h} \frac{\int_\Omega v_h \varphi_h}{\|v_h\|_{L^p(\Omega)} \|\varphi_h\|_{L^q(\Omega)}} \ge C. \tag{26}$$

We first show that $\|\nabla_h v\|_{L^p(\mathcal{T}_h)}$ can be controlled by $|v|_{W^{1,p}(\mathcal{T}_h)}$ on $W^{1,p}(\mathcal{T}_h)$.

Lemma 5.4. *Let $1 < p < \infty$. Then there exists a constant $C > 0$ independent of h such that*

$$\|\nabla_h v\|_{L^p(\mathcal{T}_h)} \lesssim |v|_{W^{1,p}(\mathcal{T}_h)} \qquad \forall v \in W^{1,p}(\mathcal{T}_h), \tag{27}$$

Proof. Let q be the Hölder conjugate of p and let $v \in W^{1,p}(\mathcal{T}_h)$ and $\varphi_h \in [V_h]^d$. From (24) we have

$$\begin{aligned}
\int_\Omega \nabla_h v \cdot \varphi_h \,\mathrm{d}x &= -\sum_{e\in\mathcal{E}_h^I}\int_e [v]\{\varphi_h\cdot\nu_e\}\,\mathrm{d}S + \sum_{T\in\mathcal{T}_h}\int_T \nabla v\cdot\varphi_h\,\mathrm{d}x\\
&\le \sum_{e\in\mathcal{E}_h^I}\int_e h_e^{\frac{1-p}{p}}|[v]|\cdot h_e^{\frac{1}{q}}|\{\varphi_h\cdot\nu_e\}|\,\mathrm{d}S + \sum_{T\in\mathcal{T}_h}\|\nabla v\|_{L^p(T)}\|\varphi_h\|_{L^q(T)}\\
&\le \sum_{e\in\mathcal{E}_h^I}\int_e \left(h_e^{1-p}|[v]|^p\right)^{\frac{1}{p}}\left(h_e|\{\varphi_h\cdot\nu_e\}|^q\right)^{\frac{1}{q}}\,\mathrm{d}S + \|\nabla v\|_{L^p(\Omega)}\|\varphi_h\|_{L^q(\Omega)}\\
&\le \left(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\|[v]\|_{L^p(e)}^p\right)^{\frac{1}{p}}\left(\sum_{e\in\mathcal{E}_h^I} h_e\|\{\varphi_h\cdot\nu_e\}\|_{L^q(e)}^q\right)^{\frac{1}{q}}\\
&\quad + \|\nabla v\|_{L^p(\Omega)}\|\varphi_h\|_{L^q(\Omega)}\\
&\lesssim \left(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\|[v]\|_{L^p(e)}^p\right)^{\frac{1}{p}}\|\varphi_h\|_{L^q(\Omega)} + \|\nabla v\|_{L^p(\mathcal{T}_h)}\|\varphi_h\|_{L^q(\Omega)}\\
&\lesssim |v|_{W^{1,p}(\mathcal{T}_h)}\|\varphi_h\|_{L^q(\Omega)}.
\end{aligned}$$

Since $\nabla_h v \in V_h$, it follows from Lemma 5.3 that

$$\|\nabla_h v\|_{L^p(\mathcal{T}_h)} \lesssim \sup_{\varphi_h\in V_h}\frac{\int_\Omega \nabla_h v\cdot\varphi_h}{\|\varphi_h\|_{L^q(\Omega)}} \lesssim |v|_{W^{1,p}(\mathcal{T}_h)}.$$

which is exactly (27). □

We next show that $|v_h|_{W^{1,p}(\mathcal{T}_h)}$ can be controlled by $\|\nabla_h v_h\|_{L^p(\mathcal{T}_h)}$ on V_h for sufficiently large γ^*.

Lemma 5.5. *Let $1 < p < \infty$. Then there exists a constant $C, \gamma^* > 0$ independent of h such that for every $v_h \in V_h$*

$$|v_h|_{W^{1,p}(\mathcal{T}_h)} \le C\|\nabla_h v_h\|_{L^p(\mathcal{T}_h)} + C\left(\sum_{e\in\mathcal{E}_h^I}\gamma_e h_e^{1-p}\|[v_h]\|_{L^p(e)}^p\right)^{1/p}, \tag{28}$$

provided that $\gamma_e > \gamma^*$.

Proof. Let q be the Hölder conjugate of p and $v_h \in V_h$. From (24) we have

$$\int_\Omega \nabla_h v_h\cdot\varphi_h\,\mathrm{d}x = -\sum_{e\in\mathcal{E}_h^I}\int_e [v_h]\{\varphi_h\cdot\nu_e\}\,\mathrm{d}S + \int_\Omega \nabla v_h\cdot\varphi_h\,\mathrm{d}x \tag{29}$$

for every $\varphi_h \in [V_h]^d$. Let $\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})$ where $\mathcal{P}_h$ is the local L^2 projection onto $\mathcal{T}_h$ defined by

$$\int_T \mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\cdot\varphi_h\,\mathrm{d}x = \int_T \nabla v_h|\nabla v_h|^{p-2}\cdot\varphi_h\,\mathrm{d}x$$

for all $\varphi_h \in V_h$ and $T \in \mathcal{T}_h$. Choosing $\varphi_h = \mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})$ in (29) yields

$$\begin{aligned}
\int_\Omega \nabla_h v_h\cdot\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\,\mathrm{d}x &= -\sum_{e\in\mathcal{E}_h^I}\int_e [v_h]\{\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\cdot\nu_e\}\,\mathrm{d}S\\
&\quad + \int_\Omega \nabla v_h\cdot\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\,\mathrm{d}x.
\end{aligned} \tag{30}$$

By the stability of $\mathcal{P}_h$ we obtain

$$\begin{aligned}(31)\quad \int_\Omega \nabla_h v_h \cdot \mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\,\mathrm{d}x &\le \|\nabla_h v_h\|_{L^p(\mathcal{T}_h)}\|\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\|_{L^q(\mathcal{T}_h)}\\ &\le \|\nabla_h v_h\|_{L^p(\mathcal{T}_h)}\|\nabla v_h|\nabla v_h|^{p-2}\|_{L^q(\mathcal{T}_h)}\\ &\le \|\nabla_h v_h\|_{L^p(\mathcal{T}_h)}\|\nabla v_h\|^{p-1}_{L^p(\mathcal{T}_h)}.\end{aligned}$$

By the standard trace and inverse inequalities for DG functions, there exists $C_1 > 0$ independent of h such that

(32)

$$\begin{aligned}&\sum_{e\in\mathcal{E}_h^I}\int_e [v_h]\{\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\cdot\nu_e\}\,\mathrm{d}S\\ &\qquad\le \Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\|[v_h]\|^p_{L^p(e)}\Big)^{\frac{1}{p}}\Big(\sum_{e\in\mathcal{E}_h^I} h_e\|\{\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\cdot\nu_e\}\|^q_{L^q(e)}\,\mathrm{d}S\Big)^{\frac{1}{q}}\\ &\qquad\le C_1\Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\|[v_h]\|^p_{L^p(e)}\Big)^{\frac{1}{p}}\|\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\|_{L^q(\mathcal{T}_h)}\\ &\qquad\le C_1\Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\|[v_h]\|^p_{L^p(e)}\Big)^{\frac{1}{p}}\|\nabla v_h\|^{p-1}_{L^p(\mathcal{T}_h)}.\end{aligned}$$

By the properties of P_h we have

$$(33)\quad \int_\Omega \nabla v_h\cdot\mathcal{P}_h(\nabla v_h|\nabla v_h|^{p-2})\,\mathrm{d}x = \int_\Omega \nabla v_h\cdot\nabla v_h|\nabla v_h|^{p-2}\,\mathrm{d}x = \|\nabla v_h\|^p_{L^p(\mathcal{T}_h)}.$$

Thus by (30)-(33) and dividing by $\|\nabla v_h\|^{p-1}_{L^p(\mathcal{T}_h)}$ we have

$$\|\nabla_h v_h\|_{L^p(\mathcal{T}_h)} \ge -C_1\Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\|[v_h]\|^p_{L^p(e)}\Big)^{\frac{1}{p}} + \|\nabla v_h\|_{L^p(\mathcal{T}_h)}.$$

Choosing $\gamma^* = C_1^p + 1$ gives us the desired estimate. The proof is complete. □

We can also prove a compactness result using the DG-FE numerical derivatives. For this, we use a discrete compactness result from Buffa and Ortner [2].

Lemma 5.6 (Theorem 5.2 and Lemma 8 of [2])**.** *For $1 < p < \infty$ and $0 < h < 1$, let $v^h \in W^{1,p}(\mathcal{T}_h)$ such that*

$$(34)\qquad \sup_{0<h<1}\big(\|v^h\|_{L^1(\Omega)} + |v^h|_{W^{1,p}(\mathcal{T}_h)}\big) < \infty.$$

Then there exists a sequence $h_j \searrow 0$ and a function $v \in W^{1,p}(\Omega)$ such that

$$(35a)\qquad v^{h_j} \to v \quad \text{in } L^q(\Omega) \qquad \forall\, 1 \le q < p^*,$$

$$(35b)\qquad v^{h_j} \to v \quad \text{in } L^q(\partial\Omega) \qquad \forall\, 1 < q < q^*,$$

$$(35c)\qquad \nabla v^{h_j} + R(v^{h_j}) \rightharpoonup \nabla v \text{ in } [L^p(\Omega)]^d,$$

where p^* is the Sobolev conjugate of p defined in (6) *and q^* is defined in* (7)*.*

We are now ready to state our compactness result, which differs from Lemma 5.6 by controlling DG functions using the DG-FE numerical derivatives as well as showing their DG-FE numerical derivatives weakly converge.

Theorem 5.7. *Let $1 < p < \infty$. There exists $\gamma^* > 0$ such that for any $v_h \in V_h$ with*

$$\sup_{0<h<1} \left(\|v_h\|_{L^p(\partial\Omega)} + \|\nabla_h v_h\|_{L^p(\mathcal{T}_h)} + \left(\sum_{e \in \mathcal{E}_h^I} \gamma_e h_e^{1-p} \|[v_h]\|_{L^p(e)}^p \right)^{\frac{1}{p}} \right) < \infty. \tag{36}$$

Then there exists a sequence $h_j \searrow 0$ and a function $v \in W^{1,p}(\Omega)$ such that

$$v_{h_j} \to v \quad \text{in } L^q(\Omega) \qquad \forall\, 1 \le q < p^*, \tag{37a}$$

$$v_{h_j} \to v \quad \text{in } L^q(\partial\Omega) \qquad \forall\, 1 < q < q^*, \tag{37b}$$

$$\nabla_{h_j} v_{h_j} \rightharpoonup \nabla v \ \text{ in } [L^p(\Omega)]^d, \tag{37c}$$

where p^* is the Sobolev conjugate of p defined in (6) *and q^* is defined in* (7).

Proof. From Lemma 5.5, we have

$$|v_h|_{W^{1,p}(\mathcal{T}_h)} \lesssim \|\nabla_h v_h\|_{L^p(\mathcal{T}_h)} + \left(\sum_{e \in \mathcal{E}_h^I} \gamma_e h_e^{1-p} \|[v_h]\|_{L^p(e)}^p \right)^{\frac{1}{p}}.$$

which shows that v_h is uniformly bounded in $W^{1,p}(\mathcal{T}_h)$. By the Poincaré-Fredrichs inequality, Theorem 10.6.12 of [1], we have

$$\|v_h\|_{L^1(\Omega)} \lesssim \|v_h\|_{L^p(\Omega)} \lesssim \|v_h\|_{L^p(\partial\Omega)} + |v_h|_{W^{1,p}(\mathcal{T}_h)}.$$

Therefore, the family $\{v_h\}$ satisfies the hypothesis of Lemma 5.6, which gives us everything in the theorem except for (37c).

To show (37c), we use the ideas from the proof of Theorem 5.2 of [2]. Let $\varphi \in [C_c^\infty(\Omega)]^d$, if we can show

$$\lim_{j\to\infty} \int_\Omega \nabla_{h_j} v_{h_j} \cdot \varphi \,\mathrm{d}x = \int_\Omega \nabla v \cdot \varphi \,\mathrm{d}x. \tag{38}$$

then we are done. To the end, let $\varphi_{h_j} \in [V_{h_j}]^d$, from (24) we have

$$\begin{aligned}
\int_\Omega \nabla_{h_j} v_{h_j} \cdot \varphi \,\mathrm{d}x &= \int_\Omega \nabla_{h_j} v_{h_j} \cdot \varphi_{h_j} \,\mathrm{d}x + \int_\Omega \nabla_{h_j} v_{h_j} \cdot (\varphi - \varphi_{h_j}) \,\mathrm{d}x \\
&= \int_\Omega \big(\nabla v_{h_j} + R(v_{h_j})\big) \cdot \varphi_{h_j} \,\mathrm{d}x + \int_\Omega \nabla_{h_j} v_{h_j} \cdot (\varphi - \varphi_{h_j}) \,\mathrm{d}x \\
&= \int_\Omega \big(\nabla v_{h_j} + R(v_{h_j})\big) \cdot \varphi \,\mathrm{d}x \\
&\quad + \int_\Omega \big(\nabla v_{h_j} + R(v_{h_j})\big) \cdot (\varphi_{h_j} - \varphi) \,\mathrm{d}x \\
&\quad + \int_\Omega \nabla_{h_j} v_{h_j} \cdot (\varphi - \varphi_{h_j}) \,\mathrm{d}x.
\end{aligned}$$

Lemma 7 of [2] and Lemma 5.4 imply the uniform boundedness of ∇v^{h_j}, $R(v^{h_j})$, and $\nabla_{h_j} v^{h_j}$ in $L^p(\Omega)$. Thus, choosing φ_{h_j} to be the piecewise constant average of φ on $T \in \mathcal{T}_{h_j}$ forces the rightmost two terms to vanish as $j \to \infty$. We then obtain (38) from (35c). The proof is complete. □

6. Numerical experiments

In the section we present some numerical tests to show the effectiveness of the proposed discontinuous Ritz method. Our prototypical example is the following $p-$Laplace energy:

$$\mathcal{J}^p(v) = \int_\Omega \Big(\frac{1}{p} |\nabla v|^p - Fv \Big) \,\mathrm{d}x, \tag{39}$$

minimized over the space $W_g^{1,p}(\Omega)$. So the density function $f(\xi, v, x) = (1/p)|\xi|^p - F(x)v$, which satisfies all of the assumptions in the theory provided $F \in L^q(\Omega)$ for some $q > p$. Moreover, the map $(\xi, v) \to f(\xi, v, x)$ is strictly convex for a.e $x \in \Omega$. Thus there is a unique minimizer $u \in W_g^{1,p}(\Omega)$. The Euler-Lagrange equation (4) of $\mathcal{J}^p$ yields the following $p-$Laplace problem:

$$-\operatorname{div}(|\nabla u|^{p-2}\nabla u) = F \qquad \text{in } \Omega, \tag{40a}$$

$$u = g \qquad \text{on } \partial\Omega. \tag{40b}$$

Note that $p = 2$ gives the standard Poisson problem; however, here p can be any number such that $1 < p < \infty$. We will test cases in both one and two-dimensions, varying the value of p. We compute the discrete solution u_h by minimizing the discrete energy (21) with $k = 1$ and using the Matlab built-in function `fminunc` with the initial guess 0 unless otherwise specified. We also let $\gamma_e \equiv 10$ for every test unless otherwise stated.

Test 1 ($d = 1$, $p > 2$). Let $p = 2.5$, $d = 1$, $\Omega = (0, 1)$ and $g = x$. Choose $F(x) = -9\sqrt{3}x^2$ so that the exact solution is $u(x) = x^3$. Table 2 shows the errors and rates in the L^p and $W^{1,p}$-norm for $u - u_h$, where $u_h \in V_h$ is the discrete minimizer of (21). The numerical results clearly indicate that the proposed DR method is converging to the correct solution and we have optimal order convergence in the $W^{1,p}$ semi-norm, but we have sub-optimal convergence rate in the L^p norm.

TABLE 2. The L^p and $W^{1,p}(\mathcal{T}_h)$ errors and rates of convergence in h for the discontinuous Ritz method (21) applied to $\mathcal{J}^p$ from (39) where $p = 2.5$ and $\gamma_e \equiv 100$.

$1/h$	$\|u - u_h\|_{L^p(\Omega)}$	rate	$\|\nabla u - \nabla_h u_h\|_{L^p(\Omega)}$	rate	iterations
10	5.12e-03	-	1.10e-01	-	72
20	3.06e-03	0.74	5.51e-02	0.99	137
40	1.67e-03	0.88	2.76e-02	1.00	276
80	8.74e-04	0.93	1.38e-02	1.00	555
160	4.49e-04	0.96	6.92e-03	1.00	1104
320	2.28e-04	0.98	3.46e-03	1.00	2123

Test 2 ($d = 1$, $p < 2$). Let $p = 1.5$, $d = 1$, $\Omega = (0, 1)$ and $g = 0$. Choose $F(x)$ such that the exact solution is $u(x) = \sin(\pi x)$. Note that

$$w := |\nabla u|^{p-2}\nabla u = \frac{\sqrt{\pi}\cos(\pi x)}{\sqrt{|\cos(\pi x)|}}$$

is not classically differentiable since $\cos(\pi x)$ is both positive and negative on $(0, 1)$, but $w \in W^{1,q}(\Omega)$ for all $1 < q < 2$ with ∇w having a discontinuity at $x = 0.5$. Table 3 shows the L^p and $W^{1,p}$ errors and rates of convergence for the DR method. We see that the rates of convergence are suboptimal for both the L^p and $W^{1,p}$ errors. This is most likely due to the degeneracy of the PDE since largest error occurs at $x = 0.5$ where w is 0. This claim is supported by Figure 1.

Test 3 (Unknown solution case). Let $p = 8.3$, $d = 1$, $\Omega = (0, 1)$ and $g = x/2$. We choose $F(x) = \frac{2000x}{(100x^2+1)^2}$. Since we do not know the exact solution to this problem, we choose u^{FE} such that

$$u^{FE} = \arg\min_{v \in S} \mathcal{J}(v_h)$$

TABLE 3. The L^p and $W^{1,p}(\mathcal{T}_h)$ errors and rates of convergence in h for the discontinuous Ritz method (21) applied to $\mathcal{J}^p$ from (39) where $p = 1.5$

$1/h$	$\|u - u_h\|_{L^p(\Omega)}$	rate	$\|\nabla u - \nabla_h u_h\|_{L^p(\Omega)}$	rate	iterations
10	8.50e-02	-	3.19e-01	-	79
20	5.77e-02	0.56	2.06e-01	0.63	142
40	4.03e-02	0.52	1.38e-01	0.57	242
80	2.85e-02	0.50	9.56e-02	0.53	415
160	2.02e-02	0.50	6.69e-02	0.51	713
320	1.43e-02	0.50	4.72e-02	0.51	1244

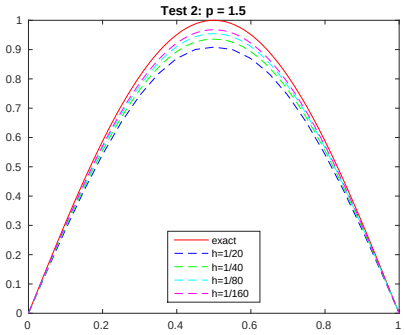

FIGURE 1. The plots of u and u_h where u is the exact minimizer for $\mathcal{J}^p(\cdot)$ from (39) with $p = 1.5$ and u_h is the discrete minimizer from (21). Here $h = 1/20, 1/40, 1/80, 1/160$.

where $S \subset W^{1,p}(\Omega)$ is the C^0 conforming Lagrange finite element space with $k = 1$ and $h = 1/640$. Table 4 shows the errors and rates in the L^p and $W^{1,p}$-norm for $u^{FE} - u_h$, where $u_h \in V_h$ is the discrete minimizer of (21). For this test we set an initial guess of $u_0 = x/2$. We see that the method is converging with a suboptimal rate of convergence in the L^p-norm.

TABLE 4. The L^p and $W^{1,p}$ errors and rates of convergence in h for the discontinuous Ritz method (21) applied to $\mathcal{J}^p$ from (39) where $p = 8.3$

$1/h$	$\|u^{FE} - u_h\|_{L^p(\Omega)}$	rate	$\|\nabla u^{FE} - \nabla_h u_h\|_{L^p(\Omega)}$	rate	iterations
10	1.95e-02	-	6.43e-01	-	95
20	9.28e-03	1.08	4.23e-01	0.61	255
40	4.46e-03	1.06	6.33e-02	2.74	1026
80	2.21e-03	1.01	2.76e-01	-2.12	1618
160	1.10e-03	1.01	2.27e-01	0.27	2763
320	5.50e-04	1.00	1.77e-01	0.35	4930

Test 4 ($d = 2$, $p > 2$)**.** Let $p = 2.5$, $d = 2$, $\Omega = (0,1)^2$. Choose F, g such that the exact solution is $u(x,y) = e^{x+y}$. For this test we choose $\gamma_e \equiv 100$. Table 5 shows the errors and rates in the L^p and $W^{1,p}$-norm for $u - u_h$, where $u_h \in V_h$ is

the discrete minimizer of (21). Again for problems with smooth solutions that lack degeneracy in the interior, the table indicates that the DR method is converging to the correct solution and we have an optimal order convergence rates in the $W^{1,p}$ semi-norm with a sub-optimal convergence rate in the L^p-norm.

TABLE 5. The L^p and $W^{1,p}$ errors and rates of convergence in h for the discontinuous Ritz method (21) applied to $\mathcal{J}^p$ from (39) where $d = 2$, $p = 2.5$, and $\gamma_e \equiv 100$.

$1/h$	$\|u - u_h\|_{L^p(\Omega)}$	rate	$\|\nabla u - \nabla_h u_h\|_{L^p(\Omega)}$	rate
4	2.01e-02	-	2.79e-01	-
8	1.04e-02	0.94	1.33e-01	1.07
16	5.32e-03	0.98	6.36e-02	1.07
32	2.68e-03	0.99	3.06e-02	1.06

References

[1] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[2] A. Buffa and C. Ortner. Compact embeddings of broken Sobolev spaces and applications. *IMA journal of numerical analysis*, 29(4):827–855, 2009.

[3] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM Journal on Numerical Analysis*, 35(6):2440–2463, 1998.

[4] B. Dacorogna. *Direct Methods in the Calculus of Variations*, volume 78. Springer Science & Business Media, 2007.

[5] B. Dacorogna. *Introduction to the Calculus of Variations*. Imperial College Press, London, third edition, 2015.

[6] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[7] X. Feng, T. Lewis, and M. Neilan. Discontinuous Galerkin finite element differential calculus and applications to numerical solutions of linear and nonlinear partial differential equations. *Journal of Computational and Applied Mathematics*, 299:68–91, 2016.

[8] X. Feng and Y. Li. Analysis of symmtric interior penalty discontinuous Galerkin methods the Allen-Cahn equation and its sharp interface limit the mean curvature flow. *IMA Journal on Numerical Analysis*, 35:1622–1651, 2015.

[9] D. Furihata and T. Matsuo. *Discrete Variational Derivative Method: A Structure-preserving Numerical Method for Partial Differential Equations*. CRC Press, 2010.

[10] T. Lewis and M. Neilan. Convergence analysis of a symmetric dual-wind discontinuous Galerkin method. *Journal of Scientific Computing*, 59(3):602–625, 2014.

[11] R. H. Nochetto, S. W. Walker, and W. Zhang. A finite element method for nematic liquid crystals with variable degree of orientation. *SIAM Journal on Numerical Analysis*, 55(3):1357–1386, 2017.

[12] S. Schnake. A Matlab toolbox for the discontinuous Galerkin finite element numerical calculus, 2014. downloadable at https://bitbucket.org/stefanschnake/dgfenumericcalculus.

[13] S. Schnake. *Numerical Methods for Non-divergence Form Second Order Elliptic Partial Differential Equations and Discontinuous Ritz Methods for Problems from the Calculus of Variations*. PhD thesis, The University of Tennessee, 2017.

[14] Stamm, B., Wihler, T.P.: A total variation discontinuous Galerkin approach for image restoration. Int. J. Numer. Anal. Model. 12(1), 81C93 (2015)

Department of Mathematics, The University of Tennessee, Knoxville, TN 37996, U.S.A.
E-mail: xfeng@math.utk.edu

Department of Mathematics, The University of Tennessee, Knoxville, TN 37996, U.S.A., Department of Mathematics, University of Oklahoma, Norman, OK 73019, U.S.A.
E-mail: sschnake@ou.edu