ESAIM: M2AN 58 (2024) 2035–2059
https://doi.org/10.1051/m2an/2024064

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

NUMERICAL ANALYSIS OF A CLASS OF PENALTY DISCONTINUOUS GALERKIN METHODS FOR NONLOCAL DIFFUSION PROBLEMS

Qiang Du[1], Lili Ju[2], Jianfang Lu[3,*] and Xiaochuan Tian[4]

Abstract. In this paper, we consider a class of discontinuous Galerkin (DG) methods for one-dimensional nonlocal diffusion (ND) problems. The nonlocal models, which are integral equations, are widely used in describing many physical phenomena with long-range interactions. The ND problem is the nonlocal analog of the classic diffusion problem, and as the interaction radius (horizon) vanishes, then the nonlocality disappears and the ND problem converges to the classic diffusion problem. Under certain conditions, the exact solution to the ND problem may exhibit discontinuities, setting it apart from the classic diffusion problem. Since the DG method shows its great advantages in resolving problems with discontinuities in computational fluid dynamics over the past several decades, it is natural to adopt the DG method to compute the ND problems. Based on [Q. Du, L. Ju, J. Lu and X. Tian, *Commun. Appl. Math. Comput.* **2** (2020) 31–55], we develop the DG methods with different penalty terms, ensuring that the proposed DG methods have local counterparts as the horizon vanishes. This indicates the proposed methods will converge to the existing DG schemes as the horizon vanishes, which is crucial for achieving *asymptotic compatibility*. Rigorous proofs are provided to demonstrate the stability, error estimates, and asymptotic compatibility of the proposed DG schemes. To observe the effect of the nonlocal diffusion, we also consider the time-dependent convection–diffusion problems with nonlocal diffusion. We conduct several numerical experiments, including accuracy tests and Burgers' equation with nonlocal diffusion, and various horizons are taken to show the good performance of the proposed algorithm and validate the theoretical findings.

Mathematics Subject Classification. 65M60, 65R20, 45A05.

Received February 19, 2024. Accepted August 5, 2024.

1. Introduction

Nonlocal modeling has become quite popular in describing some physical phenomena involving nonlocal interactions of finite range in recent years. Unlike the classic local partial differential equation models, the nonlocal models can describe the physical phenomena in a setting with reduced regularity requirements and allow the singularities and discontinuities to occur naturally. Additionally, due to the finite interaction range,

Keywords and phrases. Nonlocal diffusion, asymptotic compatibility, discontinuous Galerkin, interior penalty.

[1] Department of Applied Physics and Applied Mathematics and Data Science Institute, Columbia University, New York, NY 10027, USA.

[2] Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA.

[3] School of Mathematics, South China University of Technology, Guangzhou, Canton 510641, P.R. China.

[4] Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA.

*Corresponding author: jflu@scut.edu.cn

the nonlocal models would be more computationally efficient compared to some integro-differential equations characterized by an infinite range of interactions. The benefits of nonlocal models have led to their widespread use in various fields, such as crack and fracture in solid mechanics [24, 49, 50], traffic flows [41], nonlocal wave equation [31,39], nonlocal convection–diffusion problems [29,56], phase transitions [25,33], and image processing [37, 38]. Nonlocal volume-constrained diffusion models have a strong connection with the fractional Laplacian and fractional derivatives [18,27]. The boundary conditions of the nonlocal diffusion (ND) problems are defined on a nonzero volume region outside the domain, serving as a natural extension of those in differential equation problems. Since most singular phenomena consist of both smooth and nonsmooth regions, a natural approach is to apply nonlocal models in the nonsmooth regions and local models in the smooth regions. Consequently, if local models are used near the boundary, the boundary conditions would remain the classic ones. This leads to the seamless coupling of the local and nonlocal models, see [20,32] and the references cited therein.

Currently, there are several research topics on nonlocal models, including nonlocal vector calculus [19,28,40], nonlocal trace spaces [36, 55], nonlocal modeling and its mathematical investigations and numerical simulations, to name a few, [21, 24–26, 29–31, 33, 34, 41, 44, 52, 56, 61]. For a comprehensive literature review, readers can refer to [18, 23, 27] and the references therein. Among the various numerical methods for nonlocal problems, the discontinuous Galerkin (DG) finite element method is a natural choice due to its effectiveness in handling singularities and discontinuities. In 1973, Reed and Hill introduced the first DG method to solve the steady transport equation [45]. Around 1990, Cockburn and Shu *et al.* combined the third-order total variation diminishing Runge–Kutta method [47, 48] in temporal discretization with DG method in spatial discretization and successfully solved hyperbolic conservation laws [10–12, 14, 15]. Since then, the DG method has gained significant attention and become widely used in many fields, such as aeroacoustics, oceanography, meteorology, electromagnetism, granular flows, turbulent flows, viscoelastic flows, magneto-hydrodynamics, oil recovery simulation, semiconductor device simulation, transport of contaminants in porous media and weather forecasting, etc. Meanwhile, the study on the DG method for diffusion problems was developed independently in the 1970s and there exist various DG methods for diffusion problems. To name a few, there are symmetric interior penalty Galerkin method [22], nonsymmetric interior penalty Galerkin method [46], Baumann-Oden's method [3], Babuška–Zlámal's method [2], local DG method [13], ultra-weak DG method [8], recovery DG method [58], direct DG method [42] and hybridizable DG method [16], sparse grid DG method [60], weak Galerkin method [59], embedded DG method [17], etc. In particular, in [1] Arnold *et al.* provided a general framework to analyze the DG methods with interior penalty and revealed the key aspects of constructing these methods.

Even though many DG methods have been constructed and proposed for classic diffusion problems previously, these DG methods cannot be trivially extended to ND problems. The major difficulty lies in the absence of differential operators in the nonlocal diffusion problems, which means integration by parts cannot be used, and jumps do not appear in the weak formulations. In [9], Chen and Gunzburger proposed a discontinuous finite element method for peridynamic models in the continuous finite element framework using the piecewise polynomial finite element space. They observed that the convergence of the numerical solutions depends on the choice of the horizon and mesh size when using the piecewise constant discretizations. In [54], Tian and Du constructed a nonconforming DG method for nonlocal variational problems. However, none of the previously mentioned DG methods are *asymptotically compatible* (see, *e.g.*, [53, 57]). In [34], the authors developed a DG method for ND problems, which is a nonlocal analog of the local DG method [13] and achieved asymptotic compatibility. Later in [35], the authors proposed a penalty DG method, which is a nonlocal analog of Babuška–Zlámal's DG method [2]. In this paper, we extend this work by constructing a more general DG method for ND problems. We provide a general framework for the penalty DG methods applied to ND problems and artificially construct the penalty terms involving jumps. With the artificial penalty terms, we are able to recover several aforementioned well-known DG methods. Theoretical results on boundedness, stability, and *a priori* error estimates are also provided. To observe the nonlocal diffusion effect, we consider a convection–diffusion problem with nonlocal diffusion. With the standard DG discretization for the convective term and the proposed methods for the ND term, we obtain the L^2-stability in the semi-discrete case. Several numerical examples, including

accuracy tests, nonsmooth ND problems, and Burgers' equations with nonlocal diffusion, are presented to validate the good performance of the proposed algorithm.

The rest of the paper is organized as follows. In Section 2, we first introduce the nonlocal diffusion problem and its variational form, then propose the discontinuous Galerkin methods with some artificial penalty terms. In Section 3, we study the boundedness, stability, and *a priori* error estimates for the proposed schemes. The convection–diffusion model with nonlocal diffusion is considered and a semi-discrete analysis is given. In Section 4, we show some numerical examples, including the smooth and nonsmooth problems, and the time-dependent convection–diffusion problem with nonlocal diffusion. Concluding remarks are given in Section 5.

2. Method formulation

In this section, we first introduce the one-dimensional ND problem and its variational form. Next, we construct the penalty DG methods for the ND problem with artificial penalty terms. The penalties on the jumps not only ensure that the integral is well-defined but can also be adjusted to create nonlocal analogs of several existing DG methods for classic diffusion problems.

2.1. Problem description

Consider the one-dimensional steady-state ND problem with the nonlocal volume constraint in the following

$$\begin{cases} \mathcal{L}_\delta u = f_\delta, & x \in \Omega \triangleq (a,b), \\ u = 0, & x \in \Omega_\delta \triangleq [a-\delta, a] \cup [b, b+\delta], \end{cases} \tag{2.1}$$

where $\delta > 0$ is a constant. The operator $\mathcal{L}_\delta$ is defined as

$$\mathcal{L}_\delta u(x) := -2 \int_{x-\delta}^{x+\delta} (u(y) - u(x)) \widehat{\gamma}_\delta(x,y)\, \mathrm{d}y.$$

The kernel function $\widehat{\gamma}_\delta(x,y)$ is nonnegative and symmetric. For simplicity, in the paper, we consider a special case that

$$\begin{cases} \widehat{\gamma}_\delta(x,y) = \gamma_\delta(s) = \gamma_\delta(-s), & s = x - y, \\ s^2 \gamma_\delta(s) \in L^1_{loc}(\mathbb{R}). \end{cases} \tag{2.2}$$

The natural energy space associated with (2.1) is

$$\mathcal{S} = \left\{ v \in L^2\left(\widetilde{\Omega}\right) : \|v\|_{\mathcal{S}} < \infty, \quad v = 0 \text{ on } \Omega_\delta \right\},$$

where $\widetilde{\Omega} = \Omega \cup \Omega_\delta$ and the semi-norm $\|v\|_{\mathcal{S}}$ is defined as

$$\|v\|_{\mathcal{S}}^2 = 2 \int_0^\delta \gamma_\delta(s) \int_{\widetilde{\Omega}} \left(E_s^+ v(x)\right)^2 \mathrm{d}x\, \mathrm{d}s,$$

where $E_s^+ w(x) = w(x+s) - w(x)$. In fact, the semi-norm $\|\cdot\|_{\mathcal{S}}$ is a norm on $\mathcal{S}$ (see *e.g.* [54]). Then the variation form of (2.1) is as follows:

$$\text{Find } u \in \mathcal{S} \text{ such that } B(u,v) = (f,v), \quad \forall\, v \in \mathcal{S}, \tag{2.3}$$

where the bilinear form is given as

$$B(u,v) = 2 \int_0^\delta \gamma_\delta(s) \int_{a-\delta}^{b+\delta} E_s^+ u(x)\, E_s^+ v(x)\, \mathrm{d}x\, \mathrm{d}s, \tag{2.4}$$

and $(\cdot,\cdot)$ is the usual L^2 product on $\widetilde{\Omega}$. Note that the integral in $B(u,v)$ requires the values of u outside $\widetilde{\Omega}$, therefore we take the zero extension of u such that $u=0$ on $\widetilde{\Omega}^c$.

Since $s^2\gamma_\delta(s)\in L^1_{loc}(\mathbb{R})$, without loss of generality we assume that

$$\int_{-\delta}^{\delta} s^2\gamma_\delta(s)\,\mathrm{d}s = 1.$$

When $\delta\to 0$, the nonlocal diffusion problem (2.1) becomes the heat equation with Dirichlet boundary condition as follows:

$$\begin{cases} -u_{xx} = f, & x\in\Omega,\\ u(a)=u(b)=0. \end{cases} \tag{2.5}$$

We refer the readers to [18, 23, 27] for more details.

2.2. Penalty discontinuous Galerkin methods

To construct the penalty DG method, we first take the partition of the domain $\widetilde{\Omega}$ as $\mathcal{T}_h = \left\{I_j = \left(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right)\right\}_{j=-m+1}^{N+m}$, with

$$x_{\frac{1}{2}} = a,\ x_{N+\frac{1}{2}} = b,\ x_{-m-\frac{1}{2}} \le a-\delta < x_{-m+\frac{1}{2}},\ x_{N+m-\frac{1}{2}} < b+\delta \le x_{N+m+\frac{1}{2}}. \tag{2.6}$$

Assume the partition $\mathcal{T}_h$ is regular, *i.e.*, there exists a constant $\nu>0$ such that

$$\nu h \le \rho. \tag{2.7}$$

where h,ρ are given as

$$h = \max_j h_j,\ \rho = \min_j h_j,\ h_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}. \tag{2.8}$$

Then we define the finite element space as

$$V_h = V_h^k = \left\{ v\in L^2\left(\widetilde{\Omega}\right) :\ \ v|_{I_j}\in\mathcal{P}_k(I_j),\ j=1,\cdots,N,\ v|_{\Omega_\delta}=0 \right\}, \tag{2.9}$$

where $\mathcal{P}_k(I_j)$ is the space of polynomials on I_j whose degrees are at most k. Following [35], we divide $B(u,v)$ in (2.4) into three parts as follows:

$$B(u,v) = B_1(u,v) + B_2(u,v) + B_3(u,v),\quad \forall\, u,v\in\mathcal{S}, \tag{2.10}$$

where the above three terms are given as

$$\begin{aligned} B_1(u,v) &= 2\int_0^{\hat{h}} \gamma_\delta(s)\sum_j \int_{I^s_{j,1}} E_s^+u(x)E_s^+v(x)\,\mathrm{d}x\,\mathrm{d}s,\\ B_2(u,v) &= 2\int_0^{\hat{h}} \gamma_\delta(s)\sum_j \int_{I^s_{j,2}} E_s^+u(x)E_s^+v(x)\,\mathrm{d}x\,\mathrm{d}s,\\ B_3(u,v) &= 2\int_{\hat{h}}^{\delta} \gamma_\delta(s)\sum_j \int_{I_j} E_s^+u(x)E_s^+v(x)\,\mathrm{d}x\,\mathrm{d}s, \end{aligned} \tag{2.11}$$

where $I^s_{j,1}$ and $I^s_{j,2}$ are given as

$$I^s_{j,1} = \left(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}-s\right),\quad I^s_{j,2} = \left(x_{j+\frac{1}{2}}-s, x_{j+\frac{1}{2}}\right),\quad \hat{h} = \min\{\rho,\delta\}.$$

In [35], we showed $B(u,v)$ may not be well-defined in the discrete space V_h since $B_2(u,v)$ would cause troubles. Since $I^s_{j,2} = \left(x_{j+\frac{1}{2}} - s, x_{j+\frac{1}{2}}\right)$, then $u(x)$ and $u(x+s)$ are in the different elements, and this may lead to $\gamma_\delta(s)(u(x+s)-u(x))(v(x+s)-v(x))$ not integrable on $(0,\hat{h})$ when $u,v \in V_h$. We now take the similar treatment in [35] and put the boundary terms together, then we can obtain the bilinear form at the discrete level in the following.

$$B_h(u_h,v_h) = E(u_h,v_h) + J(u_h,v_h) + \mu P(u_h,v_h), \tag{2.12}$$

where $\mu > 0$ is taken to be large enough to ensure the stability and $E(u_h,v_h)$, $P(u_h,v_h)$ are defined as

$$\begin{aligned}
E(u_h,v_h) &= 2\sum_j \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,1}} E_s^+ u_h(x) E_s^+ v_h(x)\,\mathrm{d}x\,\mathrm{d}s \\
&\quad + 2\sum_j \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} \left(E_s^+ u_h(x) - [\![u_h]\!]_{j+\frac{1}{2}}\right)\left(E_s^+ v_h(x) - [\![v_h]\!]_{j+\frac{1}{2}}\right)\mathrm{d}x\,\mathrm{d}s \\
&\quad + 2\sum_j \int_{\hat{h}}^{\delta} \gamma_\delta(s) \int_{I_j} E_s^+ u_h(x) E_s^+ v_h(x)\,\mathrm{d}x\,\mathrm{d}s, \\
P(u_h,v_h) &= \int_0^{\hat{h}} s^2 \gamma_\delta(s)\,\mathrm{d}s \sum_j [\![u_h]\!]_{j+\frac{1}{2}} [\![v_h]\!]_{j+\frac{1}{2}},
\end{aligned}$$

where $[\![w]\!]_{j+\frac{1}{2}} = w\left(x^+_{j+\frac{1}{2}}\right) - w\left(x^-_{j+\frac{1}{2}}\right)$. If we define $g_v(x,s)$ in the following:

$$g_v(x,s) = \begin{cases} E_s^+ v(x) - [\![v]\!]_{j+\frac{1}{2}}, & \text{when } x \in I^s_{j,2},\ s \in (0,\hat{h}), \\ E_s^+ v(x), & \text{elsewhere.} \end{cases} \tag{2.13}$$

Then we have

$$E(u_h,v_h) = 2\int_0^{\delta} \gamma_\delta(s) \sum_j \int_{I_j} g_{u_h}(x,s) g_{v_h}(x,s)\,\mathrm{d}x\,\mathrm{d}s.$$

$J(u_h,v_h)$ consists of boundary terms obtained by an integration by parts in the nonlocal sense and different choices of $J(u_h,v_h)$ and μ lead to various DG schemes. For instance, we show three kinds of DG formulations in the following by choosing different $J(u_h,v_h)$ and μ, and the local limits of these DG methods have been known for many years.

– Nonlocal version of Babuška–Zlámal method [2] (nBZ):

$$\begin{aligned} J(u_h,v_h) &= 0, \\ \mu &= O\left(h^{-2k-1}\right). \end{aligned} \tag{2.14}$$

  Here k is the degree of the polynomials in V_h^k.

– Nonlocal version of IP method [22] (nIP):

$$\begin{aligned}
J(u_h,v_h) &= 2\sum_j [\![v_h]\!]_{j+\frac{1}{2}} \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} g_{u_h}(x,s)\,\mathrm{d}x\,\mathrm{d}s \\
&\quad + 2\sum_j [\![u_h]\!]_{j+\frac{1}{2}} \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} g_{v_h}(x,s)\,\mathrm{d}x\,\mathrm{d}s, \\
\mu &= O\left(h^{-1}\right).
\end{aligned} \tag{2.15}$$

– Nonlocal version of NIPG method [46] (nNIPG):

$$\begin{aligned} J(u_h, v_h) = & 2\sum_j [\![v_h]\!]_{j+\frac{1}{2}} \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} g_{u_h}(x,s)\,\mathrm{d}x\,\mathrm{d}s \\ & - 2\sum_j [\![u_h]\!]_{j+\frac{1}{2}} \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} g_{v_h}(x,s)\,\mathrm{d}x\,\mathrm{d}s, \\ \mu = & O\left(h^{-1}\right). \end{aligned} \tag{2.16}$$

In particular, the nBZ method has already been analyzed in [35]. In the rest of the paper, we focus on the two DG schemes, nIP and nNIPG.

Remark 2.1. As previously mentioned, several DG methods exist for diffusion problems in the literature. Although only three DG methods are listed above, this approach can be extended to several penalty DG methods in [1] without substantial difficulty. It is worth noting that the penalty term $P(u_h, v_h)$ may require suitable modifications of the lifting operator. We refer readers to [1] for more details.

Now we present the penalty DG formulation for the ND problem (2.1) as follows:

$$\text{Find } u_h \in V_h \text{ such that} \quad B_h(u_h, v_h) = (f_\delta, v_h), \quad \forall\, v_h \in V_h, \tag{2.17}$$

where $B_h(u_h, v_h)$ is given in (2.12), and $J(u_h, v_h)$ and μ can be taken either one from (2.14), (2.15) or (2.16).

3. Boundedness, stability and *a priori* error estimates

In this section, we consider the boundedness and stability of the DG methods. The nIP ((2.17) and (2.15)) and nNIPG schemes ((2.17) and (2.16)) are consistent, while the nBZ scheme ((2.17) and (2.14)) is not. To control the inconsistency error in the nBZ method, the so-called superpenalty technique was applied to estimate the inconsistent term, see, *e.g.*, [1, 35]. In this section, we focus on nIP and nNIPG methods, deriving the boundedness, stability, and *a priori* error estimates for these two DG methods. Throughout this section, we let $C > 0$ represent a generic constant independent of h and δ but with possibly different values. Now let us define the semi-norms for $v \in V(h) \triangleq V_h + \mathcal{S}$ as follows:

$$\begin{aligned} |v|^2_{E,h} &= 2\sum_j \int_0^{\delta} \gamma_\delta(s) \int_{I_j} g_v(x,s)^2\,\mathrm{d}x\,\mathrm{d}s, \\ |v|^2_{J,h} &= 2\sum_j h_j \int_0^{\hat{h}} \gamma_\delta(s)\,\frac{1}{s} \int_{I^s_{j,2}} g_v(x,s)^2\,\mathrm{d}x\,\mathrm{d}s, \\ |v|^2_{P,h} &= \int_0^{\hat{h}} s^2 \gamma_\delta(s)\,\mathrm{d}s \sum_j [\![v]\!]^2_{j+\frac{1}{2}}. \end{aligned} \tag{3.1}$$

To consider the boundedness and stability for the bilinear form $B_h(\cdot,\cdot)$, we define the norm for $v \in V(h)$ as follows:

$$|||v|||^2 = |v|^2_{E,h} + |v|^2_{J,h} + \mu\,|v|^2_{P,h}. \tag{3.2}$$

To see that $|||\cdot|||$ is a norm on V_h, we have the following proposition in [35].

Proposition 3.1. *For the general kernels γ_δ satisfying (2.2), it holds that for some constant $C > 0$ independent of δ and h such that*

$$\|v_h\|_{L^2} \le C\,|||v_h|||, \quad \forall\, v_h \in V_h. \tag{3.3}$$

3.1. Boundedness

The boundedness is straightforward after we define the norm $|||\cdot|||$ on $V(h)$. For instance, we consider the term $J(u_h, v_h)$ in (2.15). In fact, $\forall\, v, w \in V_h + \mathcal{S}$, by Cauchy–Schwarz inequality we have

$$\begin{aligned}
&\sum_j [\![v]\!]_{j+\frac{1}{2}} \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} g_w(x,s)\,\mathrm{d}x\,\mathrm{d}s \\
&\le \left(\int_0^{\hat{h}} s^2 \gamma_\delta(s)\,\mathrm{d}s \sum_j \frac{1}{h_j} [\![v]\!]^2_{j+\frac{1}{2}} \right)^{\frac{1}{2}} \left(\sum_j h_j \int_0^{\hat{h}} \gamma_\delta(s) \frac{1}{s} \int_{I^s_{j,2}} g_w(x,s)^2\,\mathrm{d}x\,\mathrm{d}s \right)^{\frac{1}{2}} \\
&\le (2\rho)^{-\frac{1}{2}} |v|_{P,h}\, |w|_{J,h},
\end{aligned} \tag{3.4}$$

where ρ is defined in (2.8). Similarly, we have

$$\sum_j [\![w]\!]_{j+\frac{1}{2}} \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} g_v(x,s)\,\mathrm{d}x\,\mathrm{d}s \le (2\rho)^{-\frac{1}{2}}\, |w|_{P,h}\, |v|_{J,h}\,. \tag{3.5}$$

From (3.4), (3.5) and (2.15), we obtain

$$J(v,w) \le \sqrt{2}\rho^{-\frac{1}{2}} \left(|v|_{P,h}\, |w|_{J,h} + |w|_{P,h}\, |v|_{J,h} \right). \tag{3.6}$$

Therefore, for any $\mu \ge 1/\rho$, by Cauchy–Schwarz inequality we have

$$\begin{aligned}
B_h(v,w) &\le |v|_{E,h}\, |w|_{E,h} + \sqrt{2}\rho^{-\frac{1}{2}} \left(|w|_{P,h}\, |v|_{J,h} + |v|_{P,h}\, |w|_{J,h} \right) + \mu\, |v|_{P,h}\, |w|_{P,h} \\
&\le 2\, |||v|||\, |||w|||, \quad \forall\, v, w \in V(h).
\end{aligned} \tag{3.7}$$

3.2. Stability

We now show the DG methods (2.17) are stable, *i.e.* $\exists\, C_s > 0$ independent of h and δ such that

$$B_h(v_h, v_h) \ge C_s |||v_h|||^2, \quad \forall\, v_h \in V_h. \tag{3.8}$$

From the definition of the bilinear form $B_h(\cdot,\cdot)$ in (2.12), for any $v_h \in V_h$ we have

$$B_h(v_h, v_h) = |v_h|^2_{E,h} + J(v_h, v_h) + \mu\, |v_h|^2_{P,h}\,.$$

therefore, we can obtain the stability once $J(v_h, v_h)$ is controlled. A Lemma is presented below, which is crucial in deriving the stability of the nIP method.

Lemma 3.1. *For any $v_h \in V_h$, there exists a constant $C_0 > 0$ independent of h and δ, such that*

$$|v_h|^2_{J,h} \le C_0\, |v_h|^2_{E,h}\,. \tag{3.9}$$

The proof of Lemma 3.1 is given in Appendix A.

If (3.9) holds, upon using $J(v_h, v_h) \le 2\,\rho^{-\frac{1}{2}}\, |v_h|_{P,h}\, |v_h|_{J,h}$ from (3.6) and Cauchy–Schwarz inequality we then have

$$\begin{aligned}
B_h(v_h, v_h) &\ge |v_h|^2_{E,h} - 2\,\rho^{-\frac{1}{2}}\, |v_h|_{P,h}\, |v_h|_{J,h} + \mu\, |v_h|^2_{P,h} \\
&\ge |v_h|^2_{E,h} - 2C_0\rho^{-1} |v_h|^2_{P,h} - (2C_0)^{-1} |v_h|^2_{J,h} + \mu\, |v_h|^2_{P,h} \\
&\ge \frac{1}{2} |v_h|^2_{E,h} + \left(\mu - 2C_0\rho^{-1}\right) |v_h|^2_{P,h} \\
&\ge (2C_0+2)^{-1} \left(|v_h|^2_{E,h} + |v_h|^2_{J,h} + \mu\, |v_h|^2_{P,h} \right) \\
&\quad + \left(\mu - 2C_0\rho^{-1} - \mu\,(2C_0+2)^{-1} \right) |v_h|^2_{P,h}\,.
\end{aligned} \tag{3.10}$$

Therefore, if we take μ sufficiently large, for instance $\mu \geq (1-(2C_0+2)^{-1})^{-1}2C_0\rho^{-1}$, such that

$$\mu - 2C_0\rho^{-1} - \mu(2C_0+2)^{-1} \geq 0,$$

then from (3.10) we immediately obtain desired result (3.8) with $C_s = (2C_0+2)^{-1}$.

3.3. An *a priori* error estimate

In the previous section, we have obtained the boundedness and stability of the DG methods (2.17), we now consider the consistency and approximation error. In the error estimates, we first make the assumption that the exact solution u is smooth, which means u does not have discontinuities inside the domain $\widetilde{\Omega}$ so that the continuous interpolation can be defined. We take the continuous interpolant $u_I \in V_h$ of the exact solution u so that $u - u_I$ will be zero at the element interfaces, then we have the following approximation properties [35]:

$$\|u-u_I\|_{L^2} \leq Ch^{k+1}\ |u|_{H^{k+1}} \quad \text{and} \quad |||u-u_I||| \leq Ch^k\ |u|_{H^{k+1}}\,, \tag{3.11}$$

where $C>0$ is independent of h and δ, and $|u|_{H^{k+1}} = \left(\int_{\widetilde{\Omega}} (\partial_x^{k+1}u)^2 \mathrm{d}x\right)^{1/2}$ is the semi-norm. We refer the readers to [5] for more details.

In [35], we know that with enough smoothness of the exact solution u, the bilinear form $B(u,v_h)$ is well-defined and $B(u,v_h) = (f_\delta, v_h)$. For the nIP method and nNIPG method, if they are consistent, *i.e.* $B_h(u,v_h) = B(u,v_h)$, we then have

$$B_h\left(u,v_h\right) = B\left(u,v_h\right) = \left(f_\delta, v_h\right) = B_h\left(u_h,v_h\right), \quad \forall\, v_h \in V_h.$$

In fact, we can check the consistency with some calculations that

$$\begin{aligned}
E\left(u,v_h\right) + J\left(u,v_h\right) = {} & 2\sum_j \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,1}} E_s^+ u(x) E_s^+ v_h(x)\,\mathrm{d}x\,\mathrm{d}s \\
& + 2\sum_j \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} E_s^+ u(x) \left(E_s^+ v_h(x) - [\![v_h]\!]_{j+\frac{1}{2}}\right) \mathrm{d}x\,\mathrm{d}s \\
& + 2\sum_j \int_{\hat{h}}^{\delta} \gamma_\delta(s) \int_{I_j} E_s^+ u(x) E_s^+ v_h(x)\,\mathrm{d}x\,\mathrm{d}s \\
& + 2\sum_j [\![v_h]\!]_{j+\frac{1}{2}} \int_0^{\hat{h}} \gamma_\delta(s) \int_{I^s_{j,2}} E_s^+ u(x)\,\mathrm{d}x\,\mathrm{d}s \\
= {} & B_1\left(u,v_h\right) + B_2\left(u,v_h\right) + B_3\left(u,v_h\right) = B\left(u,v_h\right).
\end{aligned}$$

Note that we use the smoothness of u that $[\![u]\!]_{j+\frac{1}{2}} = 0$, $\forall\, j$, which also leads to $P(u,v_h) = 0$, thus we have $B_h(u,v_h) = E(u,v_h) + J(u,v_h) + \mu P(u,v_h) = B(u,v_h)$. Therefore, the nIP method is consistent, as well as the conventional IP method, and it also holds for the nNIPG method.

So far, we have obtained the boundedness, stability, approximation properties and consistency of the DG methods (2.17), we now proceed to derive the error estimates. For the consistent DG method nIP and nNIPG, we have

$$\begin{aligned}
C_s|||u_I - u_h|||^2 &\leq B_h\left(u_I - u_h, u_I - u_h\right) \\
&= B_h\left(u_I - u, u_I - u_h\right) + B_h\left(u - u_h, u_I - u_h\right) \\
&\leq 2\,|||u_I - u|||\,|||u_I - u_h|||.
\end{aligned} \tag{3.12}$$

With (3.11), (3.12) and triangle inequality, we then obtain

$$\begin{aligned} |||u-u_h||| &\le |||u-u_I||| + |||u_I-u_h||| \\ &\le \left(1+\frac{2}{C_s}\right) |||u-u_I||| \le Ch^k \, |u|_{H^{k+1}} . \end{aligned} \tag{3.13}$$

For the nBZ method, we also have a similar result as (3.13). Since the nBZ method is not consistent, we need to make an extra effort to control the inconsistent term. In fact, the nBZ method relies on a rather heavy penalty and it may require some preconditioning since the condition number of the stiff matrix would be large due to this penalty. For more details, one can refer to [35]. We now summarize the above results as follows.

Theorem 3.1. *Consider the DG methods* (2.17) *for solving the ND problem* (2.1)*, with the finite element space* V_h *defined in* (2.9) *and the degrees of the piecewise polynomials* $k \ge 1$*. With suitable penalties on the jumps of element interfaces, such as in* (2.15), (2.16) *and* (2.14)*, there exists a unique numerical solution* $u_h \in V_h$ *to* (2.17)*. Assume the exact solution of* (2.1) $u \in H^{k+1}(\widetilde{\Omega})$*, then we have the following error estimate:*

$$|||u-u_h||| \le Ch^k \, \|u\|_{H^{k+1}} .$$

Remark 3.1. From (3.13), we can see the L^2 error $\|u-u_h\|_{L^2}$ is controlled by the interpolation error $|||u-u_I|||$ for the nIP and nNIPG methods. For integrable kernels with a fixed horizon δ, the discrete energy norm $|||\cdot|||$ is equivalent to the L^2 norm (see *e.g.* [43]). Thus, in this situation, we have the error estimate:

$$|||u-u_h||| \le C(\delta)h^{k+1}\|u\|_{H^{k+1}}.$$

3.4. Asymptotic compatibility

In the continuous level, when the horizon $\delta \to 0$, $f_\delta \to f_0$ in the dual space of the energy space S, and the solution u of the ND problem (2.1) converges to the solution u_{loc} of the corresponding local problem (2.5) (see *e.g.* [27]), *i.e.*

$$\|u-u_{loc}\|_{L^2} \to 0, \quad \text{as } \delta \to 0. \tag{3.14}$$

It is desirable to preserve such a limiting behavior in the numerical approximations, termed as *asymptotic compatibility* [53], such that $u_h \to u_{loc}$ when $\delta, h \to 0$ simultaneously. In [52], Tian and Du studied several existing numerical schemes and showed some of the numerical discretizations might not preserve such a limiting behavior. Later in [53], Tian and Du established an abstract mathematical framework for the numerical studies of a class of parametrized problems. We show that the DG method (2.17) is also asymptotically compatible under some appropriate conditions, stated in the following theorem.

Theorem 3.2. *Consider the DG methods* (2.17) *for solving the ND problem* (2.1)*, with the finite element space* V_h^k *defined in* (2.9)*,* $k \ge 1$*. Assume the exact solution* $u \in H^{1+\beta}(\widetilde{\Omega})$*,* $0 < \beta < 1$*, and* $\|u\|_{H^{1+\beta}(\widetilde{\Omega})}$ *is uniformly bounded with respect to the parameter* δ*. Then the DG methods* (2.17) *are asymptotically compatible,* i.e.

$$\|u_h-u_{loc}\|_{L^2} \to 0, \quad \text{as } \delta, h \to 0. \tag{3.15}$$

Remark 3.2. The assumption of the solution $u \in H^{1+\beta}(\widetilde{\Omega})$, $0 < \beta < 1$ is valid for the ND problem (2.1) with a truncated fractional kernel. The $H^{1+\beta}$ regularity of solution of such problem with any fixed δ was shown in Theorem 3.4 from [6]. However, it is quite difficult to obtain a δ-independent bound for the $H^{1+\beta}$ norm of u for general kernel and data f_δ, thus the uniform estimate in δ of $H^{1+\beta}$ regularity still remains an open question to be explored.

To prove the Theorem 3.2, we first introduce a result in Chapter 14 from [5] as follows, which plays a key role in obtaining the asymptotic compatibility in the fractional space.

Lemma 3.2. *Suppose T is a linear operator that maps $H^p(\Omega)$ to $L^2(\Omega)$ and $H^q(\Omega)$ to $L^2(\Omega)$, where p, q are some positive integers. then T maps $H^{(1-\theta)p+\theta q}(\Omega)$ to $L^2(\Omega)$, $\forall\, 0<\theta<1$. Moreover,*

$$\|T\|_{H^{(1-\theta)p+\theta q}\to L^2} \le \|T\|^{1-\theta}_{H^p\to L^2}\ \|T\|^{\theta}_{H^q\to L^2}\,.$$

In the proof of Theorem 3.2, we would like to consider the interpolation between $\mathcal{L}(H_0^1(\Omega)\cap H^2(\Omega), L^2(\Omega))$ and $\mathcal{L}(H_0^1(\Omega), L^2(\Omega))$, and this is the case when $p=1, q=2$ in Lemma 3.2.

Proof of Theorem 3.2. From the discrete Poincaré's inequality (3.3) and (3.12), we have

$$\|u_I-u_h\|_{L^2} \le C\,|||u_I-u_h||| \le C\,|||u_I-u|||.$$

Therefore, by the triangle inequality and (3.3), we can obtain

$$\|u-u_h\|_{L^2} \le \|u-u_I\|_{L^2} + \|u_I-u_h\|_{L^2} \le C\,|||u_I-u|||\,,$$

where $u_I\in V_h$ is the continuous interpolant of the exact solution u. For $u\in H^2(\widetilde{\Omega})$, from (3.13) we have

$$\|u-u_h\|_{L^2} \le Ch\,\|u\|_{H^2}\,.$$

Now we claim that

$$\|u-u_h\|_{L^2} \le C\ \|u\|_{H^1}\,. \tag{3.16}$$

If (3.16) holds true, then we can denote a linear operator $Tu := u-u_h$ and obtain

$$\|T\|_{H^1\to L^2}\le C,\quad \|T\|_{H^2\to L^2}\le Ch.$$

With $p=1, q=2$ in the Lemma 3.2, we conclude

$$\|T\|_{H^{1+\beta}\to L^2}\le Ch^{\beta},\quad \forall\, 0<\beta<1.$$

which indicates $\|u-u_h\|_{L^2}\le Ch^{\beta}\|u\|_{H^{1+\beta}}$, where C is independent of δ and h. Therefore, $\|u-u_h\|_{L^2}\to 0$ as $h\to 0$. since $\|u_{loc}-u\|_{L^2}\to 0$ as $\delta\to 0$, then

$$\|u_{loc}-u_h\|_{L^2} \le \|u_{loc}-u\|_{L^2} + \|u-u_h\|_{L^2} \to 0,\quad \text{as } \delta, h\to 0.$$

This indicates the DG methods (2.17) are asymptotically compatible for the exact solution $u\in H^{1+\beta}(\widetilde{\Omega})$, $0<\beta<1$.

It remains to prove (3.16). For the consistent DG method nIP and nNIPG, for any $v_h\in V_h$ we have

$$\begin{aligned} C_s|||v_h-u_h|||^2 &\le B_h\left(v_h-u_h, v_h-u_h\right)\\ &= B_h\left(v_h-u, v_h-u_h\right) + B_h\left(u-u_h, v_h-u_h\right)\\ &\le 2\,|||v_h-u|||\,|||v_h-u_h|||, \end{aligned} \tag{3.17}$$

which implies

$$|||v_h-u_h||| \le \frac{2}{C_s}|||v_h-u|||,\quad \forall\, v_h\in V_h.$$

Together with the Proposition 3.1, we have

$$\|v_h-u_h\|_{L^2} \le C\,|||v_h-u_h||| \le C\,|||v_h-u|||.$$

By triangle inequality and above inequality, we obtain

$$\begin{aligned}\|u-u_h\|_{L^2} &\le \|u-v_h\|_{L^2}+\|v_h-u_h\|_{L^2}\\ &\le \|u-v_h\|_{L^2}+C\,|||v_h-u|||.\end{aligned} \tag{3.18}$$

Since (3.18) holds true for any $v_h \in V_h$, then we take $v_h = 0$ and obtain

$$\|u-u_h\|_{L^2} \le \|u\|_{L^2}+C\,|||u||| = \|u\|_{L^2}+C\,\|u\|_{\mathcal{S}} \le C\,\|u\|_{H^1},$$

which proves the claim (3.16). The estimate $\|u\|_{\mathcal{S}} \le C\|u\|_{H^1}$ used in the last inequality can be found in, *e.g.*, [4]. □

3.5. Application to convection–diffusion problems

In this subsection, we consider the time-dependent convection–diffusion problem with nonlocal diffusion as follows:

$$\begin{cases} u_t + f(u)_x + \sigma\mathcal{L}_\delta u = f_s, & x\in\Omega,\ t>0,\\ u(x,0)=u_0(x), & x\in\Omega, \end{cases} \tag{3.19}$$

with periodic or compactly supported boundary conditions. $f(u)$ is the flux function and $f_s = f_s(x,t)$ is the source function. In the numerical approximation of convection-dominated problems, one of the computational challenges is the sharp transitions of the numerical solution. In particular, when $\sigma = 0$ the solution of (3.19) may evolve into shock discontinuities even with the smooth initial condition. There are many studies on the DG discretization of $f(u)_x$, see *e.g.* [51] and the references therein. Now assume we have the partition of the domain the same as in (2.6), we then construct the semi-discrete DG methods for (3.19): seek $u_h(\cdot,t)\in V_h$ such that

$$\int_{I_j}(u_h)_t\,v_h\,\mathrm{d}x + A_j(u_h,v_h) + \sigma B_{h,j}(u_h,v_h) = \int_{I_j} f_s v_h\,\mathrm{d}x, \quad \forall\, v_h\in V_h, \tag{3.20}$$

where $A_j(u_h,v_h)$ is defined as

$$A_j(u_h,v_h) = \hat{f}_{j+\frac12}(v_h)^-_{j+\frac12} - \hat{f}_{j-\frac12}(v_h)^+_{j-\frac12} - \int_{I_j} f(u_h)(v_h)_x\,\mathrm{d}x$$

with $\hat{f}_{j+\frac12}$ is the monotone flux and $\sum_j B_{h,j}(u_h,v_h) = B_h(u_h,v_h)$ with $B_h(u_h,v_h)$ defined in (2.12). Take $v_h = u_h$ in (3.20) and sum it over j, we have

$$\frac{\mathrm{d}}{2\mathrm{d}t}\|u_h(\cdot,t)\|^2_{L^2} + \sum_j A_j(u_h,u_h) + \sigma B_h(u_h,u_h) = (f_s,u_h).$$

Denote $F(u) = \int^u f(s)\,\mathrm{d}s$, we then have

$$\begin{aligned}\sum_j A_j(u_h,u_h) &= \sum_j\left(\hat{f}_{j+\frac12}(u_h)^-_{j+\frac12} - \hat{f}_{j-\frac12}(u_h)^+_{j-\frac12} - \left(F((u_h)^-_{j+\frac12}) - F((u_h)^+_{j-\frac12})\right)\right)\\ &= \sum_j\left(\hat{f}_{j+\frac12}(u_h)^-_{j+\frac12} - \hat{f}_{j+\frac12}(u_h)^+_{j+\frac12} - \left(F((u_h)^-_{j+\frac12}) - F((u_h)^+_{j+\frac12})\right)\right)\\ &= \sum_j\int_{(u_h)^-_{j+\frac12}}^{(u_h)^+_{j+\frac12}}\left(f(u) - \hat{f}\left((u_h)^-_{j+\frac12},(u_h)^+_{j+\frac12}\right)\right).\end{aligned}$$

By the above semi-discrete analysis, we can obtain the L^2-boundedness of the numerical solution, stated in the following theorem.

Theorem 3.3. *For the DG methods (3.20), the numerical solution satisfies*

$$\frac{\mathrm{d}}{2\mathrm{d}t}\left\|u_h(\cdot,t)\right\|_{L^2}^2+\sum_j\Theta_j+\sigma B_h\left(u_h,u_h\right)=\left(f_s,u_h\right), \tag{3.21}$$

where Θ_j *is given as*

$$\Theta_j=\int_{(u_h)^-_{j+\frac{1}{2}}}^{(u_h)^+_{j+\frac{1}{2}}}\left(f(u)-\hat{f}\left((u_h)^-_{j+\frac{1}{2}},(u_h)^+_{j+\frac{1}{2}}\right)\right)\mathrm{d}u\geq 0.$$

Thanks to the monotone flux, we are able to obtain $\Theta_j\geq 0$. We refer the readers to *e.g.* [51] for more details. From the previous analysis, we have $B_h(u_h,u_h)\geq C_s|||u_h|||^2\geq C\|u_h\|^2$. Therefore, when $f_s=0$, we can see that the numerical solution decays exponentially as time evolves.

Remark 3.3. For the time-dependent convection–diffusion problems, we recommend treating the diffusion term implicitly to have a relaxed CFL condition. Particularly, when both convection and diffusion co-exist, we could use the implicit-explicit (IMEX) Runge–Kutta method, such that the convection term could be treated explicitly and the diffusion term could be treated implicitly.

4. Numerical experiments

In this section, we show some numerical results to validate the theoretical results presented in the previous section. For simplicity, we take the uniform mesh, *i.e.* $h_j=h$. We consider the kernel function as

$$\gamma_\delta(s)=\frac{3-\alpha}{2\delta^{3-\alpha}}|s|^{-\alpha}\qquad\text{on }(-\delta,\delta).$$

Then $s^2\gamma_\delta(s)$ is integrable for $\alpha<3$. In the numerical tests, we take $\alpha=1/2,5/2$ such that γ_δ could have different kinds of singularities. We use the five-point Gauss–Legendre quadrature when computing the integrals for $s>\hat{h}$. And for those integrals that s close to 0, we use the exact integration because they are the improper integrals involving singularities. For both nIP and nNIPG methods, we take the penalty parameter μ as $\mu=5/h$ unless otherwise specified. We choose different values of δ in the numerical tests to show the good performance of the proposed numerical methods. For the time-dependent problems, we adopt the nIP method for nonlocal diffusion term, and the time discretization method is the 4th order implicit-explicit Runge–Kutta method with 6 stages [7]. Since we treat the convective term explicitly and the diffusive term implicitly, we take the CFL condition as $\tau=O(h)$ where τ is the time step. Specifically, the CFL number is taken as $\frac{0.3}{2k+1}$ for the numerical simulation of the time-dependent convection–diffusion equation.

Example 1. For the steady-state problem (2.1), we take the source term as

$$f_\delta(x)=-2\int_{-\delta}^{\delta}\gamma_\delta(s)\big(g(x+s)-g(x)\big)\,\mathrm{d}s,\quad x\in(0,\pi),$$

where $g(x)$ is defined by

$$g(x)=\begin{cases}\sin^6(x), & x\in(0,\pi),\\ 0, & \text{elsewhere.}\end{cases}$$

Thus the exact solution is $u(x)=g(x)$. The computational domain is $\Omega=(0,\pi)$.

TABLE 1. L^2 errors and convergence orders produced by the nIP scheme (2.17) and (2.15) when $k = 1$ in Example 1.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	3.996E−03	–	1.697E−03	–	1.706E−03	–	1.703E−03	–
	36	1.803E−03	1.963	7.483E−04	2.019	7.516E−04	2.022	7.502E−04	2.021
	48	1.019E−03	1.982	4.199E−04	2.008	4.214E−04	2.011	4.204E−04	2.013
	60	6.540E−04	1.989	2.685E−04	2.004	2.693E−04	2.007	2.688E−04	2.004
	72	4.548E−04	1.993	1.864E−04	2.002	1.868E−04	2.005	1.865E−04	2.005
	84	3.344E−04	1.995	1.369E−04	2.002	1.372E−04	2.003	1.370E−04	2.003
	96	2.562E−04	1.996	1.048E−04	2.001	1.050E−04	2.002	1.049E−04	2.001
$\frac{5}{2}$	24	3.996E−03	–	1.998E−03	–	2.129E−03	–	2.096E−03	–
	36	1.803E−03	1.963	8.435E−04	2.126	9.417E−04	2.012	8.980E−04	2.090
	48	1.019E−03	1.982	4.613E−04	2.098	5.288E−04	2.006	4.939E−04	2.078
	60	6.540E−04	1.989	2.900E−04	2.080	3.381E−04	2.004	3.112E−04	2.071
	72	4.548E−04	1.993	1.990E−04	2.067	2.347E−04	2.002	2.136E−04	2.064
	84	3.344E−04	1.995	1.449E−04	2.058	1.724E−04	2.002	1.554E−04	2.060
	96	2.562E−04	1.996	1.102E−04	2.051	1.320E−04	2.001	1.181E−04	2.056

TABLE 2. L^2 errors and convergence orders produced by the nIP scheme (2.17) and (2.15) when $k = 2$ in Example 1.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	1.797E−04	–	1.012E−04	–	1.049E−04	–	1.041E−04	–
	36	5.358E−05	2.984	2.936E−05	3.052	3.166E−05	2.954	3.090E−05	2.996
	48	2.266E−05	2.992	1.211E−05	3.079	1.345E−05	2.977	1.294E−05	3.025
	60	1.161E−05	2.995	6.078E−06	3.089	6.906E−06	2.986	6.586E−06	3.026
	72	6.724E−06	2.997	3.458E−06	3.094	4.003E−06	2.991	3.773E−06	3.056
	84	4.236E−06	2.998	2.146E−06	3.096	2.524E−06	2.993	2.370E−06	3.016
	96	2.838E−06	2.998	1.419E−06	3.096	1.692E−06	2.995	1.575E−06	3.061
$\frac{5}{2}$	24	9.746E−05	–	7.991E−05	–	7.997E−05	–	7.995E−04	–
	36	2.855E−05	3.028	2.366E−05	3.002	2.368E−05	3.002	2.367E−05	3.002
	48	1.200E−05	3.014	9.978E−05	3.001	9.985E−05	3.001	9.980E−05	3.001
	60	6.131E−05	3.008	5.108E−06	3.001	5.112E−06	3.001	5.109E−06	3.001
	72	3.545E−06	3.006	2.956E−06	3.000	2.958E−06	3.000	2.956E−06	3.001
	84	2.231E−06	3.004	1.861E−06	3.000	1.863E−06	3.000	1.862E−06	3.000
	96	1.494E−06	3.003	1.247E−06	3.000	1.248E−06	3.000	1.247E−06	3.000

From the results reported in Tables 1, 2 and 3, we can see the optimal order of convergence for nIP scheme (2.17) and (2.15) with various α and δ as the mesh is refined. We also observe the designed order of convergence for nNIPG scheme (2.17) and (2.16) for the results reported in Tables 4, 5 and 6. Note that the order is not optimal for an even degree in the classic NIPG method, which coincides with the case $\delta = 10^{-6}$ in Table 5. However, for $\gamma_\delta(s)$ is locally integrable ($\alpha = 1/2$) with fixed horizon $\delta = \pi/6$, the energy norm is equivalent to the L^2 norm. Then, from Remark 3.1, we have the optimal convergence order of 3 in the energy norm. The optimal rate of convergence can also be found in the case $\gamma_\delta(s)$ is locally integrable with horizon $\delta = \sqrt{h}$.

TABLE 3. L^2 errors and convergence orders produced by the nIP scheme (2.17) and (2.15) when $k = 3$ in Example 1.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	1.189E−05	–	2.668E−06	–	2.672E−06	–	2.672E−06	–
	36	1.357E−06	5.353	5.198E−07	4.034	5.206E−07	4.034	5.204E−07	4.035
	48	3.576E−07	4.635	1.637E−07	4.016	1.639E−07	4.017	1.638E−07	4.017
	60	1.339E−07	4.403	6.692E−08	4.009	6.699E−08	4.010	6.696E−08	4.010
	72	6.112E−08	4.300	3.224E−08	4.006	3.226E−08	4.007	3.225E−08	4.007
	84	3.194E−08	4.209	1.739E−08	4.004	1.740E−08	4.005	1.740E−08	4.004
	96	1.834E−08	4.157	1.019E−08	4.003	1.020E−08	4.004	1.019E−08	4.004
$\frac{5}{2}$	24	1.189E−05	–	3.182E−06	–	3.185E−06	–	3.184E−06	–
	36	1.357E−06	5.353	6.338E−07	3.980	6.345E−07	3.979	6.342E−07	3.980
	48	3.576E−07	4.635	2.011E−07	3.990	2.014E−07	3.990	2.012E−07	3.991
	60	1.339E−07	4.403	8.246E−08	3.994	8.259E−08	3.994	8.251E−08	3.994
	72	6.112E−08	4.300	3.979E−08	3.996	3.986E−08	3.995	3.982E−08	3.996
	84	3.194E−08	4.209	2.149E−08	3.997	2.154E−08	3.995	2.150E−08	3.996
	96	1.834E−08	4.157	1.260E−08	3.997	1.264E−08	3.990	1.261E−08	3.995

TABLE 4. L^2 errors and convergence orders produced by the nNIPG scheme (2.17) and (2.16) when $k = 1$ in Example 1.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	2.107E−03	–	1.704E−03	–	1.711E−03	–	1.709E−03	–
	36	9.325E−04	2.011	7.498E−04	2.025	7.519E−04	2.028	7.513E−04	2.026
	48	5.237E−04	2.005	4.204E−04	2.011	4.213E−04	2.014	4.208E−04	2.015
	60	3.349E−04	2.003	2.687E−04	2.006	2.691E−04	2.008	2.690E−04	2.006
	72	2.325E−04	2.002	1.865E−04	2.004	1.867E−04	2.006	1.866E−04	2.006
	84	1.708E−04	2.002	1.369E−04	2.003	1.371E−04	2.004	1.370E−04	2.003
	96	1.307E−04	2.001	1.048E−04	2.002	1.049E−04	2.003	1.049E−04	2.002
$\frac{5}{2}$	24	2.107E−03	–	1.710E−03	–	1.716E−03	–	1.714E−03	–
	36	9.325E−04	2.011	7.527E−04	2.024	7.567E−04	2.020	7.548E−04	2.023
	48	5.237E−04	2.005	4.218E−04	2.013	4.244E−04	2.010	4.230E−04	2.013
	60	3.349E−04	2.003	2.695E−04	2.008	2.713E−04	2.005	2.702E−04	2.008
	72	2.325E−04	2.002	1.869E−04	2.006	1.883E−04	2.004	1.874E−04	2.007
	84	1.708E−04	2.002	1.372E−04	2.005	1.382E−04	2.003	1.376E−04	2.005
	96	1.307E−04	2.001	1.050E−04	2.004	1.058E−04	2.002	1.053E−04	2.004

Example 2. Consider a non-smooth case for (2.1). We take the locally integrable kernel γ_δ that $\delta = 1/8$, $\alpha = 1/2$. $g(x)$ is taken as

$$g(x) = \begin{cases} 1, & x \in \left(\frac{1}{4}, \frac{3}{4}\right), \\ 0, & \text{elsewhere.} \end{cases}$$

With the definition of f_δ similar in the Example 1, we have the solution $u(x) = g(x)$. The computational domain is $\Omega = (0, 1)$.

TABLE 5. L^2 errors and convergence orders produced by the nNIPG scheme (2.17) and (2.16) when $k = 2$ in Example 1.

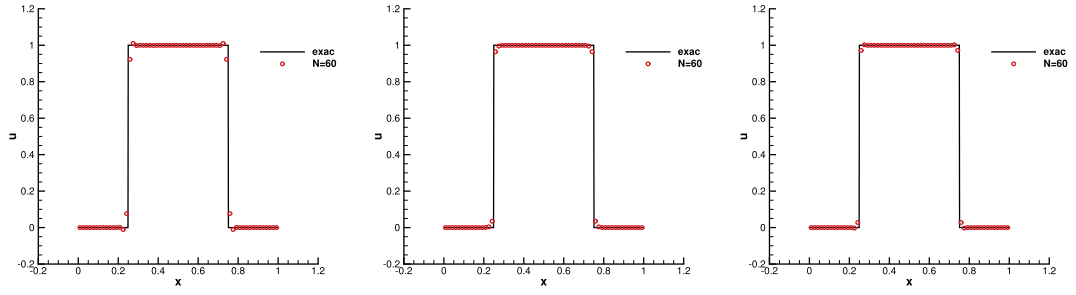
α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	5.449E−04	–	1.085E−04	–	1.136E−04	–	1.123E−04	–
	36	2.398E−04	2.024	3.168E−05	3.036	3.541E−05	2.876	3.348E−05	2.984
	48	1.345E−04	2.012	1.312E−05	3.065	1.565E−05	2.839	1.409E−05	3.008
	60	8.591E−05	2.007	6.600E−06	3.077	8.417E−06	2.779	7.179E−06	3.023
	72	5.961E−05	2.005	3.761E−06	3.084	5.134E−06	2.711	4.127E−06	3.036
	84	4.377E−05	2.003	2.337E−06	3.088	3.416E−06	2.642	2.592E−06	3.018
	96	3.350E−05	2.003	1.546E−06	3.091	2.422E−06	2.577	1.726E−06	3.047
$\frac{5}{2}$	24	4.259E−04	–	3.580E−04	–	4.305E−04	–	4.128E−04	–
	36	1.870E−04	2.030	1.311E−04	2.478	1.899E−04	2.018	1.655E−04	2.254
	48	1.047E−04	2.015	6.397E−05	2.494	1.065E−04	2.009	8.640E−05	2.259
	60	6.688E−05	2.009	3.662E−05	2.500	6.810E−05	2.006	5.218E−05	2.260
	72	4.639E−05	2.006	2.321E−05	2.502	4.726E−05	2.004	3.456E−05	2.259
	84	3.406E−05	2.004	1.578E−05	2.503	3.470E−05	2.003	2.440E−05	2.259
	96	2.607E−05	2.003	1.130E−05	2.503	2.656E−05	2.002	1.805E−05	2.258

TABLE 6. L^2 errors and convergence orders produced by the nNIPG scheme (2.17) and (2.16) when $k = 3$ in Example 1.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	9.843E−06	–	2.678E−06	–	2.683E−06	–	2.682E−06	–
	36	1.915E−06	4.037	5.208E−07	4.038	5.217E−07	4.039	5.213E−07	4.039
	48	6.027E−07	4.019	1.639E−07	4.019	1.641E−07	4.020	1.640E−07	4.020
	60	2.462E−07	4.011	6.697E−08	4.011	6.706E−08	4.012	6.701E−08	4.011
	72	1.186E−07	4.008	3.225E−08	4.007	3.229E−08	4.008	3.227E−08	4.008
	84	6.396E−08	4.005	1.740E−08	4.005	1.741E−08	4.006	1.740E−08	4.005
	96	3.747E−08	4.004	1.019E−08	4.004	1.020E−08	4.004	1.020E−08	4.004
$\frac{5}{2}$	24	9.843E−06	–	4.049E−06	–	4.311E−06	–	4.242E−06	–
	36	1.915E−06	4.037	7.450E−07	4.175	8.337E−07	4.052	7.933E−07	4.135
	48	6.027E−07	4.019	2.271E−07	4.129	2.617E−07	4.028	2.435E−07	4.106
	60	2.463E−07	4.011	9.091E−08	4.103	1.067E−08	4.019	9.771E−08	4.092
	72	1.186E−07	4.008	4.315E−08	4.087	5.130E−08	4.018	4.640E−08	4.084
	84	6.396E−08	4.005	2.302E−08	4.076	2.760E−08	4.022	2.474E−08	4.080
	96	3.747E−08	4.004	1.337E−08	4.068	1.612E−08	4.029	1.435E−08	4.079

Figure 1 plots of the numerical solution produced by the nIP scheme (2.17) and (2.15), from which we can see there are no obvious oscillations in the numerical solutions, indicating the proposed algorithm can handle the singularities and discontinuities well.

Example 3. In this example, we consider the time-dependent convection–diffusion problems (3.19) with $f(u) = u$ and the periodic boundary condition. The coefficient σ is taken to be $1/2$. With a suitable choice of source function f_s, we have the exact solution as $u(x,t) = e^{-t}\sin^6(x)$. The computational domain $\Omega = (0,\pi)$ and the final time is $T = 2.2$.



(a) (b) (c)

FIGURE 1. Plots of the numerical solution in Example 2 produced by the nIP scheme (2.17) and (2.15), $\delta = 1/8, \alpha = 1/2$, $N = 60$. Solid line: exact solution. Red circles: numerical solutions. (a) $k = 1$. (b) $k = 2$. (c) $k = 3$.

TABLE 7. L^2 errors and convergence orders produced by the scheme (3.20) and (2.15) when $k = 1$ in Example 3, $\sigma = 1/2$.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	4.361E−04	–	2.196E−04	–	1.971E−04	–	2.000E−04	–
	36	1.963E−04	1.969	1.049E−04	1.823	8.513E−05	2.071	8.854E−05	2.009
	48	1.109E−04	1.984	6.214E−05	1.819	4.732E−05	2.041	4.979E−05	2.001
	60	7.114E−05	1.990	4.126E−05	1.835	3.010E−05	2.027	3.194E−05	1.989
	72	4.946E−05	1.994	2.942E−05	1.855	2.083E−05	2.019	2.221E−05	1.993
	84	3.636E−05	1.995	2.205E−05	1.872	1.527E−05	2.013	1.636E−05	1.985
	96	2.785E−05	1.997	1.714E−05	1.886	1.168E−05	2.010	1.254E−05	1.988
$\frac{5}{2}$	24	4.361E−04	–	2.207E−04	–	2.347E−04	–	2.311E−04	–
	36	1.963E−04	1.969	9.316E−05	2.127	1.036E−04	2.016	9.897E−05	2.092
	48	1.109E−04	1.984	5.096E−05	2.097	5.815E−05	2.008	5.443E−05	2.078
	60	7.114E−05	1.990	3.205E−05	2.078	3.717E−05	2.005	3.430E−05	2.070
	72	4.946E−05	1.994	2.199E−05	2.067	2.580E−05	2.004	2.354E−05	2.064
	84	3.636E−05	1.995	1.602E−05	2.056	1.895E−05	2.002	1.714E−05	2.059
	96	2.785E−05	1.997	1.218E−05	2.049	1.450E−05	2.002	1.303E−05	2.055

From the results reported in Tables 7, 8 and 9, we show the error and order of convergence of $\|u - u_h\|$ produced by the scheme (3.20) and (2.15). For $k = 3$, we take the penalty parameter $\mu = 7/h$.

Example 4. We consider the convection–diffusion problem in Example 3 with piecewise constant initial values

$$u_0(x) = \begin{cases} u_l, & x < 0, \\ u_r, & x > 0. \end{cases}$$

We consider two kinds of initial conditions:

(i) $u_l = 0$, $u_r = 1$.
(ii) $u_l = 1$, $u_r = 0$.

To see the nonlocal diffusion effect, we take four different values of σ. We take locally integrable kernel with $\alpha = 1/2$, $\delta = 1/8$. The computational domain is $\Omega = (-9, 9)$, and the final time is $T = 2$.

TABLE 8. L^2 errors and convergence orders produced by the scheme (3.20) and (2.15) when $k = 2$ in Example 3, $\sigma = 1/2$.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	1.973E−05	–	1.386E−05	–	1.380E−05	–	1.393E−05	–
	36	5.900E−06	2.977	3.985E−06	3.074	3.971E−06	3.072	4.160E−06	2.981
	48	2.498E−06	2.987	1.635E−06	3.096	1.627E−06	3.101	1.731E−06	3.048
	60	1.282E−06	2.991	8.219E−07	3.083	8.166E−07	3.090	8.818E−07	3.022
	72	7.427E−07	2.993	4.698E−07	3.068	4.663E−07	3.073	5.021E−07	3.089
	84	4.681E−07	2.995	2.933E−07	3.055	2.910E−07	3.060	3.157E−07	3.011
	96	3.138E−07	2.996	1.953E−07	3.046	1.936E−07	3.049	2.091E−07	3.084
$\frac{5}{2}$	24	1.090E−05	–	9.065E−06	–	9.029E−06	–	9.035E−06	–
	36	3.184E−06	3.034	2.675E−06	3.010	2.659E−06	3.016	2.663E−06	3.013
	48	1.336E−06	3.019	1.125E−06	3.009	1.118E−06	3.012	1.120E−06	3.011
	60	6.821E−06	3.013	5.752E−07	3.008	5.711E−07	3.009	5.723E−07	3.009
	72	3.941E−07	3.009	3.324E−07	3.008	3.300E−07	3.008	3.307E−07	3.008
	84	2.479E−07	3.007	2.091E−07	3.007	2.076E−07	3.007	2.080E−07	3.007
	96	1.659E−07	3.006	1.400E−07	3.007	1.390E−07	3.006	1.392E−07	3.006

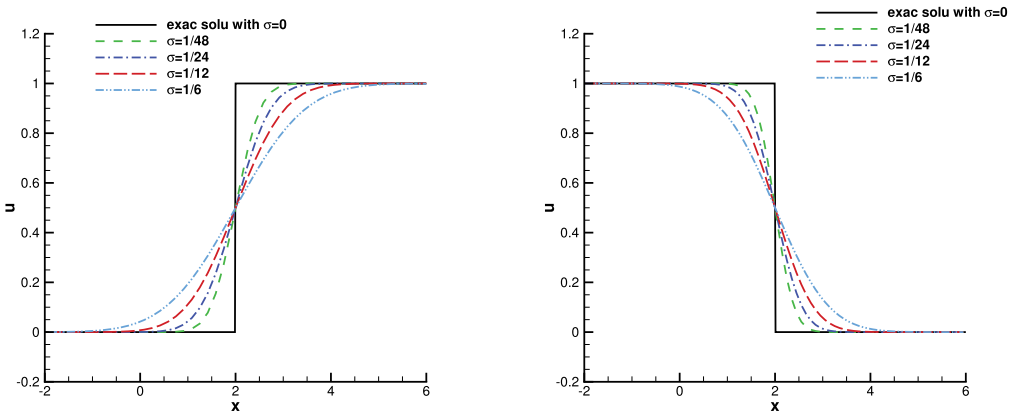
TABLE 9. L^2 errors and convergence orders produced by the scheme (3.20) and (2.15) when $k = 3$ in Example 3, $\sigma = 1/2$.

α	N	$\delta = 10^{-6}$		$\delta = \pi/6$		$\delta = 2.5h$		$\delta = \sqrt{h}$	
		L^2 error	Order	L^2 error	Order	L^2 error	Order	L^2 error	Order
$\frac{1}{2}$	24	5.539E−07	–	3.290E−07	–	3.051E−07	–	3.083E−07	–
	36	1.060E−07	4.078	6.818E−08	3.882	5.843E−08	4.076	5.986E−08	4.042
	48	3.256E−08	4.103	2.258E−08	3.842	1.829E−08	4.037	1.887E−08	4.014
	60	1.307E−08	4.091	9.575E−09	3.844	7.455E−09	4.022	7.724E−09	4.002
	72	6.216E−09	4.075	4.740E−09	3.857	3.586E−09	4.015	3.726E−09	3.999
	84	3.324E−09	4.061	2.610E−09	3.872	1.932E−09	4.010	2.014E−09	3.992
	96	1.936E−09	4.050	1.553E−09	3.886	1.132E−09	4.008	1.181E−09	3.994
$\frac{5}{2}$	24	5.539E−07	–	3.557E−07	–	3.558E−07	–	3.557E−07	–
	36	1.060E−07	4.078	7.051E−08	3.991	7.055E−08	3.990	7.052E−08	3.991
	48	3.256E−08	4.103	2.233E−08	3.996	2.235E−08	3.995	2.234E−08	3.996
	60	1.307E−08	4.091	9.152E−09	3.998	9.162E−09	3.997	9.154E−09	3.998
	72	6.216E−09	4.075	4.415E−09	3.999	4.420E−09	3.998	4.416E−09	3.999
	84	3.324E−09	4.061	2.383E−09	3.999	2.387E−09	3.997	2.384E−09	3.998
	96	1.936E−09	4.050	1.397E−09	3.999	1.400E−09	3.995	1.398E−09	3.998

From Table 10, we can see the errors and orders of convergence for the Riemann problems in Example 4. The reference solution is computed with the refined mesh $h = 1/50$ and the L^2 errors are computed by the absolute errors $\|u_h - u_{\text{ref}}\|_{L^2([-2,6])}$. The orders are around 3 and even higher when h becomes smaller. This phenomenon may be because the spatial error is small compared to the temporal error as the mesh is refined. Figure 2 plots the numerical solution produced by the scheme (3.20) and (2.15), from which we can see the numerical solutions become smooth with the nonlocal diffusion term, compared to the one without the nonlocal diffusion. We can also see that the transition width becomes larger as we increase the coefficient of the nonlocal diffusion term. This indicates that nonlocal diffusion can be another option when adding artificial diffusion to numerical schemes, especially for physical problems containing long-range interactions.

TABLE 10. L^2 errors and convergence orders produced by the scheme (3.20) and (2.15) when $k = 2$ in Example 4, $\delta = 1/8$, $\sigma = 1/5$. The L^2 errors are computed on $[-2, 6]$.

	h	1/2	1/3	1/4	1/5	1/6	1/7	1/8
I.C. (i)	L^2 error	2.223E−04	6.897E−05	2.913E−05	1.456E−05	8.084E−06	4.815E−06	2.997E−06
	Order	–	2.886	2.996	3.107	3.228	3.362	3.549
I.C. (ii)	L^2 error	2.223E−04	6.897E−05	2.913E−05	1.456E−05	8.084E−06	4.815E−06	2.997E−06
	Order	–	2.886	2.996	3.107	3.228	3.362	3.549



(a) (b)

FIGURE 2. Plots of the numerical solution in Example 4 produced by the scheme (3.20) and (2.15), $\delta = 1/8, \alpha = 1/2$, $k = 2$, $h = 1/6$. Solid line: exact solution with $\sigma = 0$. Green dash line: $\sigma = 1/48$. Blue dash-dot line: $\sigma = 1/24$. Red long dash line: $\sigma = 1/12$. Light blue dash dot dot line: $\sigma = 1/6$. (a) I.C. (i). (b) I.C. (ii)

TABLE 11. L^2 errors and convergence orders produced by the scheme (3.20) and (2.15) when $k = 2$ in Example 5, $\delta = \pi/6$, $\sigma = \pi/15$.

N	48	60	72	84	96	108	120
L^2 error	3.965E−05	2.903E−05	2.238E−05	1.731E−05	1.334E−05	1.040E−05	8.351E−05
Order	–	1.397	1.426	1.667	1.951	2.113	2.083

Example 5. Consider the viscous Burgers' equation with $f(u) = u^2/2$ in (3.19). We take the locally integrable kernel with $\alpha = 1/2$, $\delta = \pi/6$. The computational domain is $\Omega = (0, 2\pi)$ and the final time is $T = 1.6$. Note that at time $T = 1.6$, the exact solution of the inviscid Burgers' equation ($\sigma = 0$) contains a shock discontinuity inside the domain.

In Example 5, we investigate the nonlocal diffusion term for a convection–diffusion equation with the nonlinear convective term. From Table 11, we can see the errors and orders of convergence for the Example 5. The reference solution is computed with the refined mesh $h = \pi/250$ and the L^2 errors are computed by the absolute errors $\|u_h - u_{\text{ref}}\|_{L^2([0,2\pi])}$. The convergence rate is around 2 and is not optimal, which could be associated with the lack of regularity of the exact solution. We still observe the same phenomenon as in Example 4 from the numerical solutions plotted in Figure 3, which again demonstrates the effect of the nonlocal diffusion.

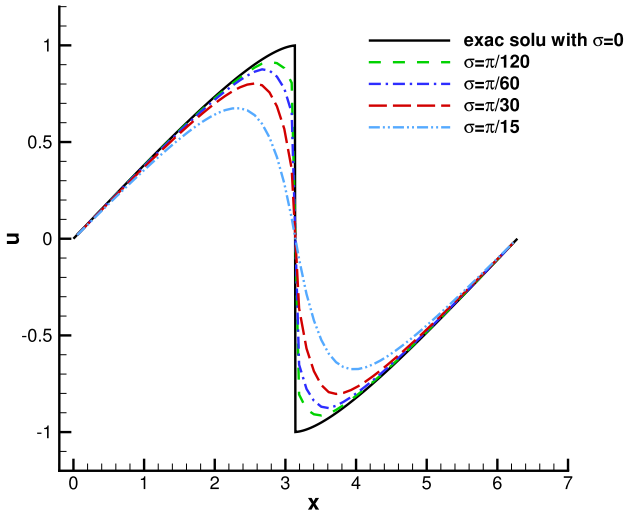


FIGURE 3. Plot of the numerical solution in Example 5 produced by the scheme (3.20) and (2.15), $\delta = \pi/6, \alpha = 1/2, k = 2, N = 60$. Solid line: exact solution with $\sigma = 0$. Green dash line: $\sigma = \pi/120$. Blue dash dot line: $\sigma = \pi/60$. Red long dash line: $\sigma = \pi/30$. Light blue dash dot dot line: $\sigma = \pi/15$.

5. Conclusion

In this paper, we propose a class of penalty discontinuous Galerkin (DG) methods for nonlocal diffusion (ND) problems. The ND problem is an integral equation with a possibly singular kernel, which can describe physical systems with long-range interactions and allow solutions with low regularity. Therefore, it is natural to adopt the DG method to compute the ND problem when the solution contains singularities and discontinuities. The proposed DG methods in this paper have corresponding local counterparts, indicating that these methods work at least for the vanished horizon. We also present theoretical results on boundedness, stability, *a priori* error estimates, and asymptotic compatibility. To illustrate the nonlocal diffusion effect, we consider the time-dependent convection–diffusion equation with nonlocal diffusion and conduct a semi-discrete analysis. Numerical tests demonstrate that our methods are high-order, stable, and asymptotically compatible. Currently, the DG methods are developed for one-dimensional ND problems, and extending these methods to multidimensional problems is part of our future work.

Appendix A. The technical proof of Lemma 3.1

Proof. First let us consider two cases: $0 < s < \varepsilon\rho$ and $\varepsilon\rho < s < \hat{h}$, where $\varepsilon \in (0,1)$ will be determined later. In fact, with the definition of $|\cdot|_{J,h}$ in (3.1), we have

$$|v_h|_{J,h}^2 = J_1 + J_2, \tag{A.1}$$

where J_1 and J_2 are defined as

$$J_1 = 2\sum_j h_j \int_0^{\varepsilon\rho} \gamma_\delta(s)\,\frac{1}{s}\int_{I_{j,2}^s} \left(E_s^+ v_h - [\![v_h]\!]_{j+\frac{1}{2}}\right)^2 \mathrm{d}x\,\mathrm{d}s,$$

$$J_2 = 2\sum_j h_j \int_{\varepsilon\rho}^{\hat{h}} \gamma_\delta(s)\,\frac{1}{s}\int_{I_{j,2}^s} \left(E_s^+ v_h - [\![v_h]\!]_{j+\frac{1}{2}}\right)^2 \mathrm{d}x\,\mathrm{d}s.$$

We now consider J_1, *i.e.* the case $0 < s < \varepsilon\rho$, and we have

$$\begin{aligned} J_1 \le 4\sum_j h_j \int_0^{\varepsilon\rho} \gamma_\delta(s)\,\frac{1}{s}\int_{I^s_{j,2}} \left(v_h(x+s) - v_h(x^+_{j+\frac12})\right)^2 \mathrm{d}x\,\mathrm{d}s \\ + 4\sum_j h_j \int_0^{\varepsilon\rho} \gamma_\delta(s)\,\frac{1}{s}\int_{I^s_{j,2}} \left(v_h(x) - v_h(x^-_{j+\frac12})\right)^2 \mathrm{d}x\,\mathrm{d}s. \end{aligned} \tag{A.2}$$

In fact, we have

$$\begin{aligned} \int_{I^s_{j,2}} \left(v_h(x+s) - v_h(x^+_{j+\frac12})\right)^2 \mathrm{d}x &\le \int_{I^s_{j,2}} \left(\int_{x_{j+\frac12}}^{x+s} \partial_y v_h(y)\,\mathrm{d}y\right)^2 \mathrm{d}x \\ &\le \int_{I^s_{j,2}} \left(\int_{x_{j+\frac12}}^{x+s} 1\,\mathrm{d}y\right)^2 \mathrm{d}x\ \|\partial_y v_h\|^2_{L^\infty(I^s_{j+1,2})} \\ &= \frac{1}{3}s^3\ \|\partial_y v_h\|^2_{L^\infty(I^s_{j+1,2})}. \end{aligned} \tag{A.3}$$

Similarly, we also have

$$\int_{I^s_{j,2}} \left(v_h(x) - v_h(x^-_{j+\frac12})\right)^2 \mathrm{d}x \le \frac{1}{3}s^3\ \|\partial_y v_h\|^2_{L^\infty(I^s_{j,2})}. \tag{A.4}$$

Plug (A.3) and (A.4) into (A.2), with inverse estimates we then obtain

$$\begin{aligned} J_1 &\le \frac{4}{3}\int_0^{\varepsilon\rho} s^2\gamma_\delta(s)\,\mathrm{d}s \sum_j h_j \left(\|\partial_y v_h\|^2_{L^\infty(I^s_{j+1,2})} + \|\partial_y v_h\|^2_{L^\infty(I^s_{j,2})}\right) \\ &\le C\int_0^{\varepsilon\rho} s^2\gamma_\delta(s)\,\mathrm{d}s \sum_j \|\partial_y v_h\|^2_{L^2(I_j)}. \end{aligned} \tag{A.5}$$

Next we consider the $|\cdot|_{E,h}$ defined in (3.1). Similar to (A.1), we divide the $|v_h|^2_{E,h}$ into three parts:

$$|v_h|^2_{E,h} = E_1 + E_2 + E_3, \tag{A.6}$$

where $E_1 = E_{1,1} + E_{1,2}$, $E_2 = E_{2,1} + E_{2,2}$, and E_3 are given as

$$\begin{aligned} E_{1,1} &= 2\sum_j \int_0^{\varepsilon\rho} \gamma_\delta(s) \int_{I^s_{j,1}} g_{v_h}(x,s)^2\,\mathrm{d}x\,\mathrm{d}s, \quad E_{1,2} = \sum_j \int_0^{\varepsilon\rho} \gamma_\delta(s) \int_{I^s_{j,2}} g_{v_h}(x,s)^2\,\mathrm{d}x\,\mathrm{d}s, \\ E_{2,1} &= 2\sum_j \int_{\varepsilon\rho}^{\hat h} \gamma_\delta(s) \int_{I^s_{j,1}} g_{v_h}(x,s)^2\,\mathrm{d}x\,\mathrm{d}s, \quad E_{2,2} = 2\sum_j \int_{\varepsilon\rho}^{\hat h} \gamma_\delta(s) \int_{I^s_{j,2}} g_{v_h}(x,s)^2\,\mathrm{d}x\,\mathrm{d}s, \\ E_3 &= 2\sum_j \int_{\hat h}^{\delta} \gamma_\delta(s) \int_{I_j} g_{v_h}(x,s)^2\,\mathrm{d}x\,\mathrm{d}s. \end{aligned}$$

We claim that for ε sufficiently small but independent of h and δ, we have

$$E_{1,1} \ge C\int_0^{\varepsilon\rho} s^2\gamma_\delta(s)\,\mathrm{d}s \sum_j \|\partial_y v_h\|^2_{L^2(I_j)}. \tag{A.7}$$

If (A.7) holds, then J_1 can be bounded by $E_{1,1}$. Now we prove that (A.7) is true when ε is appropriately chosen.

From the definition of $g_{v_h}(x,s)$ in (2.13), we know that $g_{v_h}(x,s) = E_s^+ v_h(x)$ on $I_{j,1}^s \times (0, \varepsilon\rho)$. By Taylor expansion, we obtain

$$E_s^+ v_h(x) = v_h(x+s) - v_h(x) = \sum_{l=1}^{k} \frac{\partial_x^l v_h(x)}{l!} s^l.$$

Thus, with Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\int_{I_{j,1}^s} g_{v_h}(x,s)^2 \,\mathrm{d}x &= \int_{I_{j,1}^s} \left(\sum_{l=1}^{k} \frac{\partial_x^l v_h(x)}{l!} s^l \right)^2 \mathrm{d}x \\
&= \int_{I_{j,1}^s} \left(\partial_x v_h(x) s\right)^2 \mathrm{d}x + 2 \int_{I_{j,1}^s} \left(\partial_x v_h(x) s\right) \left(\sum_{l=2}^{k} \frac{\partial_x^l v_h(x)}{l!} s^l \right) \mathrm{d}x \\
&\quad + \int_{I_{j,1}^s} \left(\sum_{l=2}^{k} \frac{\partial_x^l v_h(x)}{l!} s^l \right)^2 \mathrm{d}x \\
&\geq \frac{1}{2} s^2 \int_{I_{j,1}^s} \left(\partial_x v_h(x)\right)^2 \mathrm{d}x - \int_{I_{j,1}^s} \left(\sum_{l=2}^{k} \frac{\partial_x^l v_h(x)}{l!} s^l \right)^2 \mathrm{d}x.
\end{aligned} \tag{A.8}$$

Note that $s \in (0, \varepsilon\rho)$, then with Cauchy–Schwarz inequality and inverse estimates we have

$$\begin{aligned}
&\int_{I_{j,1}^s} \left(\sum_{l=2}^{k} \frac{\partial_x^l v_h(x)}{l!} s^l \right)^2 \mathrm{d}x \\
&= \int_{I_{j,1}^s} \left(\sum_{l=2}^{k} \partial_x^l v_h(x) (h_j - s)^{l-1} \cdot (h_j - s)^{-l+1} \frac{s^l}{l!} \right)^2 \mathrm{d}x \\
&\leq s^2 \left(\sum_{l=2}^{k} \frac{1}{(l!)^2} \left(\frac{s}{h_j - s} \right)^{2l-2} \right) \left(\sum_{l=2}^{k} (h_j - s)^{2l-2} \int_{I_{j,1}^s} \left(\partial_x^l v_h(x)\right)^2 \mathrm{d}x \right) \\
&\leq C_2 \left(\sum_{l=2}^{k} \left(\frac{\varepsilon}{1-\varepsilon} \right)^{2l-2} \right) s^2 \int_{I_{j,1}^s} \left(\partial_x v_h(x)\right)^2 \mathrm{d}x,
\end{aligned} \tag{A.9}$$

where $C_2 > 0$ is independent of h and δ. Since

$$\sum_{l=2}^{k} \left(\frac{\varepsilon}{1-\varepsilon} \right)^{2l-2} = \varepsilon_1 \frac{1 - \varepsilon_1^{k-1}}{1 - \varepsilon_1}, \quad \varepsilon_1 = \left(\frac{\varepsilon}{1-\varepsilon} \right)^2,$$

we then take ε statisfying $\varepsilon \leq 1/(1 + \sqrt{4C_2 + 1})$ such that $\varepsilon_1 \leq 1/(4C_2 + 1)$, which implies

$$C_2 \varepsilon_1 \frac{1 - \varepsilon_1^{k-1}}{1 - \varepsilon_1} \leq \frac{C_2 \varepsilon_1}{1 - \varepsilon_1} \leq \frac{1}{4}.$$

Therefore, plugging (A.9) into (A.8) with small ε satisfying the above inequality, we obtain

$$\int_{I_{j,1}^s} g_{v_h}(x,s)^2 \,\mathrm{d}x \geq \frac{1}{4} s^2 \int_{I_{j,1}^s} \left(\partial_x v_h(x)\right)^2 \mathrm{d}x. \tag{A.10}$$

If we take the Taylor expansion of $E_s^+ v_h(x)$ that

$$E_s^+ v_h(x) = -\sum_{l=1}^{k} \frac{(\partial_x^l v_h)(x+s)}{l!}(-s)^l,$$

we can still obtain a similar result in the following.

$$\int_{I_{j,1}^s} g_{v_h}(x,s)^2 \,\mathrm{d}x \geq \frac{1}{4}\, s^2 \int_{I_{j,1}^s} \left((\partial_x v_h)(x+s)\right)^2 \mathrm{d}x. \tag{A.11}$$

Therefore, with (A.10) and (A.11) we can obtain

$$\begin{aligned} E_{1,1} &= \sum_j \int_0^{\varepsilon\rho} \gamma_\delta(s) \left(2 \int_{I_{j,1}^s} g_{v_h}(x,s)^2 \,\mathrm{d}x\right) \mathrm{d}s \\ &\geq \frac{1}{4} \int_0^{\varepsilon\rho} s^2 \gamma_\delta(s) \sum_j \left(\int_{I_{j,1}^s} (\partial_x v_h(x))^2 \,\mathrm{d}x + \int_{I_{j,1}^s} ((\partial_x v_h)(x+s))^2 \,\mathrm{d}x\right) \mathrm{d}s \\ &\geq \frac{1}{4} \int_0^{\varepsilon\rho} s^2 \gamma_\delta(s) \sum_j \int_{I_j} (\partial_x v_h(x))^2 \,\mathrm{d}x \,\mathrm{d}s, \end{aligned} \tag{A.12}$$

which is exactly the result we want, as we claim in (A.7). Now let us fix $\varepsilon = \varepsilon_0$ so that (A.7) holds, we then consider the another case: $\varepsilon_0\rho < s < \hat{h}$. In fact, we have

$$\begin{aligned} J_2 &\leq 2 \sum_j h_j \int_{\varepsilon_0\rho}^{\hat{h}} \gamma_\delta(s) \frac{1}{\varepsilon_0 \rho} \int_{I_{j,2}^s} \left(E_s^+ v_h - [\![v_h]\!]_{j+\frac{1}{2}}\right)^2 \mathrm{d}x \,\mathrm{d}s \\ &\leq \frac{2\nu}{\varepsilon_0} \sum_j h_j \int_{\varepsilon_0\rho}^{\hat{h}} \gamma_\delta(s) \int_{I_{j,2}^s} \left(E_s^+ v_h - [\![v_h]\!]_{j+\frac{1}{2}}\right)^2 \mathrm{d}x \,\mathrm{d}s = \frac{\nu}{\varepsilon_0} E_{2,2}, \end{aligned} \tag{A.13}$$

where ν is defined in (2.7). Therefore, by (A.5), (A.6), (A.7) and (A.13), we obtain (3.9) and complete the proof. □

Funding

Q. Du's research is partially supported by US National Science Foundation Grant DMS-2012562. L. Ju's research is partially supported by US National Science Foundation Grant DMS-2109633. J. Lu's research is partially supported by Science and Technology Program of Guangzhou 2023A04J1300 and the Fundamental Research Funds for the Central Universities. X. Tian's research is partially supported by US National Science Foundation Grant DMS-2111608.

References

[1] D.N. Arnold, F. Brezzi, B. Cockburn and L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* **39** (2002) 1749–1779.

[2] I. Babuška and M. Zlámal, Nonconforming elements in the finite element method with penalty. *SIAM J. Numer. Anal.* **10** (1973) 863–875.

[3] C.E. Baumann and J.T. Oden, A discontinuous *hp* finite element method for convection–diffusion problems. *Comput. Methods Appl. Mech. Eng.* **175** (1999) 311–341.

[4] J. Bourgain, H. Brezis and P. Mironescu, Another look at Sobolev spaces, edited by J.L. Menaldi, E. Rofman and A. Sulem. In: *Optimal Control and Partial Differential Equations: In Honor of Professor Alain Benssoussan's 60th Birthday.* IOS Press (2001) 439–455.

[5] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods, 3rd edition. Springer, New York (2008).
[6] O. Burkovska and M. Gunzburger, Regularity analyses and approximation of nonlocal variational equality and inequality problems. *J. Math. Anal. Appl.* **478** (2019) 1027–1048.
[7] M.P. Calvo, J. de Frutos and J. Novo, Linearly implicit Runge–Kutta methods for advection–reaction–diffusion equations. *Appl. Numer. Math.* **37** (2001) 535–549.
[8] Y. Cheng and C.-W. Shu, A discontinuous Galerkin finite element method for time dependent partial differential equations with higher order derivatives. *Math. Comput.* **77** (2008) 699–730.
[9] X. Chen and M. Gunzburger, Continuous and discontinuous finite element methods for a peridynamics model of mechanics. *Comput. Methods Appl. Mech. Eng.* **200** (2011) 1237–1250.
[10] B. Cockburn and C.-W. Shu, TVB Runge–Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws II: general framework. *Math. Comput.* **52** (1989) 411–435.
[11] B. Cockburn and C.-W. Shu, The Runge–Kutta local projection P^1-discontinuous-Galerkin finite element method for scalar conservation laws. *ESAIM:M2AN* **25** (1991) 337–361.
[12] B. Cockburn and C.-W. Shu, The Runge–Kutta discontinuous Galerkin finite element method for conservation laws V: multidimensional systems. *J. Comput. Phys.* **141** (1998) 199–224.
[13] B. Cockburn and C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection–diffusion systems. *SIAM J. Numer. Anal.* **35** (1998) 2440–2463.
[14] B. Cockburn, S.-Y. Lin and C.-W. Shu, TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one dimensional systems. *J. Comput. Phys.* **84** (1989) 90–113.
[15] B. Cockburn, S. Hou and C.-W. Shu, TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: the multidimensional case. *Math. Comput.* **54** (1990) 545–581.
[16] B. Cockburn, J. Gopalakrishnan and R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.* **47** (2009) 1319–1365.
[17] B. Cockburn, J. Guzmán, S.-C. Soon and H.K. Stolarski, An analysis of the embedded discontinuous Galerkin method for second-order elliptic problems. *SIAM J. Numer. Anal.* **47** (2009) 2686–2707.
[18] M. D'Elia, Q. Du, C. Glusa, M. Gunzburger, X. Tian and Z. Zhou, Numerical methods for nonlocal and fractional models. *Acta Numer.* **29** (2020) 1–124.
[19] M. D'Elia, C. Flores, X. Li, P. Radu and Y. Yu, Helmholtz–Hodge decompositions in the nonlocal framework: well-posedness analysis and applications. *J. Peridyn. Nonlocal Model.* **2** (2020) 401–418.
[20] M. D'Elia, X. Li, P. Seleson, X. Tian and Y. Yu, A review of local-to-nonlocal coupling methods in nonlocal diffusion and nonlocal mechanics. *J. Peridyn. Nonlocal Model.* **4** (2022) 1–50.
[21] W. Deng and J.S. Hesthaven, Local discontinuous Galerkin methods for fractional diffusion equations. *ESAIM:M2AN* **47** (2013) 1845–1864.
[22] J. Douglas and T. Dupont, Interior penalty procedures for elliptic and parabolic galerkin methods. In Vol. 58 *Lecture Notes in Physics*. Springer, Berlin (1976).
[23] Q. Du, Nonlocal modeling, analysis, and computation. In Vol. 94 *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2019).
[24] Q. Du and K. Zhou, Mathematical analysis for the peridynamic nonlocal continuum theory. *ESAIM:M2AN* **45** (2011) 217–234.
[25] Q. Du and J. Yang, Asymptotically compatible Fourier spectral approximations of nonlocal Allen–Cahn equations. *SIAM J. Numer. Anal.* **54** (2016) 1899–1919.
[26] Q. Du and X. Tian, Mathematics of smoothed particle hydrodynamics: a study via nonlocal Stokes equations. *Found. Comput. Math.* **20** (2020) 801–826.
[27] Q. Du, M. Gunzburger, R.B. Lehoucq and K. Zhou, Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Rev.* **54** (2012) 667–696.
[28] Q. Du, M. Gunzburger, R. Lehoucq and K. Zhou, A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. *Math. Model. Methods Appl. Sci.* **23** (2013) 493–540.
[29] Q. Du, Z. Huang and R. Lehoucq, Nonlocal convection–diffusion volume-constrained problems and jump processes. *Discrete Contin. Dyn. Syst. B* **19** (2014) 373–389.
[30] Q. Du, Z. Huang and P.G. LeFloch, Nonlocal conservation laws. A new class of monotonicity-preserving models. *SIAM J. Numer. Anal.* **55** (2017) 2465–2489.
[31] Q. Du, J. Zhang and C. Zheng, Nonlocal wave propagation in unbounded multi-scale media. *Commun. Comput. Phys.* **24** (2018) 1049–1072.

[32] Q. Du, X.H. Li, J. Lu and X. Tian, A quasi-nonlocal coupling method for nonlocal and local diffusion models. *SIAM J. Numer. Anal.* **56** (2018) 1386–1404.

[33] Q. Du, L. Ju, X. Li and Z. Qiao, Maximum principle preserving exponential time differencing schemes for the nonlocal Allen–Cahn equation. *SIAM J. Numer. Anal.* **57** (2019) 875–898.

[34] Q. Du, L. Ju and J. Lu, A discontinuous Galerkin method for one-dimensional time-dependent nonlocal diffusion problems. *Math. Comput.* **88** (2019) 123 –147.

[35] Q. Du, L. Ju, J. Lu and X. Tian, A discontinuous Galerkin method with penalty for one-dimensional nonlocal diffusion problems. *Commun. Appl. Math. Comput.* **2** (2020) 31–55.

[36] Q. Du, X. Tian, C. Wright and Y. Yu, Nonlocal trace spaces and extension results for nonlocal calculus. *J. Funct. Anal.* **282** (2022) No. 109453.

[37] G. Gilboa and S. Osher, Nonlocal linear image regularization and supervised segmentation. *Multiscale Model. Simul.* **6** (2007) 595–630.

[38] G. Gilboa and S. Osher, Nonlocal operators with applications to image processing. *Multiscale Model. Simul.* **7** (2008) 1005–1028.

[39] Q. Guan and M. Gunzburger, Stability and accuracy of time-stepping schemes and dispersion relations for a nonlocal wave equation. *Numer. Method Partial Differ. Equ.* **31** (2015) 500–516.

[40] M. Gunzburger and R. Lehoucq, A nonlocal vector calculus with application to nonlocal boundary value problems. *Multiscale Model. Simul.* **8** (2010) 1581–1598.

[41] K. Huang and Q. Du, Stability of a nonlocal traffic flow model for connected vehicles. *SIAM J. Appl. Math.* **82** (2022) 221–243.

[42] H. Liu and J. Yan, The direct discontinuous Galerkin (DDG) methods for diffusion problems. *SIAM J. Numer. Anal.* **47** (2008/2009) 675–698.

[43] T. Mengesha and Q. Du, The bond-based peridynamic system with Dirichlet-type volume constraint. *Proc. R. Soc. Edinb. Sect. A* **144** (2014) 161–186.

[44] L. Qiu, W. Deng and J.S. Hesthaven, Nodal discontinuous Galerkin methods for fractional diffusion equations on 2D domain with triangular meshes. *J. Comput. Phys.* **298** (2015) 678–694.

[45] W.H. Reed and T.R. Hill, *Triangular mesh methods for the neutron transport equation*, Los Alamos Scientific Laboratory report LA-UR-73-479, NM (1973).

[46] B. Rivière, M. F. Wheeler and V. Girault, Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems I. *Comput. Geosci.* **3** (1999) 337–360.

[47] C.-W. Shu, Total-variation-diminishing time discretizations. *SIAM J. Sci. Statist. Comput.* **9** (1988) 1073–1084.

[48] C.-W. Shu and S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.* **77** (1988) 439–471.

[49] S.A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids* **48** (2000) 175–209.

[50] S.A. Silling and R.B. Lehoucq, Peridynamic theory of solid mechanics. *Adv. Appl. Mech.* **44** (2010) 73–168.

[51] C.-W. Shu, Discontinuous Galerkin methods: general approach and stability, edited by S. Bertoluzza, S. Falletta, G. Russo and C.-W. Shu. Birkhauser. In: *Numerical Solutions of Partial Differential Equations*. Advanced Courses in Mathematics CRM Barcelona, Basel (2009) 149–201.

[52] X. Tian and Q. Du, Analysis and comparison of different approximations to nonlocal diffusion and linear peridynamic equations. *SIAM J. Numer. Anal.* **51** (2013) 3458–3482.

[53] X. Tian and Q. Du, Asymptotically compatible schemes and applications to robust discretization of nonlocal models. *SIAM J. Numer. Anal.* **52** (2014) 1641–1665.

[54] X. Tian and Q. Du, Nonconforming discontinuous Galerkin methods for nonlocal variational problems. *SIAM J. Numer. Anal.* **53** (2015) 762–781.

[55] X. Tian and Q. Du, Trace theorems for some nonlocal function spaces with heterogeneous localization. *SIAM J. Math. Anal.* **49** (2017) 1621–1644.

[56] H. Tian, L. Ju and Q. Du, Nonlocal convection–diffusion problems and finite element approximations. *Comput. Methods Appl. Mech. Eng.* **289** (2015) 60–78.

[57] X. Tian and Q. Du, Asymptotically compatible schemes for robust discretization of parametrized problems with applications to nonlocal models. *SIAM Rev.* **62** (2020) 199–227.

[58] B. van Leer and S. Nomura, Discontinuous Galerkin for diffusion. In: *17th AIAA Computational Fluid Dynamics Conference* (2005) 5108.

[59] J. Wang and X. Ye, A weak Galerkin finite element method for second-order elliptic problems. *J. Comput. Appl. Math.* **241** (2013) 103–115.

[60] Z. Wang, Q. Tang, W. Guo and Y. Cheng, Sparse grid discontinuous Galerkin methods for high-dimensional elliptic equations. *J. Comput. Phys.* **314** (2016) 244–263.

[61] Q. Xu and J.S. Hesthaven, Discontinuous Galerkin method for fractional convection–diffusion equations. *SIAM J. Numer. Anal.* **52** (2014) 405–423.