Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2016-0226

Vol. **23**, No. 1, pp. 168-197
January 2018

Error Estimates for the Iterative Discontinuous Galerkin Method to the Nonlinear Poisson-Boltzmann Equation

Peimeng Yin[1], Yunqing Huang[2] and Hailiang Liu[1,*]

[1] *Mathematics Department, Iowa State University, Ames, IA 50011, USA.*
[2] *Key Laboratory of Intelligent Computing & Information Processing of Ministry of Education; Hunan Key Laboratory for Computation and Simulation in Science and Engineering; School of Mathematics and Computational Science, Xiangtan University, Xiangtan, 411105, P.R. China.*

Communicated by Chi-Wang Shu

Received 28 November 2016; Accepted (in revised version) 29 March 2017

Abstract. This paper is devoted to the error estimate for the iterative discontinuous Galerkin (IDG) method introduced in [P. Yin, Y. Huang and H. Liu. *Commun. Comput. Phys.* 16: 491–515, 2014] to the nonlinear Poisson-Boltzmann equation. The total error includes both the iteration error and the discretization error of the direct DG method to linear elliptic equations. For the DDG method, the energy error is obtained by a constructive approach through an explicit global projection satisfying interface conditions dictated by the choice of numerical fluxes. The L^2 error of order $\mathcal{O}(h^{m+1})$ for polynomials of degree m is further recovered. The bounding constant is also shown to be independent of the iteration times. Numerical tests are given to validate the established convergence theory.

AMS subject classifications: 65D15, 65N30, 35J05, 35J25

Key words: Poisson-Boltzmann equation, DG methods, global projection, energy error estimates, L^2 error estimates.

1 Introduction

This paper is devoted to the error estimate for the iterative discontinuous Galerkin (IDG) schemes, introduced in [36] to solve the boundary value problem for the nonlinear Poisson-Boltzmann (PB) equation,

$$-\lambda^2\Delta u = f(x) + e^{-u} \quad \text{in } \Omega \subset \mathbb{R}^d, \tag{1.1a}$$

$$u = g(x) \quad \text{on } \partial\Omega. \tag{1.1b}$$

*Corresponding author. *Email addresses:* pemyin@iastate.edu (P. Yin), huangyq@xtu.edu.cn (H. Huang), hliu@iastate.edu (H. Liu)

In this model equation, u is the unknown in the bounded domain Ω, with f and g given; $\lambda>0$ is a physical parameter, called scaled Debye length. In many physically relevant settings, this parameter is very small [16]. This PB equation appears in many applications, including semiconductor modeling [13,30] and charged particles in solutions [15,17].

There are two main challenges in numerically solving the PB problem (1.1), one is nonlinearity of the model, another is smallness of the parameter $\lambda \ll 1$. Resolution of the former requires some iteration techniques, instead of a direct discretization by standard methods; for the latter, one needs to properly choose initial guess for the iteration to converge. These two issues have been properly resolved by the IDG algorithm introduced in [36]. Such an algorithm involves two steps: (i) the nonlinear PB equation is iteratively approximated by a series of linear PB equations, and (ii) each linear PB equation is solved by the direct Discontinuous Galerkin (DDG) method following [23,24]. As illustrated in [36], the iterative DG method has linear complexity in terms of the degree of freedom even for small λ. Also, $(m+1)$th order of accuracy for P^m polynomials was numerically observed in [36]. This work aims to obtain the optimal L^2 error for the IDG method rigorously.

Our main result may be stated as follows: for smooth solution u to the nonlinear PB equation (1.1), let u_h^n be the numerical solution to the linearized PB equation at step n, generated by the DDG method using polynomials of degree m over computational cells of size h. For some appropriate iterative step and initial guess u^0 for the iteration, there exists $0<\mu<1$ such that the following estimate holds,

$$\|u_h^n-u\|_{L^2(\Omega)} \le \mu^n \|u^0-u\|_{L^2(\Omega)} + Ch^{m+1},$$

for some constant C independent of h and n. The idea to obtain this estimate is to split the error into iteration error $\|u^n-u\|$ and discretization error $\|u^n-u_h^n\|$ of the DDG method for the linearized PB equation with solution u^n at $n-$th step. The iteration error

$$\|u^n-u\|_{L^2(\Omega)} \le \mu^n \|u^0-u\|_{L^2(\Omega)}$$

was already obtained in [36], hence the main task of this work is to estimate $\|u^n-u_h^n\| \le Ch^{m+1}|u^n|_{m+1}$ for each linearized PB equation, and uniform boundedness of $|u^n|_{m+1}$ in terms of n.

Due to the nonlinearity of the problem, bounding $|u^n|_{m+1}$ uniformly in n is more involved. There are three key ingredients in our analysis: (i) the elliptic regularity gives $\|u^{n+1}\|_{s+1} \le \|F\|_s$ with F involving u^n, f and the nonlinear term e^{-u^n}, see Lemma 3.1; (ii) the Moser-type estimate is used to bound $|e^{-u^n}|_s$ by $|u^n|_s$ and $\|e^{-u^n}\|_{L^\infty}$, see Lemma 3.2; and (iii) the point-wise bound in (2.7) for the iterative solutions is used to bound $\|u^n\|_{L^\infty}$ by $\max\{\|u^0\|_{L^\infty}, \|u\|_{L^\infty}\}$ as shown in (3.10). These together suffice to deduce the desired uniform bound of $|u^n|_{m+1}$.

Obtaining error estimates for various DG methods has been a main subject of research. For the linear Poisson equation with the Dirichlet boundary condition, a unified analysis was presented in [1] to obtain the best possible error estimates for a class of existing DG methods, including the optimal L^2 estimates for the method of Babuska [5], the

IPDG method [12], the method of Bassi-Reby [4], the LDG method [10], and the method of Brezzi et al. [6] due to their consistency and stability, as well as the sub-optimal L^2 estimates for the method of Bauman-Oden [3] and the NIPG method [33] due to certain inconsistency. For the IPDG method to linearized PB equations, the techniques in [1] can well be carried over; yet the optimal L^2 error given in [2] is, instead, concluded from some results in [32].

Note that the numerical flux in the DDG method involves interface jumps in second derivatives, analysis in [1] is not applicable in straightforward manner. For the DDG method applied to diffusion equations with periodic boundary condition, a novel global projection is introduced in [21] to estimate the optimal L^2 error. In this work we modify the projection in [21] to match with the given Dirichlet boundary condition. Such a modified projection may not be optimal, yet it is sufficient for us to obtain the optimal error in the energy norm. Recovery of the optimal error in L^2 follows from the usual duality-based lift technique, which is valid for the DDG method if the exact solution $u \in H^s$ for $s \geq 3$. For multi-dimensional non-structure meshes, the explicit projection is not available, we refer to [22] for techniques exploiting certain implicit projection defined by the DDG method on the associated elliptic problem. The projection error in [22] is obtained by refining the analysis in [1]. In contrast, the explicit projection essentially used in our constructive approach also gives an alternative approximation of the original elliptic problem.

The discontinuous Galerkin (DG) method we discuss in this paper is a class of finite element methods, using a completely discontinuous piecewise polynomial space for the numerical solution and the test functions. One main advantage of the DG method is the flexibility afforded by local approximation spaces combined with the suitable design of numerical fluxes crossing cell interfaces. More general information about DG methods for elliptic, parabolic, and hyperbolic PDEs can be found in the recent books and lecture notes, see, e.g., [19, 32, 34]. In particular, we refer to [2, 14, 36] for some DG methods applied to the PB equations. The idea of DDG methods for second order PDEs is to directly force the weak solution formulation of the PDE into the DG function space for both the numerical solution and test functions. A key feature in the DDG schemes proposed in [23, 24] lies in numerical flux choices for the solution gradient, which involves higher order derivatives evaluated crossing interfaces. The DDG method has been proven to be optimally accurate [21] and superconvergent [7] when applied to time-dependent diffusion equations [21, 23, 24]. DDG methods for second order linear elliptic equations were first studied in [18], and further extended to solve nonlinear Poisson-Boltzmann equation in [36]. This work is to develop convergence theory for the DDG method for elliptic equations [18, 36]. The DDG method has been extended to various applications from Fokker-Planck type equations [25–28] in biophysics to fluid equations such as the compressible Navier-Stokes equation [11].

The rest of the paper is organized as follows: in Section 2, we describe the formulation of the iterative DG methods and present the main results of this paper. In Section 3, we reformulate the DDG method for the model problem and construct a global projection.

Based on the projection, we obtain the error estimates of the DDG method in energy norm. In Section 4, the optimal L^2 error estimates of the DDG method is recovered from its energy norm. Numerical tests are given in Section 5 to examine the optimal error estimate of the DDG method and the IDG method. In Section 6, concluding remarks are given. In Appendix A, we present a detailed proof of Theorem 4.1 in Section 3, and part of the facts is shown in Appendix B.

Throughout this paper, we adopt standard notations for Sobolev spaces such as $H^s(D)$ on sub-domain $D\subset\Omega$ equipped with the norm $\|\cdot\|_{s,D}$ and semi-norm $|\cdot|_{s,D}$. When $D=\Omega$, we omit the index D, and when $s=0$ we omit the index $s=0$ too. We use the notation $A\lesssim B$ to indicate that A can be bounded by B multiplied by a constant independent of the mesh size. $A\sim B$ stands for $A\lesssim B$ and $B\lesssim A$.

2 The IDG scheme and main results

Recall that the following iteration was introduced in [36]: starting with an initial guess u^0, find $u^n(n=1,2,\cdots)$ iteratively by solving the linearized PB equation

$$\begin{cases} -\lambda^2\Delta u^{n+1}+k^n u^{n+1}=k^n u^n+f(x)+e^{-u^n} & \text{in } \Omega,\\ u^{n+1}=g(x) & \text{on } \partial\Omega, \end{cases} \tag{2.1}$$

where for the purpose of error estimate $g(x)$ and $f(x)$ are assumed as smooth as needed, and $k^n=e^{-\operatorname{essinf}(u^n)}$ is a properly chosen step parameter to enforce the convergence of the iteration.

Depending on the size of λ, there are two cases to consider for choosing the initial guess:

(i) If $\lambda=\mathcal{O}(1)$, or $\lambda\ll 1$ but $f(x)\geq 0$ for all $x\in\Omega$, we take

$$u^0=w, \tag{2.2}$$

where w solves

$$\begin{cases} -\lambda^2\Delta w=f(x) & \text{in } \Omega,\\ w=g(x) & \text{on } \partial\Omega. \end{cases} \tag{2.3}$$

(ii) If $\lambda\ll 1$, and there exists $x_0\in\Omega$, such that $f(x_0)<0$, we take

$$u^0=\min\left\{\ln\left(\frac{1}{\operatorname{esssup}(-f_0)}\right),\operatorname{essinf}(g(x))\right\},\quad x\in\Omega, \tag{2.4}$$

where $f_0=\frac{f-|f|}{2}$. Effectiveness of these choices has been numerically verified in [36].

2.1 Convergence rate of the iteration

Following [8], we define

$$M=\{v|\ v\in H^1(\Omega),\ e^{-v}\in L^\infty(\Omega),\ v=g(x) \text{ on } \partial\Omega\}.$$

The weak solution formulation of (1.1) is to find $u\in M$, such that

$$B(u,v)=(f(x),v)+(e^{-u},v),\ \ \forall v\in H_0^1(\Omega), \tag{2.5}$$

where $(u,v)=\int_\Omega uvdx$ denotes the inner product in the L^2 space, and

$$B(u,v)=(\lambda^2\nabla u,\nabla v)$$

denotes the bilinear operator generated from the diffusion. Correspondingly, the weak solution formulation of (2.1) is: from $u^n\in M$, we find $u^{n+1}\in M$ such that

$$B(u^{n+1},v)+k^n(u^{n+1},v)=k^n(u^n,v)+(f(x),v)+(e^{-u^n},v),\ \ \forall v\in H_0^1(\Omega), \tag{2.6}$$

where u^0 is the chosen initial guess. For each $u^n\in M$, (2.6) is shown in [36] to admit a unique solution $u^{n+1}\in M$.

Regarding the convergence rate of the solution sequence $\{u^n\}$ to u, we obtained the following result in [36].

Theorem 2.1 (Convergence rate)**.** *The solution sequence* $\{u^n\},(n=0,1,\cdots)$ *of* (2.6) *converges to the solution* u *of* (2.5) *in* M *and satisfies*

$$u^0\leqslant u^1\leqslant\cdots\leqslant u^n\leqslant u^{n+1}\leqslant\cdots\leqslant u\quad in\ \Omega\quad a.e. \tag{2.7}$$

Moreover,

$$\|u^n-u\|\leq\mu^n\|u^0-u\|,$$

where

$$0<\mu=\frac{\alpha}{\sqrt{1+\frac{2C_\Omega\lambda^2}{k^0}}}<\alpha$$

and C_Ω *is a constant depending on* Ω*, with*

$$\alpha=1-\frac{e^{-\operatorname{esssup}u}}{k^0}<1.$$

The iteration (2.1) is solved in [36] by the direct discontinuous Galerkin (DDG) method. We present the scheme here for the one dimensional case. To discretize the iterative weak formulation, we partition the domain $\Omega=[a,b]$ into computational elements $I_j=(x_{j-1/2},x_{j+1/2})$, $h=x_{j+1/2}-x_{j-1/2}$, with $x_{1/2}=a$ and $x_{N+1/2}=b$. We denote by v^+ and v^- the right and left limits of function v at $x_{j+1/2}$, and define

$$[v]_{j+1/2}=v^+_{j+1/2}-v^-_{j+1/2},\quad \{v\}=\frac{v^+_{j+1/2}+v^-_{j+1/2}}{2}.$$

Define a discontinuous Galerkin finite element space

$$V_h=\{v\in L^2(\Omega):v|_{I_j}\in\mathbb{P}^m(I_j),j=1,2,\cdots,N\},$$

where $\mathbb{P}^m(I_j)$ denotes the set of polynomials of degree no more than m on I_j.

In one dimensional case, (2.1) becomes

$$\begin{cases} -\lambda^2 u_{xx}^{n+1}+k^n u^{n+1}=k^n u^n+f(x)+e^{-u^n} \quad \text{in } \Omega, \\ u^{n+1}(a)=g_1, \\ u^{n+1}(b)=g_2. \end{cases} \tag{2.8}$$

The DDG method for (2.8) is to find $u_h^{n+1}\in V_h$ such that for all $v\in V_h$, $j=1,\cdots,N$,

$$\lambda^2\int_{I_j} u_{hx}^{n+1}v_x dx+k^n\int_{I_j} u_h^{n+1}v dx+\lambda^2\left(-\widehat{(u_{hx}^{n+1})}v+\widehat{(u_h^{n+1}-u_h^{n+1})}v_x\right)\Big|_{\partial I_j}=\int_{I_j} F^n v dx, \tag{2.9}$$

where $F^n=k^n u_h^n+f(x)+e^{-u_h^n}$, $k^n=e^{-\min(u^n)}$ and $v|_{\partial I_j}=v_{j+1/2}^- - v_{j-1/2}^+$. The numerical flux for (2.9) at interior interfaces is given by

$$\widehat{u_{hx}}=\beta_0\frac{[u_h]}{h}+\{u_{hx}\}+\beta_1 h[u_{hxx}], \quad \widehat{u}_h=\{u_h\}, \tag{2.10}$$

where (β_0,β_1) are the admissible parameters. On the boundary $x_{1/2}$ and $x_{N+1/2}$, the numerical flux has the following form

$$\text{at } x_{1/2}, \qquad \widehat{u_{hx}}=\beta_0\frac{u_h^+-g_1}{h}+u_{hx}^+, \quad \widehat{u}_h=g_1, \tag{2.11a}$$

$$\text{at } x_{N+1/2}, \qquad \widehat{u_{hx}}=\beta_0\frac{g_2-u_h^-}{h}+u_{hx}^-, \quad \widehat{u}_h=g_2. \tag{2.11b}$$

Remark 2.1. Numerical tests indicate that the boundary numerical flux of the following form

$$\text{at } x_{1/2}, \qquad \widehat{u_{hx}}=\beta_e\frac{u_h^+-g_1}{h}+u_{hx}^+, \quad \widehat{u}_h=(1-\nu)u_h^++\nu g_1, \tag{2.12a}$$

$$\text{at } x_{N+1/2}, \qquad \widehat{u_{hx}}=\beta_e\frac{g_2-u_h^-}{h}+u_{hx}^-, \quad \widehat{u}_h=(1-\nu)u_h^-+\nu g_2, \tag{2.12b}$$

where $\nu\in[0,1]$ is also admissible. The DDG scheme adopted in [36] corresponds to the case $\nu=\frac{1}{2}$, that is,

$$\widehat{u}_h|_{x_{1/2}}=\frac{u_h^++g_1}{2}, \quad \widehat{u}_h|_{x_{N+1/2}}=\frac{u_h^-+g_2}{2},$$

for which the scheme is well-posed for

$$\beta_0>\Gamma(\beta_1), \quad \beta_e>\frac{9}{8}m^2. \tag{2.13}$$

With the class of boundary fluxes in (2.12), a similar verification shows that the scheme is well-posed if

$$\beta_0>\Gamma(\beta_1), \quad \beta_e>\frac{(1+\nu)^2}{2}m^2, \tag{2.14}$$

where

$$\Gamma(\beta_1)=\sup_{v\in P^{m-1}([-1,1])}\frac{(v(1)-2\beta_1\partial_\xi v(1))^2}{\frac{1}{2}\int_{-1}^{1}v^2(\xi)d\xi}. \tag{2.15}$$

Remark 2.2. A quantitative estimate for $\Gamma(\beta_1)$ is presented in [21] as

$$\Gamma(\beta_1)=m^2\left(1-\beta_1(m^2-1)+\frac{\beta_1^2}{3}(m^2-1)^2\right),\quad m\geq 1. \tag{2.16}$$

For the stability of the DDG scheme with numerical fluxes (2.10) and (2.11) it suffices to select $\beta_0>\beta_0^*$, where

$$\beta_0^*=\max\{\Gamma(\beta_1),2m^2\}. \tag{2.17}$$

The main result of this work is as follows:

Theorem 2.2. *If $\beta_0>\beta_0^*$, then the solution sequence $u_h^n\in V_h$ converges to the smooth solution u^n as mesh is refined. Moreover,*

$$\|u_h^n-u^n\|\leq Ch^{m+1}|u^n|_{m+1},$$

where C is independent of h and n.

Remark 2.3. Depending on different cases where u^0 is either taken as in (2.4) or obtained by finding a unique $u^0\in M$ such that

$$B(u^0,v)=(f(x),v),\quad \forall v\in H_0^1(\Omega). \tag{2.18}$$

In the latter case, u_h^0 is obtained from solving (2.18) by the same DDG method.

Once this is achieved, the triangle inequality leads to the error estimate for the IDG method,

$$\|u_h^n-u\|\leq\|u_h^n-u^n\|+\|u^n-u\|\leq Ch^{m+1}|u^n|_{m+1}+\mu^n\|u^0-u\|.$$

The remaining task is to estimate the error for the DG discretization, i.e., to prove Theorem 2.2, and to show that $|u^n|_{m+1}$ is uniformly bounded in n. It suffices to prove the latter for large enough n. Such a result can be summarized in the following.

Theorem 2.3. *If $f\in H^s(\Omega)$, $s\geq m-1$, then the solution to* (2.8) *with large n admits the following estimate*

$$|u^n|_{s+2}\leq C,$$

where C may depend on $\|u^0\|_{L^\infty},\|u\|_{L^\infty},\|f\|_s$, $g_i(i=1,2)$, or the parameter λ, but independent of n.

Remark 2.4. This result is also valid in multi-dimensional setting on rectangular grids. Yet for simplicity of presentation the proof in Section 3 is only given for one-dimensional case, for which bounding constants in the estimate can be made precise.

2.2 Scheme reformulation

In order to prove Theorems 2.2 and 2.3, we set

$$\begin{cases} U(x)=u^{n+1}(x)-\left(g_1\frac{x-b}{a-b}+g_2\frac{x-a}{b-a}\right), \\ k=\frac{k^n}{\lambda^2}, \\ F(x)=ku^n+\frac{1}{\lambda^2}\left(f(x)+e^{-u^n}\right)-k\left(g_1\frac{x-b}{a-b}+g_2\frac{x-a}{b-a}\right), \end{cases} \tag{2.19}$$

then (2.8) reduces to

$$\begin{cases} -U_{xx}+kU=F(x) \quad \text{in } \Omega, \\ U(a)=0, \quad U(b)=0. \end{cases} \tag{2.20}$$

Summation of the DDG scheme (2.9) over each computational cell $I_j, j=1,2,\cdots,N$, by (2.19) gives a global DDG formulation for (2.20): find $u_h\in V_h$ such that

$$A(u_h,v)=\int_\Omega Fvdx, \quad \forall v\in V_h, \tag{2.21}$$

where

$$A(u_h,v)=\sum_{j=1}^{N}\int_{I_j}u_{hx}v_xdx+k\int_\Omega u_hvdx+\sum_{j=0}^{N}(\widehat{u_{hx}}[v]+\{v_x\}[u_h])_{j+1/2}, \tag{2.22}$$

the numerical flux in (2.22) is defined as the same as (2.10) for all interior cell interfaces. On the boundary $x_{1/2}$ and $x_{N+1/2}$, the numerical flux has the following form

$$\widehat{u_{hx}}_{1/2}=\beta_0\frac{[u_h]_{1/2}}{h}+(u_{hx})^+_{1/2}, \quad [u_h]_{1/2}=(u_h)^+_{1/2}, \tag{2.23a}$$

$$\widehat{u_{hx}}_{N+1/2}=\beta_0\frac{[u_h]_{N+1/2}}{h}+(u_{hx})^-_{N+1/2}, \quad [u_h]_{N+1/2}=-u^-_{h(N+1/2)}. \tag{2.23b}$$

For the test function v, we abuse the notation as follows

$$[v]_{1/2}=v^+_{1/2}, \quad [v]_{N+1/2}=-v^-_{N+1/2}, \tag{2.24a}$$

$$\{v_x\}_{1/2}=(v_x)^+_{1/2}, \quad \{v_x\}_{N+1/2}=(v_x)^-_{N+1/2}, \tag{2.24b}$$

so to have a compact formulation in (2.22). For smooth solution $U\in H^{s+1}(\Omega)(s\geq m\geq 1)$ to (2.20), the DDG method (2.21) can be shown to be consistent with (2.20) in the sense that

$$A(U,v)=\int_\Omega Fvdx, \quad \forall v\in V_h. \tag{2.25}$$

This when combined with (2.21) yields the Galerkin orthogonality:

$$A(U-u_h,v)=0, \quad \forall v\in V_h. \tag{2.26}$$

For any function $v\in V_h$, it has been proved in [36] for $\beta_0>\beta_0^*$,

$$A(v,v)\geq\gamma\|v\|_E^2+k\|v\|^2,\quad\forall v\in V_h, \tag{2.27}$$

where $\gamma\in(0,1)$ is some positive constant and

$$\|v\|_E^2=\sum_{j=1}^{N}\int_{I_j}|v_x|^2dx+\frac{\beta_0}{h}\sum_{j=0}^{N}[v]_{j+1/2}^2. \tag{2.28}$$

3 Uniform boundedness of $|u^n|_{s+2}$

In order to prove Theorem 2.3, we first prepare the following two lemmas.

Lemma 3.1. *Let U be the solution of* (2.20) *with $F\in H^s$ for any $s\geq 0$, then*

$$\|U\|\leq\frac{(b-a)^2}{\pi^2}\|F\|,\quad\|U_x\|\leq\frac{b-a}{\pi}\|F\|. \tag{3.1}$$

Moreover, for $C_s=1+k+\cdots+k^{\left[\frac{s}{2}\right]}+(s-2[s/2])(b-a)k^{\left[\frac{s}{2}\right]+1}/\pi$, we have

$$|U|_{s+2}\leq C_s\|F\|_s. \tag{3.2}$$

Proof. By Poincare's inequality we have $\|U\|\leq\frac{b-a}{\pi}\|U_x\|$; integration of (2.20) against U gives

$$\|U_x\|^2+k\|U\|^2=\int_a^b FUdx\leq\|F\|\|U\|, \tag{3.3}$$

these together yield (3.1). Taking square of (2.20) and integration gives

$$\|U_{xx}\|^2+2k\|U_x\|^2+k^2\|U\|^2=\|F\|^2, \tag{3.4}$$

hence $\|U_{xx}\|\leq\|F\|$. For $F\in H^s(\Omega)$ with $s>0$ we use $-\partial_x^{s+2}U=\partial_x^sF-k\partial^sU$ to obtain recursively

$$\begin{aligned}\|\partial_x^{s+2}U\|&\leq\|\partial_x^sF\|+k\|\partial_x^sU\|\\&\leq(|F|_s+k|F|_{s-2}+\cdots+k^r|F|_{s-2r})+(s-2r)k^{r+1}\|U_x\|\quad(r=[s/2])\\&\leq(1+k+\cdots+k^r)\|F\|_s+(s-2r)(b-a)k^{r+1}/\pi\|F\|\\&\leq C_s\|F\|_s,\end{aligned}$$

where $\|U_{xx}\|\leq\|F\|$ has been used in the case when s is even. □

Lemma 3.2. *For $v(x)\in H^i(\Omega)\cap L^\infty(\Omega)$ with $i\geq 1$, we have*

$$|e^v|_i\leq C\|e^v\|_{L^\infty}\left(\|v\|_{L^\infty}^{i-1}+1\right)(|v|_i+\|v\|_{L^\infty}), \tag{3.5}$$

where C depends upon $b-a$ and i, but is independent of v.

This is the Moser-type calculus inequalities, see [29, Proposition 2.1] for functions defined in the whole space. For the case in a bounded domain, a modification is needed to yield a slightly different bound. In the one-dimensional case, we present a simple, self-contained proof as below.

Proof. By Faà di Bruno's formula

$$\partial_x^i g(v)=\sum \frac{i!}{l_1!l_2!\cdots l_i!}\partial_v^l g(v)\prod_{j=1}^{i}\left(\frac{\partial_x^j v}{j!}\right)^{l_j}, \tag{3.6}$$

where the Diophantine equation $\sum_{j=1}^{i} jl_j=i$ holds with $\sum_{j=1}^{i} l_j=l$. For $g(v)=e^v$, we have

$$|e^v|_i\le \|e^v\|_{L^\infty}\sum \frac{i!}{l_1!l_2!\cdots l_i!}\prod_{j\in J}\left(\frac{1}{j!}\right)^{l_j}\left\|\prod_{j\in J}\left(\partial_x^j v\right)^{l_j}\right\|, \tag{3.7}$$

where we have set $J:=\{j,1\le j\le i,l_j\neq 0\}$. By Hölder's inequality,

$$\left\|\prod_{j\in J}\left(\partial_x^j v\right)^{l_j}\right\|\le\prod_{j\in J}\left\|\left(\partial_x^j v\right)^{l_j}\right\|_{L^{p_j}}=\prod_{j\in J}\left\|\left(\partial_x^j v\right)\right\|_{L^{l_jp_j}}^{l_j}, \tag{3.8}$$

where $\sum_{j\in J}\frac{1}{p_j}=\frac{1}{2}$. From the celebrated Gagliardo-Nirenberg inequality when applied to a bounded interval Ω [31],

$$\left\|\left(\partial_x^j v\right)\right\|_{L^{l_jp_j}}\le C_1\|v\|_{L^\infty}^{1-\alpha_j}\|\partial_x^i v\|^{\alpha_j}+C_2\|v\|_{L^r}, \tag{3.9}$$

where $r>0$ is arbitrary, $\alpha_j\in(0,1)$ satisfy

$$\frac{1}{l_jp_j}-j=\left(\frac{1}{2}-i\right)\alpha_j,$$

so that $\sum_{j\in J}\alpha_j l_j=1$, with C_1,C_2 depending only on $|\Omega|=b-a$, j, i, l_j, p_j, it follows that

$$\left\|\left(\partial_x^j v\right)\right\|_{L^{l_jp_j}}\le C\|v\|_{L^\infty}^{1-\alpha_j}(\|\partial_x^i v\|^{\alpha_j}+\|v\|_{L^\infty}^{\alpha_j}).$$

This when combined with (3.8) yields

$$\begin{aligned}\left\|\prod_{j\in J}\left(\partial_x^j v\right)^{l_j}\right\|&\le C\prod_{j\in J}\|v\|_{L^\infty}^{l_j-l_j\alpha_j}\prod_{j\in J}(\|\partial_x^i v\|^{\alpha_j}+\|v\|_{L^\infty}^{\alpha_j})^{l_j}\\&\le C\|v\|_{L^\infty}^{l-1}\prod_{j\in J}2^{1-\alpha_j}(\|\partial_x^i v\|+\|v\|_{L^\infty})^{l_j\alpha_j}\\&\le C\prod_{j\in J}2^{1-\alpha_j}\|v\|_{L^\infty}^{l-1}(|v|_i+\|v\|_{L^\infty}),\end{aligned}$$

which upon insertion into (3.7) leads to the desired estimate. □

Proof of Theorem 2.3. Let c_i $(i=0,\cdots,3)$ denote constants which may depend on $\|u^0\|_{L^\infty}$, $\|u\|_{L^\infty}$, $\|f\|_s$, $g_i (i=1,2)$, and physical parameters, but independent of n. From (2.7) it follows that $k^n = e^{-essinf(u^n)} \le k^0$ and $k \le k^0/\lambda^2$, as well as

$$\|u^n\|_{L^\infty} \le \max\{\|u^0\|_{L^\infty}, \|u\|_{L^\infty}\}, \tag{3.10}$$

hence $\|u^n\|_{L^\infty} + k^n + C_s \le c_0$. Using the regularity estimate (3.2) we have

$$|u^{n+1}|_{s+2} = |U|_{s+2} \le C_s \|F\|_s. \tag{3.11}$$

Here the relation $U(x) = u^{n+1}(x) - \tilde{g}$ has been used, with

$$\tilde{g}(x) = g_1 \frac{x-b}{a-b} + g_2 \frac{x-a}{b-a}.$$

Note that for $s=0$,

$$|F|_0 = \|F\| \le \frac{1}{\lambda^2}\left(k^n \|u^n - \tilde{g}\| + \|f\| + \|e^{-u^n}\|\right) \le c_1.$$

For $s \ge 1$, using (3.5) we have

$$\begin{aligned} |F|_s &\le \frac{1}{\lambda^2}\left(k^n |u^n - \tilde{g}|_s + |f|_s + |e^{-u^n}|_s\right) \\ &\le \frac{1}{\lambda^2}\left(k^n |u^n - \tilde{g}|_s + |f|_s + \|e^{-u^n}\|_{L^\infty}\left(\|u^n\|_{L^\infty}^{s-1} + 1\right)\left(|u^n|_s + \|u^n\|_{L^\infty}\right)\right) \\ &\le c_2 (|u^n|_s + 1). \end{aligned}$$

This when inserted into (3.11) gives

$$\begin{aligned} \|u^{n+1}\|_{s+2} &\le c_3(\|u^n\|_s + 1) - 1 \\ &\le c_3^{[s/2]+1}(\|u^{n-[s/2]}\|_{s-2[s/2]} + 1) - 1. \end{aligned}$$

For s even or odd, the right hand side is always bounded by $\|u^{l+1}\|_1$ for some $l = n-1-[s/2]$. Note that

$$\|u^{l+1}\|_1 \le \|U\|_1 + \|\tilde{g}\|_1 \le C\|F\| + \|\tilde{g}\|_1, \tag{3.12}$$

which is also uniformly bounded. This completes the proof of Theorem 2.3.

4 Error estimates of the DDG method

With a bit abuse of the notation in Section 4.1 and Appendix A, u is also used to denote any given function, instead of only the solution of the original PDE problem (1.1).

4.1 Global projection

A key step in obtaining the desired estimate in Theorem 2.2 is the global projection P which we introduce below. For a given piecewise smooth function $u\in L^2(\Omega)$, we define $Pu\in V_h(m\geq 1)$ to satisfy the following relations,

$$\int_{I_j} Pu(x)v(x)=\int_{I_j} u(x)v(x),\quad \forall v\in P^{m-2}(I_j),\ j=1,\cdots,N, \tag{4.1a}$$

$$\widehat{(Pu)_x}=\widehat{u_x}\quad \text{at } x_{j+1/2} \text{ for } j=1,\cdots,N-1, \tag{4.1b}$$

$$\{Pu\}:=\{u\}\quad \text{at } x_{j+1/2} \text{ for } j=1,\cdots,N-1, \tag{4.1c}$$

$$Pu(x_{1/2}^+)=u(x_{1/2}^+),\quad Pu(x_{N+1/2}^-)=u(x_{N+1/2}^-). \tag{4.1d}$$

For smooth $u\in H^{s+1}(\Omega)$,$s\geq 2$, we have $\widehat{u_x}_{j+1/2}=u_x(x_{j+1/2})$ and $\{u\}_{j+1/2}=u(x_{j+1/2})$. Note that in case of $m=1$, the relation (4.1a) is redundant.

Lemma 4.1. *For (β_0,β_1) such that $\beta_0>\beta_0^*$, (4.1) admits a unique projection P.*

Proof. For given u, Pu solves a linear system, hence its existence is implied by uniqueness. It suffices to prove $Pu\equiv 0$ if we take $u=0$.

From (4.1a-c) with $u=0$ and integration by parts, it follows that for $w=Pu$,

$$\begin{aligned}0&=-\sum_{j=1}^{N}\int_{I_j} ww_{xx}dx+\sum_{j=1}^{N-1}\left(\widehat{w_x}[w]-\{w\}[w_x]\right)_{j+1/2}\\&=\sum_{j=1}^{N}\int_{I_j} w_xw_xdx+\sum_{j=1}^{N-1}\left(\widehat{w_x}[w]+\{w_x\}[w]\right)_{j+1/2}+(ww_x)_{1/2}^+-(ww_x)_{N+1/2}^-.\end{aligned} \tag{4.2}$$

From (4.1d) we see that $w(x_{1/2}^+)=0$ and $w(x_{N+1/2}^-)=0$, we can replace the last two terms by

$$(\hat{w}_x+\{w_x\})_{1/2}w_{1/2}^+-(\hat{w}_x+\{w_x\})_{N+1/2}(w)_{N+1/2}^-,$$

where $\hat{w}_x$ is defined as given in (2.23) for u_h and $\{w_x\}$ is defined as given in (2.24) for v, so that

$$\begin{aligned}0&=\sum_{j=1}^{N}\int_{I_j}|w_x|^2dx+\sum_{j=1}^{N-1}\left(\widehat{w_x}[w]+\{w_x\}[w]\right)_{j+1/2}\\&\quad+\widehat{w_x}w_{1/2}^+-\hat{w}w_{N+1/2}^-+w_{1/2}^+(w_x)_{1/2}^+-w_{N+1/2}^-(w_x)_{N+1/2}^-\\&\geq\gamma\left[\sum_{j=1}^{N}\int_{I_j}|w_x|^2dx+\frac{\beta_0}{h}\left(\sum_{j=1}^{N-1}[w]_{j+1/2}^2+(w_{1/2}^+)^2+(w_{N+1/2}^-)^2\right)\right]\\&=\gamma\|w\|_E^2,\end{aligned}$$

for some $\gamma\in(0,1)$, provided $\beta_0>\beta_0^*$, by (2.27). We thus conclude $w\equiv 0$. ∎

As for the projection defined in (4.1), we have the following error estimate.

Theorem 4.1 (The projection error estimate)**.** *Let P denote the one dimensional global projection* (4.1)*, then*

$$Pv=v, \quad \forall v\in V_h. \tag{4.3}$$

Moreover, if $u|_{I_j}\in H^{s+1}(I_j)$ for $j=1,2,\cdots,N$, then we have the following estimates

$$\sum_{j=1}^{N}\|\partial_x^p(Pu-u)\|_{0,I_j}^2\leq Ch^{2(\min\{s,m\}-p+1)}\sum_{j=1}^{N}|u|_{s+1,I_j}^2, \tag{4.4a}$$

$$\sum_{j=1}^{N}|\partial_x^p(Pu-u)|_{\infty,I_j}^2\leq Ch^{2(\min\{s,m\}-p)+1}\sum_{j=1}^{N}|u|_{s+1,I_j}^2, \tag{4.4b}$$

for any $0\leq p\leq\min\{s,m\}$, where C is independent of h.

Proof. For $\forall u\in V_h$, a direct verification shows that Pu when taken as u satisfies (4.1). Uniqueness result stated in Lemma 4.1 asserts that $Pu=u$, as long as $u\in V_h$. The detailed proof of estimates in (4.4) is deferred to Appendix A for completeness. □

4.2 Error in energy norm

Recall the following estimate

$$|w|_{\infty,I_j}^2\leq 2h^{-1}\|w\|_{0,I_j}^2+h\|w_x\|_{0,I_j}^2, \quad \forall w\in H^1(I_j), \tag{4.5}$$

which can be verified by a direct integration.

For the error in energy norm, we have

Theorem 4.2. *Let u_h be the numerical solution to the DDG scheme* (2.21) *with (β_0,β_1) satisfying $\beta_0>\beta_0^*$, and U be the smooth solution to problem* (2.20)*, then*

$$\|U-u_h\|_E\leq Ch^m|U|_{m+1}, \tag{4.6}$$

where C is independent of h and n.

Proof. By Galerkin orthogonality (2.26), we have

$$A(U-u_h,v)=0, \quad \forall v\in V_h, \tag{4.7}$$

which can be rewritten as

$$A(e,v)=A(\epsilon,v), \tag{4.8}$$

where

$$e=PU-u_h, \quad \epsilon=PU-U. \tag{4.9}$$

For ϵ, we use Theorem 4.1 to get

$$\begin{aligned}\|\epsilon\|_E^2 &= \sum_{j=1}^{N}\int_{I_j}|\epsilon_x|^2dx+\frac{\beta_0}{h}\sum_{j=0}^{N}[\epsilon]_{j+1/2}^2\\ &\le 4\left(\sum_{j=1}^{N}\|\epsilon_x\|_{I_j}^2+\frac{\beta_0}{h}\sum_{j=1}^{N}|\epsilon|_{\infty,I_j}^2\right)\le C_1h^{2m}|U|_{m+1}^2,\end{aligned}\tag{4.10}$$

where C_1 is independent of h and n.

On the other hand, a direct calculation of the right hand side of (4.8) gives

$$\begin{aligned}A(\epsilon,v) &= \sum_{j=1}^{N}\int_{I_j}\epsilon_xv_xdx+k\int_{\Omega}\epsilon vdx+\sum_{j=0}^{N}\left(\widehat{\epsilon_x}[v]+\{v_x\}[\epsilon]\right)_{j+1/2}\\ &= -\sum_{j=1}^{N}\int_{I_j}\epsilon v_{xx}dx+k\int_{\Omega}\epsilon vdx+\sum_{j=1}^{N-1}\left(\widehat{\epsilon_x}[v]-\{\epsilon\}[v_x]\right)_{j+1/2}\\ &\quad+\left[\left(\beta_0\frac{\epsilon^+}{h}+\epsilon_x^+\right)v_{1/2}^++\left(\beta_0\frac{\epsilon^-}{h}-\epsilon_x^-\right)v_{N+1/2}^-\right]\\ &= k\int_{\Omega}\epsilon vdx+\epsilon_x^+v_{1/2}^+-\epsilon_x^-v_{N+1/2}^-,\end{aligned}\tag{4.11}$$

where we have used the fact that

$$\int_{I_j}\epsilon v_{xx}dx=0,\quad v\in V_h,\quad j=1,2,\cdots,N,$$

and at $x_{j+1/2}$,$j=1,2\cdots N-1$,

$$\widehat{\epsilon_x}=\widehat{(PU)_x}-\widehat{U_x}=0,\quad \{\epsilon\}=\{PU\}-U=0.$$

At $x_{1/2}$, we have $\epsilon_{1/2}^+=PU(x_{1/2}^+)-U(x_{1/2}^+)=0$, and also at $x_{N+1/2}$, $\epsilon_{N+1/2}^-=0$.

Taking $v=e$ in (4.11) and using (2.27), we have

$$\begin{aligned}k\|e\|^2+\gamma\|e\|_E^2\le A(e,e) &= k\int_{\Omega}\epsilon edx+\epsilon_x^+e_{1/2}^+-\epsilon_x^-e_{N+1/2}^-\\ &\le k\|e\|^2+\frac{k}{4}\|\epsilon\|^2+\frac{\gamma\beta_0}{2h}\left((e_{1/2}^+)^2+(e_{N+1/2}^-)^2\right)\\ &\quad+\frac{h}{2\gamma\beta_0}\left[((\epsilon_x^+)_{1/2})^2+((\epsilon_x^-)_{N+1/2})^2\right]\\ &\le k\|e\|^2+\frac{\gamma}{2}\|e\|_E^2+\frac{k}{4}\|\epsilon\|^2+\frac{h}{2\gamma\beta_0}\left[((\epsilon_x^+)_{1/2})^2+((\epsilon_x^-)_{N+1/2})^2\right].\end{aligned}$$

By the estimates stated in Theorem 4.1, we have

$$\begin{aligned}\frac{\gamma}{2}\|e\|_E^2 &\le \frac{k}{4}\|\epsilon\|^2+\frac{h}{2\gamma\beta_0}\left[((\epsilon_x^+)_{1/2})^2+((\epsilon_x^-)_{N+1/2})^2\right]\\ &\le \frac{k^0}{4\lambda^2}\|\epsilon\|^2+\frac{h}{2\gamma\beta_0}\left[|\epsilon_x|_{\infty,I_1}^2+|\epsilon_x|_{\infty,I_N}^2\right]\le C_2h^{2m}|U|_{m+1}^2.\end{aligned}\tag{4.12}$$

Hence, $\|e\|_E \leq 2C_2/\gamma h^m |U|_{m+1}$. This implies that

$$\|u-u_h\|_E \leq \|e\|_E + \|\epsilon\|_E \leq Ch^m |U|_{m+1},$$

where C is independent of h and n. □

4.3 The L^2-error estimate

In this section we recover the L^2-error estimate based on the error in energy norm using a "duality" argument.

Theorem 4.3. *Let u_h be the numerical solution to the DDG scheme* (2.21) *with $\beta_0 > \beta_0^*$, and U be the smooth solution to problem* (2.20)*, then*

$$\|U-u_h\| \leq Ch^{m+1}|U|_{m+1}, \tag{4.13}$$

where C is independent of h and n.

Proof. Consider the following problem

$$\begin{cases} -\psi_{xx}+k\psi = U-u_h & \text{in } \Omega, \\ \psi(a)=0, \quad \psi(b)=0. \end{cases} \tag{4.14}$$

Then $\psi \in H^2(\Omega)$ and there exists C such that

$$|\psi|_2 \leq C\|\theta\|, \quad \theta = U-u_h. \tag{4.15}$$

On the other hand,

$$\|\theta\|^2 = \int_a^b (-\psi_{xx}+k\psi)\theta dx.$$

Integration by parts on each cell gives

$$\begin{aligned} \|\theta\|^2 &= \sum_{j=1}^N \int_{I_j} \psi_x\theta_x dx + k\int_\Omega \psi\theta dx + \sum_{j=1}^{N-1} [\psi_x\theta] + (\psi_x\theta)^+_{1/2} - (\psi_x\theta)^-_{N+1/2} \\ &= \sum_{j=1}^N \int_{I_j} \psi_x\theta_x dx + k\int_\Omega \psi\theta dx + \sum_{j=0}^N \left(\widehat{\theta_x}[\psi] + \{\psi_x\}[\theta]\right)_{j+1/2} \\ &= A(\theta,\psi). \end{aligned}$$

Here we have used the fact $\psi_x = \{\psi_x\}, [\psi] = 0$. From (2.26), it follows that $A(\theta,v) = 0$, we proceed

$$\begin{aligned} \|\theta\|^2 &= A(\theta,\psi) - A(\theta,v) \\ &= \sum_{j=1}^N \int_{I_j} \theta_x(\psi-v)_x dx + k\int_\Omega \theta(\psi-v)dx \\ &\quad + \sum_{j=0}^N \left(\widehat{\theta_x}[\psi-v]\right)_{j+1/2} + \sum_{j=0}^N ([\theta]\{(\psi-v)_x\})_{j+1/2}. \end{aligned} \tag{4.16}$$

We choose v as the continuous piecewise linear interpolant of $\psi(x_{j+1/2})$ so that $[v]=0$ and the standard approximation result (see [9], Theorem 3.1.5)

$$\|\partial_x^q(\psi-v)\| \lesssim h^{2-q}|\psi|_2, \quad q=0,1. \tag{4.17}$$

Based on this we proceed to estimate each term in (4.16): for the first term, we have

$$\begin{aligned}\left|\sum_{j=1}^N \int_{I_j} \theta_x(\psi-v)_x dx\right| &\le \left(\sum_{j=1}^N \int_{I_j} |\theta_x|^2 dx\right)^{\frac{1}{2}} \left(\sum_{j=1}^N \int_{I_j} |(\psi-v)_x|^2 dx\right)^{\frac{1}{2}} \\ &\le \|\theta\|_E \|(\psi-v)_x\| \\ &\le Ch\|\theta\|_E |\psi|_2.\end{aligned}$$

The second term is bounded by

$$k\left|\int_\Omega \theta(\psi-v)dx\right| \le \frac{k^0}{\lambda^2}\left|\int_\Omega \theta(\psi-v)dx\right| \le Ch^2\|\theta\||\psi|_2,$$

where C is independent of h and n. For the last term, it follows

$$\begin{aligned}&\left|\sum_{j=0}^N ([\theta]\{(\psi-v)_x\})_{j+1/2}\right| \\ &\le \left(\frac{\beta_0}{h}\sum_{j=0}^N [\theta]_{j+1/2}^2\right)^{\frac{1}{2}} \left(\frac{h}{\beta_0}\right)^{\frac{1}{2}} \left(\sum_{j=0}^N \{(\psi-v)_x\}_{j+1/2}^2\right)^{\frac{1}{2}} \\ &\le C\|\theta\|_E \left(\sum_{j=1}^N \|(\psi-v)_x\|_{I_j}^2 + h^2\sum_{j=1}^N \|(\psi-v)_{xx}\|_{I_j}^2\right)^{\frac{1}{2}} \\ &\le Ch\|\theta\|_E|\psi|_2.\end{aligned} \tag{4.18}$$

Here we have used the estimate (4.5). Putting together we see that

$$\|\theta\|^2 \le C_1 h(\|\theta\|_E + h\|\theta\|)|\psi|_2 \le C_2 h(\|\theta\|_E + h\|\theta\|)\|\theta\| \Rightarrow \|\theta\| \le Ch\|\theta\|_E,$$

as long as h is small, where C, C_1, C_2 are independent of h and n. This when using the obtained error in energy norm (4.6) gives (4.13). □

4.4 Comments on extensions in dimension 2

The previous analysis can be extended to 2D rectangular meshes. First we take a tensor product of two one-dimensional projections as $\Pi = P^{(x)} \otimes P^{(y)}$, then as in [21, Theorem

7.3] we can obtain the following approximation result: if $w\in H^{s+1}(\Omega)$ for some $s\geq 1$, then

$$\sum_i\sum_j |w-\Pi w|^2_{p,K_{ij}} \leq Ch^{2(\min\{s,m\}-p+1)}\sum_i\sum_j |w|^2_{s+1,K_{ij}}, \tag{4.19a}$$

$$\sum_i\sum_j \|D^\alpha(w-\Pi w)\|^2_{p,\infty,K_{ij}} \leq Ch^{2(\min\{s,m\}-p+1/2)}\sum_i\sum_j |w|^2_{s+1,K_{ij}}, \tag{4.19b}$$

for any $0\leq p\leq \min\{s,m\}$, where C is independent of h. In the energy error estimate on 2D rectangular meshes, $A(\epsilon,e)$ contains more than just boundary terms. However, a careful estimate following [21, Lemma 6.3] can lead to the estimate of form

$$A(\epsilon,e)\leq Ch^{2\min\{s,k\}}|U|^2_{s+1}+\frac{1}{2}A(e,e)+k\|e\|^2, \tag{4.20}$$

where $e=\Pi U-u_h\in V_h\times V_h$, $\epsilon=\Pi U-U$, and $K_{ij}=[x_{i-1/2},x_{i+1/2}]\times[y_{j-1/2},y_{j+1/2}]$, C is independent of n and h. Such estimate can be used to obtain the desired order of error in the energy norm. The recovery of L^2 error is entirely similar. Without giving further details we state here the 2D result.

Theorem 4.4. *Let u_h be the numerical solution to the 2D DDG scheme on rectangular meshes with $\beta_0>\beta_0^*$, and U be the smooth solution to problem* (2.20)*, then*

$$\|U-u_h\|\leq Ch^{m+1}|U|_{m+1}, \tag{4.21}$$

where C is independent of h and n.

5 Numerical tests

In this section, we will numerically validate the boundary flux (2.12) as ν runs in $[0,1]$, through numerical convergence tests. The results show that $\nu=1$ is the best choice, with which we further present a test on the IDG method for the nonlinear Poisson-Boltzmann equation. The result is in agreement with the obtained optimal L^2 estimate.

Example 5.1. Consider the following boundary value problem

$$-u_{xx}=f,\quad u\left(-\frac{\pi}{2}\right)=e^{\sin\left(\frac{\pi^2}{2}\right)},\quad u\left(\frac{\pi}{2}\right)=e^{\sin\left(-\frac{\pi^2}{2}\right)}, \tag{5.1}$$

imposed on domain $[-\pi/2,\pi/2]$, with $f=-\pi^2e^{-\sin(\pi x)}\left(\cos^2(\pi x)+\sin(\pi x)\right)$. Its exact solution is known as

$$u=e^{-\sin(\pi x)}.$$

For simplicity, we only test three cases with $\nu=0.0$, $\nu=0.5$ and $\nu=1.0$. We first take $\beta_1=\frac{1}{2k(k+1)}$ and fixed parameters β_0, β_e satisfying (2.14) as listed in Table 1. The corresponding

Table 1: The choice of β_0, β_1 and β_e for P^m polynomials in 1D.

m	1	2	3	4
β_0	1.11	3.09	6.34	10.76
β_e	2.01	8.01	18.01	32.01
β_1	0	$\frac{1}{12}$	$\frac{1}{24}$	$\frac{1}{40}$

Table 2: 1D DDG scheme L^2 errors and orders.

m	ν	N=20	N=40		N=80		N=160	
		error	error	order	error	order	error	order
1	0.0	0.121079	0.035682	1.76	0.0094379	1.92	0.00234588	2.01
	0.5	0.107326	0.0307487	1.80	0.00793563	1.95	0.00194429	2.03
	1.0	0.102474	0.0248212	2.05	0.00526192	2.24	0.00106266	2.31
2	0.0	0.00244011	0.000366626	2.73	5.2454e-05	2.81	7.00019e-06	2.91
	0.5	0.00243223	0.000330262	2.88	4.47211e-05	2.88	5.83511e-06	2.94
	1.0	0.00242468	0.000295932	3.03	3.67582e-05	3.01	4.58213e-06	3.00
3	0.0	0.000336115	2.08618e-05	4.01	1.25643e-06	4.05	7.63528e-08	4.04
	0.5	0.000247371	1.55568e-05	3.99	9.49223e-07	4.03	5.81686e-08	4.03
	1.0	0.00013989	9.19519e-06	3.93	5.87913e-07	3.97	3.7056e-08	3.99
4	0.0	8.56477e-06	3.15246e-07	4.76	1.11719e-08	4.82	3.70713e-10	4.91
	0.5	8.56468e-06	2.91722e-07	4.88	9.78374e-09	4.90	3.17117e-10	4.95
	1.0	8.56458e-06	2.71192e-07	4.98	8.49586e-09	5.00	2.6592e-10	5.00

numerical results in Table 2 indicate that the DDG scheme is convergent with optimal L^2 errors of order h^{m+1}.

Note that (2.14) is only a sufficient condition for the DDG scheme to be stable, we hence also test $\beta_0=\beta_e=2,\beta_1=\frac{1}{12}$ for P^m $(m=1,2,3,4)$, which are used in [24]. The errors and orders given in Table 3 also show the optimal convergence, yet are inferior to the results in Table 2.

The numerical results in both Table 2 and Table 3 clearly show that $\nu=1$ is a better choice than $\nu<1$. We hence will test the nonlinear 2D nonlinear PB equation only with the choice of $\nu=1$.

Example 5.2. In this example we test the 2D nonlinear PB problem,

$$\begin{cases} -\lambda^2\Delta u=f(x)+e^{-u}, & (x,y)\in\Omega=[0,1]^2\subset\mathbb{R}^2, \\ u=g(x), & \text{on } \partial\Omega, \end{cases} \tag{5.2}$$

where $f=2\lambda^2\pi^2\cos\pi x\cos\pi y-e^{-(\cos\pi x\cos\pi y)}$ and $g=\cos\pi y\cos\pi x$ on rectangular meshes. The exact solution is

$$u=\cos\pi x\cos\pi y,$$

Table 3: 1D DDG scheme L^2 errors and orders.

m	ν	N=20	N=40		N=80		N=160	
		error	error	order	error	order	error	order
1	0.0	0.142399	0.0377084	1.92	0.00937737	2.01	0.00231008	2.02
	0.5	0.110152	0.0283125	1.96	0.0069494	2.03	0.00170145	2.03
	1.0	0.0720056	0.0154555	2.22	0.00343446	2.17	0.000797615	2.11
2	0.0	0.00295777	0.000939647	1.65	0.000155469	2.60	2.17393e-05	2.84
	0.5	0.00290126	0.00104135	1.48	0.000170568	2.61	2.36742e-05	2.85
	1.0	0.00278841	0.000349128	3.00	4.83296e-05	2.85	6.36683e-06	2.92
3	0.0	0.00379197	0.000136715	4.79	8.4193e-06	4.02	5.43934e-07	3.95
	0.5	0.00349559	0.000247371	3.82	1.55568e-05	3.99	9.49223e-07	4.03
	1.0	0.000811925	2.26798e-05	5.16	1.55717e-06	3.86	1.08557e-07	3.84
4	0.0	0.000267032	4.87809e-06	5.77	1.79582e-07	4.76	1.11469e-08	4.01
	0.5	7.17126e-05	2.96649e-06	4.60	1.25572e-07	4.56	4.78229e-09	4.71
	1.0	4.93837e-05	2.86064e-06	4.11	1.22926e-07	4.54	4.33637e-09	4.83

Table 4: 2D IDG scheme on rectangular meshes, $\lambda=1$.

m	iterations		N=4	N=8		N=16		N=32	
			error	error	order	error	order	error	order
1	10	$\|u-u_h\|_{L^2}$	0.0790251	0.0224836	1.81	0.00585643	1.94	0.00125092	2.23
		$\|u-u_h\|_{H^1}$	1.07804	0.601693	0.84	0.323835	0.89	0.141462	1.19
2	10	$\|u-u_h\|_{L^2}$	0.00183512	0.000235801	2.96	3.0053e-05	2.97	3.79997e-06	2.98
		$\|u-u_h\|_{H^1}$	0.0548449	0.0129734	2.08	0.00320584	2.02	0.000799251	2.00
3	10	$\|u-u_h\|_{L^2}$	8.72875e-05	6.4608e-06	3.76	4.39679e-07	3.88	2.86451e-08	3.94
		$\|u-u_h\|_{H^1}$	0.00493462	0.000657517	2.91	8.44388e-05	2.96	1.06975e-05	2.98

We use the iterative DG method as presented in [36], and the iteration process is terminated when $\|u_h^n-u_h^{n-1}\|<1.0\times10^{-10}$. Note that in each iteration step we apply the same DDG scheme formulation as given in [36] for 2D rectangular meshes, yet using the boundary flux as given in (2.12) in both x and y direction with $\nu=1$; while same problem is tested in [36, Example 4.2] with $\nu=1/2$. For the DDG flux parameters we use those listed in Table 1, and for P^m polynomial elements we test $m=1,2,3$. For the parameter λ we consider three cases with $\lambda=1,0.1,0.01$. The initial guess u_h^0 is obtained by solving (2.18) by the same DDG scheme when $\lambda=1$. But such a choice leads to unbounded k_n quickly when $\lambda=0.1$, 0.01, for which we simply take (2.4) for u_h^0. Note that (2.4) also works for $\lambda=1$, yet consuming more iteration steps. The numerical results on both L^2 and H^1 errors are reported in Table 4, 5 and 6, from which we see that optimal orders are obtained.

Table 5: 2D IDG scheme on rectangular meshes, $\lambda=0.1$.

m	iterations		N=4	N=8		N=16		N=32	
			error	error	order	error	order	error	order
1	49	$\|u-u_h\|_{L^2}$	0.0400215	0.0148916	1.43	0.00471664	1.66	0.00114131	2.05
		$\|u-u_h\|_{H^1}$	0.661225	0.435545	0.60	0.269079	0.69	0.132006	1.03
2	49	$\|u-u_h\|_{L^2}$	0.00160754	0.000225985	2.83	2.97089e-05	2.93	3.78872e-06	2.97
		$\|u-u_h\|_{H^1}$	0.0536654	0.0129531	2.05	0.00320562	2.01	0.00079925	2.00
3	49	$\|u-u_h\|_{L^2}$	8.43495e-05	6.38459e-06	3.72	4.38122e-07	3.87	2.86181e-08	3.94
		$\|u-u_h\|_{H^1}$	0.00482977	0.000653397	2.89	8.42983e-05	2.95	1.06929e-05	2.98

Table 6: 2D IDG scheme on rectangular meshes, $\lambda=0.01$.

m	iterations		N=4	N=8		N=16		N=32	
			error	error	order	error	order	error	order
1	113	$\|u-u_h\|_{L^2}$	0.0162307	0.00412912	1.97	0.00111791	1.89	0.000353022	1.66
		$\|u-u_h\|_{H^1}$	0.502242	0.252285	0.99	0.128061	0.98	0.0678027	0.92
2	113	$\|u-u_h\|_{L^2}$	0.00108235	0.000138494	2.97	1.97518e-05	2.81	3.1004e-06	2.67
		$\|u-u_h\|_{H^1}$	0.0554368	0.0136678	2.02	0.00331356	2.04	0.00080617	2.04
3	113	$\|u-u_h\|_{L^2}$	5.5932e-05	4.3261e-06	3.69	3.55312e-07	3.61	2.64524e-08	3.75
		$\|u-u_h\|_{H^1}$	0.00406606	0.000543387	2.90	7.61175e-05	2.84	1.0307e-05	2.88

6 Conclusion

This paper is concerned with the optimal error estimate for the iterative discontinuous Galerkin (IDG) method introduced in [36] to the nonlinear Poisson-Boltzmann equation in terms of both the iteration step n and the spatial mesh size h. The total error includes both the iteration error and the discretization error of the direct DG method to linear elliptic equations. For the DDG method, the optimal energy error is obtained by a constructive approach through an explicit global projection satisfying interface conditions dictated by the choice of numerical fluxes, followed by a careful recovery of the L^2 error of order $O(h^{m+1})$ for polynomials of degree m. Furthermore, we have shown that the bounding constant is independent of h and n, using several techniques including the elliptic regularity, the Morser-type estimate of the nonlinear source, as well as the point-wise bound of the iterative solutions. With a little further effort the results can be extended to two dimensional settings.

A Projection approximation error

We now present a detailed proof of Theorem 4.1.

We mimic the proof in [21, Section 7, Theorem 7.1] to derive bounds on the difference $u-Pu$ in terms of the Legendre coefficients of u in several steps (we only show the case of $s=m$):

Step 1: Denote by $\phi_i=L_i(\xi), i\geq 0$, the Legendre polynomial of degree i on $[-1,1]$ with $\|\phi_i\|^2=\frac{2}{2i+1}$ and $\phi_i(\pm 1)=(\pm 1)^i$ and expand the function u and Pu on I_j into the series

$$u|_{I_j}=\widetilde{u}_j(\xi):=\sum_{i=0}^{\infty}u_i^j\phi_i(\xi), \tag{A.1a}$$

$$Pu|_{I_j}=\widetilde{Pu}_j(\xi):=\sum_{i=0}^{m-2}u_i^j\phi_i(\xi)+a_{m-1}^j\phi_{m-1}(\xi)+a_m^j\phi_m(\xi), \tag{A.1b}$$

which satisfy the orthogonal property (4.1a). Hence,

$$(u-Pu)|_{I_j}=\sum_{i=m+1}^{\infty}u_i^j\phi_i(\xi)+(u_{m-1}^j-a_{m-1}^j)\phi_{m-1}(\xi)+(u_m^j-a_m^j)\phi_m(\xi). \tag{A.2}$$

In order to estimate (4.4), we find an equivalent expression as

$$\|\widetilde{u}_j-\widetilde{Pu}_j\|^2=\|\widetilde{u}_j-Q_m\widetilde{u}_j\|^2+\sum_{i=m-1}^{m}|u_i^j-a_i^j|^2\frac{2}{2i+1}, \tag{A.3a}$$

$$|(\widetilde{u}_j-\widetilde{Pu}_j)|\leq|\widetilde{u}_j-Q_m\widetilde{u}_j|_\infty+\sum_{i=m-1}^{m}|u_i^j-a_i^j|\cdot|\phi_i|_\infty, \tag{A.3b}$$

where Q_m is the standard L^2 projection from $L^2[-1,1]$ onto $P^m[-1,1]$ with $Q_m\widetilde{u}_j=\sum_{i=0}^{m}u_i^j\phi_i(\xi)$, so that

$$\widetilde{u}_j-Q_m\widetilde{u}_j=\sum_{i=m+1}^{\infty}u_i^j\phi_i(\xi).$$

Now, we show the upper bound of the right hand side of (A.3) term by term.

By the definition of Q_m, it follows that

$$\|\widetilde{u}_j-Q_m\widetilde{u}_j\|=\inf_{v\in P^m[-1,1]}\|\widetilde{u}_j-v\|\leq C|\widetilde{u}_j|_{m+1}, \tag{A.4a}$$

$$|\widetilde{u}_j-Q_m\widetilde{u}_j|_\infty\leq Ch^{m+1/2}|u|_{m+1,I_j}=C|\widetilde{u}_j|_{m+1}. \tag{A.4b}$$

Denote $\partial_\xi\widetilde{u}_j(\xi)=\sum_{i=0}^{\infty}\beta_i^j\phi_i(\xi)$, it was proved in [21, Section 7] that

$$\left|\sum_{i=m+1}^{\infty}u_i^j\phi_i(\pm 1)\right|^2\leq\frac{1}{2m+1}\|\partial_\xi\widetilde{u}_j\|^2, \tag{A.5a}$$

$$\sum_{i=1}^{m}|u_i^j|^2\leq 2\|\partial_\xi\widetilde{u}_j\|^2, \tag{A.5b}$$

where $\|\partial_{\xi}\widetilde{u}_j\|^2=\sum_{i=0}^{\infty}(\beta_i^j)^2\frac{2}{2i+1}$ and

$$u_i^j=\frac{\beta_{i-1}^j}{2i-1}-\frac{\beta_{i+1}^j}{2i+3},\quad i\geq 1. \tag{A.6}$$

We are left to show the upper bound for $u_i^j-a_i^j$ for $i=m-1,m$.

Step 2: From the interface conditions and boundary conditions (4.1b)-(4.1d) and a rearrangement of terms, we have for $j=1,\cdots N-1$,

$$\begin{aligned}
&\sum_{i=m-1}^{m}\phi_i(-1)(a_i^1-u_i^1)=\widetilde{b}_0^1:=u_{1/2}^+-\sum_{i=0}^{m}\phi_i(-1)u_i^1,\\
&\sum_{i=m-1}^{m}[\phi_i(1)(a_i^j-u_i^j)+\phi_i(-1)(a_i^{j+1}-u_i^{j+1})]=\widetilde{b}_1^j:=2\{u\}_{j+1/2}-\sum_{i=0}^{m}[\phi_i(1)u_i^j+\phi_i(-1)u_i^{j+1}],\\
&\sum_{i=m-1}^{m}[q_0(i)(a_i^j-u_i^j)+q_1(i)(a_i^{j+1}-u_i^{j+1})]=\widetilde{b}_2^j:=h\widehat{u_x}_{j+1/2}-\sum_{i=0}^{m}[q_0(i)u_i^j+q_1(i)u_i^{j+1}],\\
&\sum_{i=m-1}^{m}\phi_i(1)(a_i^N-u_i^N)=\widetilde{b}_0^N:=u_{N+1/2}^--\sum_{i=0}^{m}\phi_i(1)u_i^N,
\end{aligned} \tag{A.7}$$

where

$$q_0(i)=-\beta_0\phi_i(1)+\phi_i'(1)-4\beta_1\phi_i''(1), \tag{A.8a}$$
$$q_1(i)=\beta_0\phi_i(-1)+\phi_i'(-1)+4\beta_1\phi_i''(-1). \tag{A.8b}$$

Since this linear system is uniquely solvable, we have

$$\sum_{j=1}^{N}\sum_{i=m-1}^{m}|a_i^j-u_i^j|^2\leq C\left((\widetilde{b}_0^1)^2+(\widetilde{b}_0^N)^2+\sum_{j=1}^{N-1}((\widetilde{b}_1^j)^2+(\widetilde{b}_2^j)^2)\right). \tag{A.9}$$

Here C is independent of N. A detailed analysis of such independency will be given in Appendix B.

From

$$\widetilde{b}_0^1=\sum_{i=m+1}^{\infty}u_i^1\phi_i(-1)=\sum_{i=m+1}^{\infty}u_i^1(-1)^i$$

and (A.5a), we have

$$|\widetilde{b}_0^1|^2\leq\left|\sum_{i=m+1}^{\infty}u_i^1(-1)^i\right|^2\leq\frac{1}{2m+1}\|\partial_{\xi}\widetilde{u}_1\|^2. \tag{A.10}$$

In a similar fashion, we have

$$|\widetilde{b}_0^N|^2\leq\left|\sum_{i=m+1}^{\infty}u_i^N(-1)^i\right|\leq\frac{1}{2m+1}\|\partial_{\xi}\widetilde{u}_N\|^2. \tag{A.11}$$

As shown in [21, Section 7],

$$\sum_{j=1}^{N-1}((\widetilde{b}_1^j)^2+(\widetilde{b}_2^j)^2)\leq C(m,\beta_0,\beta_1)\sum_{j=1}^{N-1}\|\partial_\xi\widetilde{u}_j\|^2_{\min\{m,2\}}. \tag{A.12}$$

Thus, combining (A.9)-(A.12) gives

$$\sum_{j=1}^{N}\sum_{i=m-1}^{m}|a_i^j-u_i^j|^2\leq C\sum_{j=1}^{N}\|\partial_\xi\widetilde{u}_j\|_m^2. \tag{A.13}$$

Insertion of (A.4), (A.5) and (A.13) into the right hand side of (A.3) yields

$$\sum_{j=1}^{N}\|u-Pu\|_{I_j}^2\leq\frac{h}{2}\sum_{j=1}^{N}\|\widetilde{u}_j-Q_m\widetilde{u}_j\|^2+\frac{h}{2}C\sum_{j=1}^{N}\|\partial_\xi\widetilde{u}_j\|_m^2, \tag{A.14a}$$

$$\sum_{j=1}^{N}\|(u-Pu)\|_{\infty,I_j}^2\leq C\left(\sum_{j=1}^{N}\|\widetilde{u}_j-Q_m\widetilde{u}_j\|_\infty^2+\sum_{j=1}^{N}\|\partial_\xi\widetilde{u}_j\|_m^2\right). \tag{A.14b}$$

Replacing u by $u-v$, where v is an arbitrary element in V_h and taking in to account that $Pv=V$, $Q_m\widetilde{v}_j=\widetilde{v}_j$, then we have

$$\sum_{j=1}^{N}\|u-Pu\|_{I_j}^2\leq\frac{h}{2}\sum_{j=1}^{N}\|\widetilde{u}_j-Q_m\widetilde{u}_j\|^2+\frac{h}{2}C\sum_{j=1}^{N}\|\partial_\xi\widetilde{u}_j-\partial_\xi\widetilde{v}_j\|_m^2, \tag{A.15a}$$

$$\sum_{j=1}^{N}\|(u-Pu)\|_{\infty,I_j}^2\leq C\left(\sum_{j=1}^{N}\|\widetilde{u}_j-Q_m\widetilde{u}_j\|_\infty^2+\sum_{j=1}^{N}\|\partial_\xi\widetilde{u}_j-\partial_\xi\widetilde{v}_j\|_m^2\right). \tag{A.15b}$$

By Theorem 3.1.1 in [9], it follows

$$\inf_{q\in P^{m-1}([-1,1])}\|\partial_\xi\widetilde{u}_j-q\|_m\leq C|\partial_\xi\widetilde{u}_j|_m=C|\widetilde{u}_j|_{m+1}. \tag{A.16}$$

Step 3: Plugging (A.4) and (A.16) into the right hand side of (A.15), it follows that

$$\sum_{j=1}^{N}\|u-Pu\|_{I_j}^2\leq Ch\sum_{j=1}^{N}|\widetilde{u}_j|_{m+1}^2, \tag{A.17a}$$

$$\sum_{j=1}^{N}\|(u-Pu)\|_{\infty,I_j}^2\leq C\sum_{j=1}^{N}\left(|\widetilde{u}_j|_{m+1}^2+|\widetilde{u}_j|_{m+1}^2\right). \tag{A.17b}$$

Applying standard scaling to (A.17), we arrive at (4.4) with $p=0$.

Next, we show (4.4) with $p\neq0$. From (A.2) and $\partial_x=\left(\frac{h}{2}\right)^{-1}\partial_\xi$, it follows that

$$\partial_x^p(u-Pu)|_{I_j}=\left(\partial_\xi^p(\widetilde{u}-Q_m\widetilde{u})+(u_{m-1}^j-a_{m-1}^j)\phi_{m-1}^{(p)}(\xi)+(u_m^j-a_m^j)\phi_m^{(p)}(\xi)\right)\left(\frac{h}{2}\right)^{-p}. \tag{A.18}$$

Then,

$$\begin{aligned}&\|\partial_x^p(u-Pu)\|_{0,I_j}^2\\ \leq&3\left(\|\partial_\xi^p(\widetilde{u}-Q_m\widetilde{u})\|_{0,[-1,1]}^2+\sum_{i=m-1}^{m}|(u_i^j-a_i^j)|^2\|\phi_i^{(p)}(\xi)\|_{0,[-1,1]}^2\right)\left(\frac{h}{2}\right)^{-2p+1},\end{aligned}\tag{A.19a}$$

$$\begin{aligned}&\|\partial_x^p(u-Pu)\|_{\infty,I_j}^2\\ \leq&C\left(\|\partial_\xi^p(\widetilde{u}-Q_m\widetilde{u})\|_\infty^2+\sum_{i=m-1}^{m}|a_i^j-u_i^j|^2\right)\left(\frac{h}{2}\right)^{-2p}.\end{aligned}\tag{A.19b}$$

Taking summation and again replacing u by $u-v$ for any $v\in V_h$, we have

$$\begin{aligned}&\sum_{j=1}^{N}\|\partial_x^p(u-Pu)\|_{I_j}^2\\ \leq&3\sum_{j=1}^{N}\|\partial_x^p(u-Qu)\|_{I_j}^2+\left(Ch\sum_{j=1}^{N}\inf_{\widetilde{v}_j\in P^m([-1,1])}\|\partial_\xi\widetilde{u}_j-\partial_\xi\widetilde{v}_j\|_{m,[-1,1]}^2\right)\left(\frac{h}{2}\right)^{-2p},\end{aligned}\tag{A.20a}$$

$$\begin{aligned}&\sum_{j=1}^{N}\|\partial_x^p(u-Pu)\|_{\infty,I_j}^2\\ \leq&C\left(\sum_{j=1}^{N}\|\partial_x^p(u-Qu)\|_{\infty,I_j}^2+\left(\frac{h}{2}\right)^{-p}\sum_{j=1}^{N}\inf_{\widetilde{v}_j\in P^m([-1,1])}\|\partial_\xi\widetilde{u}_j-\partial_\xi\widetilde{v}_j\|_{m,[-1,1]}^2\right).\end{aligned}\tag{A.20b}$$

By plugging the standard L^2 projection error estimate

$$\sum_{j=1}^{N}\|\partial_x^p(u-Qu)\|_{I_j}^2\leq Ch^{2(m+1-p)}\sum_{j=1}^{N}|u|_{m+1,I_j}^2,\tag{A.21a}$$

$$\sum_{j=1}^{N}\|\partial_x^p(u-Qu)\|_{\infty,I_j}^2\leq Ch^{2m+1-2p}\sum_{j=1}^{N}|u|_{m+1,I_j}^2,\tag{A.21b}$$

and (A.16) with standard scaling into (A.20), the error estimates stated in (4.1a) and (4.1b) are obtained.

B A uniform bound

We now present a self-contained account to show that the bounding constant in (A.9) is independent of N. Denote the coefficient matrix of unknowns in (A.7) by A and we shall show that for any N,

$$\|A^{-1}\|\leq K,$$

uniformly in N. Note that A can be written as

$$A=\begin{bmatrix} \vec{c} & \vec{0} & \vec{0} & \cdots & \vec{0} & \vec{0} \\ \vec{d} & \vec{c} & 0 & \cdots & 0 & 0 \\ 0 & \vec{d} & \vec{c} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \vec{d} & \vec{c} \\ 0 & 0 & 0 & \cdots & \vec{0} & \vec{d} \\ \vec{e} & \vec{f} & 0 & \cdots & 0 & 0 \\ 0 & \vec{e} & \vec{f} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \vec{e} & \vec{f} \end{bmatrix}_{2N\times 2N}, \tag{B.1}$$

where

$$\begin{aligned} &\vec{c}=\begin{bmatrix}(-1)^{m-1} & (-1)^m\end{bmatrix}, \quad \vec{d}=\begin{bmatrix}1 & 1\end{bmatrix}, \\ &\vec{e}=\begin{bmatrix}q_0(m-1) & q_0(m)\end{bmatrix}, \quad \vec{f}=\begin{bmatrix}(-1)^m q_0(m-1) & (-1)^{m+1}q_0(m)\end{bmatrix}, \end{aligned} \tag{B.2}$$

where we have used the fact that

$$q_1(i)=(-1)^{i+1}q_0(i), \quad i=0,1,\cdots. \tag{B.3}$$

It suffices to show that the smallest eigenvalue of $B=AA^\top$, denoted by $\lambda_{\min}(B)$, is bounded from below for any N.

Note that

$$B=\left[\begin{array}{c|c} B_1 & B_2 \\ \hline B_2^T & B_3 \end{array}\right]_{2N\times 2N}, \tag{B.4}$$

where

$$B_1=\begin{bmatrix} 2 & & & & & \\ & 4 & & & & \\ & & 4 & & & \\ & & & \ddots & & \\ & & & & 4 & \\ & & & & & 2 \end{bmatrix}_{(N+1)\times(N+1)}, \quad B_3=\begin{bmatrix} \gamma_0 & \gamma_2 & & & & \\ \gamma_2 & \gamma_0 & \gamma_2 & & & \\ & \gamma_2 & \gamma_0 & \gamma_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \gamma_2 & \gamma_0 & \gamma_2 \\ & & & & \gamma_2 & \gamma_0 \end{bmatrix}_{(N-1)\times(N-1)}, \tag{B.5}$$

and

$$B_2=\begin{bmatrix} \gamma_1 & & & & & \\ & \gamma_1 & & & & \\ -\gamma_1 & & \gamma_1 & & & \\ & -\gamma_1 & & \gamma_1 & & \\ & & \ddots & & \ddots & \\ & & & -\gamma_1 & & \gamma_1 \\ & & & & -\gamma_1 & \\ & & & & & -\gamma_1 \end{bmatrix}_{(N+1)\times(N-1)}, \tag{B.6}$$

with

$$\gamma_0=2(q_0^2(m-1)+q_0^2(m)), \tag{B.7a}$$
$$\gamma_1=(-1)^{m-1}q_0(m-1)+(-1)^m q_0(m), \tag{B.7b}$$
$$\gamma_2=(-1)^m q_0^2(m-1)+(-1)^{m+1}q_0^2(m). \tag{B.7c}$$

Take a nonsingular P

$$P=\left[\begin{array}{c|c} I_{(N+1)\times(N+1)} & -B_1^{-1}B_2 \\ \hline 0 & I_{(N-1)\times(N-1)} \end{array}\right]_{2N\times 2N}, \tag{B.8}$$

to obtain

$$B'=P^TBP=\left[\begin{array}{c|c} B_1 & 0 \\ \hline 0 & B_3-B_2^TB_1^{-1}B_2 \end{array}\right]_{2N\times 2N}. \tag{B.9}$$

By taking $y=Px$, we have

$$\begin{aligned}\lambda_{\min}(B)&=\inf_{x\neq 0}\frac{x^TBx}{x^Tx}=\inf_{y\neq 0}\frac{y^TB'y}{y^TP^TPy}\\&=\inf_{y\neq 0}\frac{\frac{y^TB'y}{y^Ty}}{\frac{y^TP^TPy}{y^Ty}}\geq\frac{\lambda_{\min}(B')}{\lambda_{\max}(P^TP)},\end{aligned} \tag{B.10}$$

where $\lambda_{\max}(P^TP)>0$ denotes as the largest eigenvalue of P^TP. Next, we prove that $\lambda_{\max}(P^TP)$ is bounded from above and $\lambda_{\min}(B')$ is bounded from below for any N.

Note that

$$P^TP=\left[\begin{array}{c|c} I_{(N+1)\times(N+1)} & -B_1^{-1}B_2 \\ \hline -B_2^TB_1^{-1} & I_{(N-1)\times(N-1)}+B_2^T(B_1^{-1})^2B_2 \end{array}\right]_{2N\times 2N}, \tag{B.11}$$

with

$$B_1^{-1}B_2=\begin{bmatrix} \frac{\gamma_1}{2} & & & & & \\ & \frac{\gamma_1}{4} & & & & \\ -\frac{\gamma_1}{4} & & \frac{\gamma_1}{4} & & & \\ & -\frac{\gamma_1}{4} & & \frac{\gamma_1}{4} & & \\ & & \ddots & & \ddots & \\ & & & -\frac{\gamma_1}{4} & & \frac{\gamma_1}{4} \\ & & & & -\frac{\gamma_1}{4} & \\ & & & & & -\frac{\gamma_1}{2} \end{bmatrix}_{(N+1)\times(N-1)}, \tag{B.12a}$$

$$B_2^T(B_1^{-1})^2B_2=\begin{bmatrix} \frac{5\gamma_1^2}{16} & & -\frac{\gamma_1^2}{16} & & & & \\ & \frac{\gamma_1^2}{8} & & -\frac{\gamma_1^2}{16} & & & \\ -\frac{\gamma_1^2}{16} & & \frac{\gamma_1^2}{8} & & -\frac{\gamma_1^2}{16} & & \\ & \ddots & & \ddots & & \ddots & \\ & & -\frac{\gamma_1^2}{16} & & \frac{\gamma_1^2}{8} & & -\frac{\gamma_1^2}{16} \\ & & & -\frac{\gamma_1^2}{16} & & \frac{\gamma_1^2}{8} & \\ & & & & -\frac{\gamma_1^2}{16} & & \frac{5\gamma_1^2}{16} \end{bmatrix}_{(N-1)\times(N-1)}. \tag{B.12b}$$

By Gershgorin circle theorem, it follows that

$$|\lambda_{\max}(P^TP)-(P^TP)_{ii}|\le R_i,\quad i=1,2,\cdots,2N, \tag{B.13}$$

where $(P^TP)_{ii}$ is the ith diagonal entry, and

$$R_i=\sum_{j\neq i}|(P^TP)_{ij}|. \tag{B.14}$$

Thus, we have

$$0<\lambda_{\max}(P^TP)\le\max_i\Big(|(P^TP)_{ii}|+R_i\Big)=\frac{3}{8}r_1^2+\frac{3}{4}|r_1|+1. \tag{B.15}$$

From (B.9), we have

$$\begin{aligned}\lambda_{\min}(B')&=\min\{\lambda_{\min}(B_1),\lambda_{\min}(B_3-B_2^TB_1^{-1}B_2)\}\\&=\min\{2,\lambda_{\min}(B_3-B_2^TB_1^{-1}B_2)\}.\end{aligned} \tag{B.16}$$

Since

$$B_3-B_2^TB_1^{-1}B_2=\begin{bmatrix} p-r & 2q & r & & & & \\ 2q & p & 2q & r & & & \\ r & 2q & p & 2q & r & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & r & 2q & p & 2q & r \\ & & & r & 2q & p & 2q \\ & & & & r & 2q & p-r \end{bmatrix}_{(N-1)\times(N-1)}, \tag{B.17}$$

with

$$p=\gamma_0-\frac{\gamma_1^2}{2},\quad q=\frac{\gamma_2}{2},\quad r=\frac{\gamma_1^2}{4}. \tag{B.18}$$

If we take $\theta=\frac{\pi}{N}$, then the eigenvalues of (B.17) are (see [35])

$$\lambda(B_3-B_2^TB_1^{-1}B_2)=(p-2r)-\frac{1}{r}\left(q^2-(q-2r\cos j\theta)^2\right),\quad j=1,2,\cdots,N-1. \tag{B.19}$$

Since $\beta_0>\beta_0^*$, we observe that

$$q_0(m)=-\beta_0+\frac{1}{2}m(m+1)-\frac{\beta_1}{2}(m-1)m(m+1)(m+2)<0, \tag{B.20a}$$

$$q_0(m-1)=-\beta_0+\frac{1}{2}m(m-1)-\frac{\beta_1}{2}(m-2)(m-1)m(m+1)<0, \tag{B.20b}$$

then we have

$$\begin{aligned}|q|-2r&=\frac{|\gamma_2|-\gamma_1^2}{2}\\&=\frac{1}{2}\left(\left|q_0^2(m-1)-q_0^2(m)\right|-\left(q_0(m-1)-q_0(m)\right)^2\right),\end{aligned} \tag{B.21}$$

if $q_0(m)\le q_0(m-1)<0$,

$$|q|-2r=q_0(m-1)(q_0(m)-q_0(m-1))\ge 0; \tag{B.22}$$

if $q_0(m-1)\le q_0(m)<0$,

$$|q|-2r=q_0(m)(q_0(m-1)-q_0(m))\ge 0; \tag{B.23}$$

which result in

$$|q-2r\cos j\theta|\ge|q|-2r\ge 0. \tag{B.24}$$

So it follows

$$\begin{aligned}\lambda_{\min}(B_3-B_2^TB_1^{-1}B_2)&\ge(p-2r)-\frac{1}{r}\left(q^2-(|q|-2r)^2\right)\\&=p+2r-4|q|\\&=\gamma_0-2|\gamma_2|\\&=4\min\{q_0^2(m-1),q_0^2(m)\}>0\end{aligned} \tag{B.25}$$

as needed.

Acknowledgments

The authors thank the referees for valuable suggestions which led to significant improvements in this revised version. This work was supported by the National Science Foundation of USA under Grant DMS1312636 and by NSF Grant RNMS (Ki-Net) 1107291. Huang's work was supported by National Science Foundation of China under Grant 91430213.

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(5):1749–1779, 2002.

[2] L. Bedin and P. Bösing. Discontinuous Galerkin method for the linear Poisson-Boltzmann equation. *International Journal of Applied Mathematics*, 26(6):713–726, 2013.

[3] C. E. Baumann and J. T. Oden. A discontinuous hp finite element method for the Euler and Navier–Stokes equations. *Internat. J. Numer. Methods Fluids*, 31:79–95, 1999.

[4] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 131:267–279, 1997.

[5] I. Babuska and M. Zlamal. Nonconforming elements in the finite element method with penalty. *SIAM J. Numer. Anal.*, 10: 863–875, 1973.

[6] F. Brezzi, G. Manzini, D. Marini, P. Pietra and A. Russo. Discontinuous Galerkin approximations for elliptic problems. *Numer. Methods Partial Differential Equations*, 16:365–378, 2000.

[7] W. Cao, H. Liu and Z. Zhang. Superconvergence of the direct discontinuous Galerkin method for convection-diffusion equations. *Numer Methods Partial Differential Eq*, 33: 290–317, 2017.

[8] L. Chen, M. J. Holst and J. Xu. The finite element approximation of the nonlinear Poisson-Boltzmann equation. *SIAM J. Numer. Anal.*, 45(6):2298–2320, 2007.

[9] P. G. Ciarlet. The Finite Element Method for Elliptic Problems. *Studies in Mathematics and its Applications* Vol. 4, 1978.

[10] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35:2440–2463, 1998.

[11] J. Cheng, X. Yang, T. Liu and H. Luo. A direct discontinuous Galerkin method for the compressible Navier-Stokes equations on arbitrary grids. *54th AIAA Aerospace Sciences Meeting*, 1334–, 2016.

[12] J. Douglas and T. Dupont. Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods. *Lecture Notes in Phys.* 58, Springer-Verlag, Berlin, 1976.

[13] P. Degond, H. Liu, D. Savelief and M. H. Vignal. Numerical approximation of the Euler-Poisson-Boltzmann model in the quasineutral limit. *J. Sci. Comput.*, 51(1):59–86, 2012.

[14] W. Deng, X. Zhufu, J. Xu, S. Zhao. A new discontinuous Galerkin method for the nonlinear Poisson-Boltzmann equation. *Applied Mathematics Letters*, 49:126–132, 2015.

[15] B. Eisenberg and W. Liu. Poisson-Nernst-Planck systems for ion channels with permanent charges. *SIAM J. Math. Anal.*, 38:1932–1966, 2007.

[16] I. Gasser. A review on small Debye length and quasi-neutral limits in macroscopic models for charged fluids. *IMA Math. Appl.*, 136: 107–119, 2004.

[17] B. Hille. Ion Channels and Excitable Membranes. 3rd ed. (Sinauer Associates, Inc., Sunderland, MA, 2001).

[18] Y. Huang, H. Liu and N. Yi. Recovery of normal derivatives from the piecewise L^2 projection. *J. Comput. Phys*, 231(4):1230–1243, 2012.

[19] J. S. Hesthaven and T. Warburton. Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications. Springer, New York, 2007.

[20] J. L. Lions and E. Magenes. Non-homogeneous Boundary Value Problems and Applications. *New-York: Springer-Verlag*, 1972.

[21] H. Liu. Optimal error estimates of the direct discontinuous Galerkin method for convection-diffusion equations. *Math. Comp.*, 84(295): 2263–2295, 2015.

[22] H. Liu. Analysis of direct discontinuous Galerkin methods for elliptic and convection-diffusion equations. *Numerische Mathematik.*, 2017.
[23] H. Liu and J. Yan. The direct discontinuous Galerkin (DDG) method for diffusion problems. *SIAM Journal on Numerical Analysis.*, 47(1): 675–698, 2009.
[24] H. Liu and J. Yan. The Direct Discontinuous Galerkin (DDG) method for diffusion with interface corrections. *Commun. Comput. Phys.*, 8(3):541–564, 2010.
[25] H. Liu and H. Yu. The entropy satisfying dicontinuous Galerkin method for Fokker-Planck equations. *J. Sci. Comput.*, 62: 803–830, 2015.
[26] H. Liu and H. Yu. Maximum-principle-satisfying third order discontinuous Galerkin schemes for Fokker–Planck equations. *SIAM J. Sci. Comput.*, 36(5):A2296–A2325, 2014.
[27] H. Liu and Z. Wang. An entropy satisfying discontinuous Galerkin method for nonlinear Fokker-Planck equations. *J. Sci. Comput.*, 68:1217–1240, 2016.
[28] H. Liu and Z. Wang. A free energy satisfying discontinuous Galerkin method for Poisson–Nernst–Planck systems. *J. Comput. Phys.*, 238:413–437, 2017.
[29] A. Majda. Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables. *Applied Mathematical Sciences*, Vol. 53, 1984.
[30] P.A. Markowich, C.A. Ringhofer and C. Schmeiser. Semiconductor Equations. Springer-Verlag Inc., New York, 1990.
[31] L. Nirenberg, On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 13:115–162, 1959.
[32] B. Riviére. Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations. *SIAM*, Frontiers in Applied Mathematics, 2008.
[33] B. Riviére, M. F. Wheeler and V. Girault. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems I. *Comput. Geosci.*, 3:337–360, 1999.
[34] C.-W. Shu. Discontinuous Galerkin methods: general approach and stability. In *Numerical solutions of partial differential equations*, Adv. Courses Math. CRM Barcelona, pages 149–201. Birkhäuser, Basel, 2009.
[35] J. Todd. The condition of certain matrices, III. *Journal of Research of the National Bureau of Standards* 60(1):1–7, 1958.
[36] P. Yin, Y. Huang and H. Liu. An iterative discontinuous Galerkin method for solving the nonlinear Poisson Boltzmann equation. *Commun. Comput. Phys.* 16:491–515, 2014.