Computers and Mathematics with Applications 122 (2022) 35–47

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

www.elsevier.com/locate/camwa

Battling Gibbs phenomenon: On finite element approximations of discontinuous solutions of PDEs ☆

Shun Zhang

Department of Mathematics, City University of Hong Kong, Hong Kong, China

ARTICLE INFO

Keywords:
Discontinuous solution of PDEs
Gibbs phenomenon
Overshoot
Adaptive finite element methods
Singularly perturbed equation

ABSTRACT

In this paper, we want to clarify the Gibbs phenomenon when continuous and discontinuous finite elements are used to approximate discontinuous or nearly discontinuous PDE solutions from the approximation point of view. For a simple step function, we explicitly compute its continuous and discontinuous piecewise constant or linear projections on discontinuity matched or non-matched meshes. For the simple discontinuity-aligned mesh case, piecewise discontinuous approximations are always good. For the general non-matched case, we explain that the piecewise discontinuous constant approximation combined with adaptive mesh refinements is a good choice to achieve accuracy without overshoots. For discontinuous piecewise linear approximations, non-trivial overshoots will be observed unless the mesh is matched with discontinuity. For continuous piecewise linear approximations, the computation is based on a "far-away assumption", and non-trivial overshoots will always be observed under regular meshes. We calculate the explicit overshoot values for several typical cases. Numerical tests are conducted for a singularly-perturbed reaction-diffusion equation and linear hyperbolic equations to verify our findings in the paper. Also, we discuss the L^1-minimization-based methods and do not recommend such methods due to their similar behavior to L^2-based methods and more complicated implementations.

1. Introduction

For a wide range of partial differential equations, their solutions can be discontinuous or nearly discontinuous. For example, for singularly perturbed problems, when the transient layer is very sharp, the solution can be viewed as discontinuous, see [32,10]. For linear and nonlinear hyperbolic equations, discontinuous solutions can be caused by discontinuous initial data or shock forming, see [15,14]. For some cases, the location of discontinuity is known, for example, the discontinuity in the initial or boundary data. For many other cases, the exact location of discontinuity is unknown. When different numerical methods are used to solve such problems, the numerical solution is often *oscillatory* near a discontinuity, i.e., it overshoots (and undershoots). In spectral methods and Fourier analysis, it is called the Gibbs phenomenon, see [16,12]. In some literature, they are also called *wiggles*, see [37]. Such spurious oscillations or wiggles are unacceptable in many situations; see discussions on page 209 of Hesthaven [14] for examples. These oscillations are essentially caused by a simple fact: continuous functions are used to approximate a discontinuous function. In numerical methods (finite difference, finite volume, spectral, and discontinuous Galerkin methods) for hyperbolic equations, some nonlinear tricks are used to eliminate or reduce the Gibbs phenomenon: artificial viscosity, limiters, filters, and other methods, see [15,14,12].

On the other hand, in the numerical elliptic and singularly perturbed problems community using finite element methods, the Gibbs phenomenon and how to eliminate it are still quite mystical. Many numerical methods are tested for problems with discontinuity and near discontinuity. Sometimes the numerical solutions have no Gibbs phenomenon, while on the other occasions, the spurious oscillations appear for the same numerical method on a slightly different setting. There are many explanations for the phenomenon and how to reduce it. The following views are very common.

1 When the solution is discontinuous, if methods based on discontinuous piecewise polynomials such as discontinuous Galerkin (DG) methods are used, the overshoots can be eliminated or reduced. For example, in numerical computation for a singularly perturbed problem in a SIAM review paper [10] (p. 165), the authors claimed the DG methods would have no overshoots for almost discontinuous solutions.
2 The adaptive finite element method with mesh refinements based on a posteriori error estimation is a useful procedure for detect-

☆ This work was supported in part by Hong Kong Research Grants Council under the GRF Grant Projects No. CityU 11302519 and CityU 11305319.
E-mail address: shun.zhang@cityu.edu.hk.

https://doi.org/10.1016/j.camwa.2022.07.014
Received 6 January 2022; Received in revised form 10 July 2022; Accepted 21 July 2022

0898-1221/© 2022 Elsevier Ltd. All rights reserved.

ing the regions and locations with bad approximations; see [1,35]. When the mesh is fine enough, the discontinuity is resolved, and the overshoot can be reduced.

3 In the literature of finite difference/finite volume schemes, there are various *order barrier theorems*, see [11,37]. For a simple linear equation $d\phi/dx = q(x)$, Theorem 4.2.2 of [37] states that linear discretization schemes of positive type are at most of first-order accurate. Similarly, there is the famous Godunov's order barrier theorem, see [11] and Theorem 9.2.2 of [37]: for the linear transport equation with constant coefficient, the linear one-step second-order accurate numerical schemes cannot be monotonically preserving. These theorems are developed for specific equations and use the notions of positive or monotonically preserving numerical schemes. Theorem 4.2.2 of [37] is proved by a simple local Taylor expansion and calculation of local truncation error for a one-dimensional uniform numerical grid. Do we have the same order barrier for other numerical methods and other partial differential equations?

4 There is also some research on the relation between the L^1-best approximation and the Gibbs phenomenon. The paper [33] of Saff and Tashevof shows that if continuous piecewise linear polynomials approximate a jump discontinuity on a discontinuity-matched uniform mesh in one dimension in the L^1-best approximation sense, the Gibbs phenomenon vanishes. In the context of numerical methods for PDEs, papers with the L^1-based numerical methods include [22,23,19,13,17,18,25].

In this paper, we want to clarify this question from the approximation point of view by using a simple model problem. Since most finite element or DG methods are based on projections or pseudo-projections in Hilbert spaces, the best result a numerical method can achieve of a certain approximation often cannot be better than its L^2-projection. That is, if V_h is the discrete approximation space, the best we can hope for the numerical error is often in the following form (note that for many complicated problems, it is essential to use adaptive methods to achieve the desired optimality; and many methods cannot have such a good result, for example, the DG method for the linear hyperbolic equation):

$$\|u - u_h\|_0 \le C \inf_{v_h \in V_h} \|u - v_h\|_0 = C\|u - \Pi_{V_h} u\|_0, \tag{1.1}$$

where Π_{V_h} is the L^2-projection operator onto the discrete approximation space V_h and $\|\cdot\|_0$ is the L^2-norm. Since the Gibbs phenomenon is about the wiggles/overshoots/undershoots, the error is pointwise. Thus, to study the Gibbs phenomenon, we should measure the error in an L^∞ like norm. In Theorem 6.5 of [7], for the singularly perturbed reaction-diffusion problem with a two-step finite element method, the following result,

$$\|u - u_h\|_{L^\infty(K)} \le C\|u - \Pi_{V_h} u\|_{L^\infty(K)} + \text{extra terms on slightly larger domains of } K, \tag{1.2}$$

is proved, where V_h is the discontinuous piecewise polynomial space on a finite element mesh, K is an element of the finite element mesh, and Π_{V_h} is the L^2-projection onto V_h. Inspired by (1.2), at least from the approximation point of view, to study the Gibbs phenomenon of finite element approximations to (nearly) discontinuous solutions, we need to study the term $\|u - \Pi_{V_h} u_h\|_{L^\infty(K)}$ for different V_h spaces (continuous or discontinuous finite elements on discontinuity-aligned or nonaligned, uniform or adaptive meshes) with Π_{V_h} as the L^2-projection.

We also need to emphasize that the numerical stability of the numerical method for a specific equation is different from the Gibbs phenomenon or spurious oscillation, as discussed on page 119 of Wesseling [37]. The numerical stability is essential for the success of the numerical method, but it cannot guarantee the disappearance of the Gibbs phenomenon. The Gibbs phenomenon (spurious oscillation or wiggle) is essentially an approximation problem. Some tricks to increase the numerical stability, such as the artificial/numerical viscosity, will enhance both the stability and reduce the Gibbs phenomenon. We assume that the artificial/numerical viscosity is not used.

Choosing u to be the simplest discontinuous function, a step function, we try to study the properties of $\Pi_{V_h} u$ with V_h being the continuous or discontinuous finite element space on a discontinuity-aligned or nonaligned, uniform or adaptive mesh. Based on explicit computations, we find the following results. For the simple discontinuity-aligned mesh case, piecewise discontinuous approximations are always good in terms of overshoot. A special but important discontinuity-aligned mesh case is the boundary layer, where the (almost) discontinuity is on the boundary and is matched with the numerical mesh. For the general discontinuity-nonaligned mesh case, we explain that **the piecewise discontinuous constant approximation combined with adaptive mesh refinements is the best choice to achieve accuracy without overshoot**. For discontinuous piecewise linear approximations, non-trivial overshoots will be observed unless the mesh is matched with discontinuity. For continuous piecewise linear approximations, non-trivial overshoots will always be observed under regular meshes. We calculate the explicit overshoot values for several typical cases.

We also find that a special high ratio mesh obtained by adaptive mesh refinements combined with coarsening can reduce the overshoot. However, such a case is not always possible in two and three dimensions or one dimension with other error sources.

The paper is organized as follows. Section 2 discusses a simple case of piecewise discontinuous polynomial approximations on a discontinuity-aligned mesh. Section 3 describes a model problem of approximating a step function by L^2 projections onto piecewise constant and linear function spaces. Discontinuous piecewise constant and linear approximations are explicitly computed for the model problem. We also compute continuous piecewise linear approximations under a "far-away assumption" for the model problem. In Section 4, linear conforming, linear DG, and lowest-order mixed finite element methods are tested on a one-dimensional singularly perturbed problem. Several test problems of 2D linear transport problems with P0- and P1-DGFEMs are done in Section 5. In Section 6, we comment on known numerical test results in the literature. We discuss the L^1-minimization-based methods in Section 7. We present a table to summarize the results for different cases in Section 8. In Section 9, we make some concluding remarks.

2. Piecewise discontinuous polynomial approximations on a discontinuity-aligned mesh

In this section, we consider an almost trivial but important case, piecewise discontinuous polynomial approximations on a discontinuity-aligned mesh. Assume $\mathcal{T}_h = \{K\}$ is a finite element mesh on d-dimensional polygonal/polyhedral domain, and the function $u|_K \in H^2(K)$ on each element $K \in \mathcal{T}_h$. By the Sobolev embedding theorem, we gave $u|_K \in C^0(K)$. The norm $\|u\|_{L^\infty(K)}$ is also well-defined. The function is continuous inside each element, and thus the discontinuity of the function (if any) is aligned with the mesh. Let $P_k(K)$ be the space of polynomials defined on K whose degree is less or equal to integer $k \ge 0$ and let $P_k(\mathcal{T}_h) := \{v \in L^2(\Omega) : v|_K = P_k(K), \forall K \in \mathcal{T}_h\}$ be the space of piecewise discontinuous polynomial space. The diameter of the element K is defined as h_K. The notation $|\cdot|_{2,K}$ denotes the H^2-semi-norm on an element K. We have the following Theorem.

Theorem 2.1. *Assume on each element $K \in \mathcal{T}_h$, $u|_K \in H^2(K)$. Let $u_{k,K} \in P_k(K)$ be the L^2-projection of u in $P_k(K)$, then for some constant $C > 0$ independent of h,*

$$\|u - u_{k,K}\|_{L^\infty(K)} \le C h_K^{2-d/2} |u|_{2,K} \quad \text{for } d = 1, 2, 3. \tag{2.3}$$

Proof. The proof here is a modification of the argument in p.93 of Braess [4]. Since $u \in C^0(K)$, let $u_{I,K}$ be the well-defined nodal interpolation of u in $P_k(K)$. By the triangle inequality, we have

$$\|u-u_{k,K}\|_{L^\infty(K)} \le \|u-u_{I,K}\|_{L^\infty(K)} + \|u_{I,K}-u_{k,K}\|_{L^\infty(K)}. \tag{2.4}$$

By the inverse estimate (Lemma 1.138 of [9]), we get

$$\|u_{I,K}-u_{k,K}\|_{L^\infty(K)} \le Ch_K^{-2/d}\|u_{I,K}-u_{k,K}\|_{0,K}. \tag{2.5}$$

By the triangle inequality and the approximation properties of the L^2-projection and the nodal interpolation,

$$\|u_{I,K}-u_{k,K}\|_{0,K} \le \|u-u_{k,K}\|_{0,K} + \|u-u_{I,K}\|_{0,K} \le Ch_K^2|u|_{2,K}. \tag{2.6}$$

Combining above two inequalities, we have

$$\|u_{I,K}-u_{k,K}\|_{L^\infty(K)} \le Ch_K^{2-2/d}|u|_{2,K}. \tag{2.7}$$

The theorem is proved by substituting it into (2.4) and using the approximation property of the interpolation (Theorem 16.1 of [8]):

$$\|u-u_{I,K}\|_{L^\infty(K)} \le Ch_K^{2-2/d}|u|_{2,K}. \quad \square$$

Remark 2.2. Note that there are various more refined max-norm/pointwise error estimates for different methods and equations; see [8,36,20]. Most of these analyses are based on the global smoothness of the solution.

Remark 2.3. From the theorem, it is clear that for the L^2-projection of the piecewise discontinuous polynomial approximation on a discontinuity-aligned mesh, the pointwise error will decrease to zero as the mesh size decreases. Combined with adaptive mesh refinements, there is no or only neglectable Gibbs phenomenon for this case.

Remark 2.4. Note that some boundary layer problems fall into this category. When there is a sharp solution change near the boundary, the solution can be viewed as discontinuous on the Dirichlet boundary. In this case, the mesh is always aligned with the boundary (assuming the domain is polygonal/polyhedral); thus it is a discontinuity-aligned case.

From the discussion in this section, to eliminate the Gibbs phenomenon/overshoot, the boundary condition should be enforced weakly. The famous Nitsche's method [31] and discontinuous Galerkin method [2,15] are successful for this case. We also need to emphasize that weak enforcement should be chosen carefully. A simple penalization may not work since it is not weak enough, see some numerical experiments in [26].

Remark 2.5. For the Gibbs phenomenon near a discontinuity, we are only interested in the value of the numerical solution that overshoots/undershoots. It may be possible that the numerical approximation lies between the discontinuity gap and the maximum error at that location is large, but no Gibbs phenomena happen. Thus, the measurement we use to quantify the Gibbs phenomena for a nonaligned mesh is a little weaker than the standard L^∞-norm; see also our definition of os in Sections 3.2.2 and 3.3.2.

3. An illustrative one-dimensional problem

3.1. A model problem

Consider a step function u:

$$u(x) = \begin{cases} -1 & x<0, \\ 1 & x>0. \end{cases} \tag{3.1}$$

Here, the discontinuity gap $u(0+)-u(0-)$ is 2. For more general cases with a gap $2c$, $c>0$, it is easy to see that the corresponding results below are proportional to c.

Let $\mathcal{T}_h=\{K\}$ be a one-dimensional mesh on $\Omega=(-a,b)$ with elements (intervals in this 1-D setting) denoted by K. We assume that a and b are big enough positive numbers with respect to the mesh size. Consider three approximation spaces defined on $\mathcal{T}_h$:

$$P_0(\mathcal{T}_h) := \{v\in L^2(\Omega) : v|_K = P_0(K), \forall K\in\mathcal{T}_h\}, \tag{3.2}$$

$$P_1(\mathcal{T}_h) := \{v\in L^2(\Omega) : v|_K = P_1(K), \forall K\in\mathcal{T}_h\}, \tag{3.3}$$

$$S_1(\mathcal{T}_h) := \{v\in C^0(\Omega) : v|_K = P_1(K), \forall K\in\mathcal{T}_h\}. \tag{3.4}$$

Let V_h be $P_0(\mathcal{T}_h)$, $P_1(\mathcal{T}_h)$, or $S_1(\mathcal{T}_h)$. The L^2-projection $u_h\in V_h$ is defined by:

$$(u_h,v_h)_\Omega = (u,v_h)_\Omega \quad \forall v_h\in V_h.$$

We use notations $u_{h,0}$, $u_{h,1}$, and $u_{c,1}$ to denote the L^2-projections of u on $P_0(\mathcal{T}_h)$, $P_1(\mathcal{T}_h)$, and $S_1(\mathcal{T}_h)$, respectively.

3.2. Discontinuous approximations

We consider two discontinuous approximations first: P_0 and P_1 approximations. Note that, for a discontinuous projection, the approximation is exact on those elements away from the discontinuity. Thus, we only need to discuss the special interval that contains the discontinuity. Let $0\le t\le 1$ and $h>0$. We consider the interval $I_0=(-th,(1-t)h)$, see (a) of Fig. 1.

3.2.1. Discontinuous piecewise constant approximation

A simple computation shows that the constant projection of u on the interval I_0 is the following,

$$u_{h,0}|_{I_0} = 1-2t. \tag{3.5}$$

It is obvious that $-1\le 1-2t\le 1$ for $t\in[0,1]$, thus, there is no overshoot. If $t=0$ or 1, which is a discontinuity-aligned mesh, the numerical approximation is exact.

Computing the L^2-error in the interval I_0, we have,

$$\|u-u_{h,0}\|_{0,I_0} = 2\sqrt{t(1-t)h}. \tag{3.6}$$

The worst case is $\|u-u_{h,0}\|_{0,I_0}=\sqrt{h}$, when $t=1/2$. This also matches the a priori analysis that $\|u-u_{h,0}\|_{0,I_0}\le Ch^{1/2-\epsilon}$ for $u\in H^{1/2-\epsilon}(I_0)$ for an arbitrary small $\epsilon>0$.

With a good a posteriori error estimator that can identify the bad approximated elements, the discontinuity-crossing elements will be found and divided, and the L^2- and other integration-based norms of error will become smaller and smaller.

3.2.2. Discontinuous piecewise linear approximation

In this subsection, we consider the L^2-projection of u onto a linear function on I_0. To this end, let λ_{-1} and λ_1 be the linear Lagrange basis functions define on $(-th,(1-t)h)$ with $\lambda_i(z_i)=1$ and $\lambda_i(z_{-i})=0$, where $i=1$ or -1, and $z_{-1}=-th$ and $z_1=(1-t)h$.

Let the projection $u_{h,1}|_{I_0}=U_{-1}\lambda_{-1}+U_1\lambda_1$. The projection problem is

$$\begin{pmatrix} (\lambda_{-1},\lambda_{-1})_{I_0} & (\lambda_1,\lambda_{-1})_{I_0} \\ (\lambda_{-1},\lambda_1)_{I_0} & (\lambda_1,\lambda_1)_{I_0} \end{pmatrix}\begin{pmatrix} U_{-1} \\ U_1 \end{pmatrix} = \begin{pmatrix} (u,\lambda_{-1})_{I_0} \\ (u,\lambda_1)_{I_0} \end{pmatrix}.$$

All terms of the matrix problem can be computed exactly. Solving the projection problem, we get

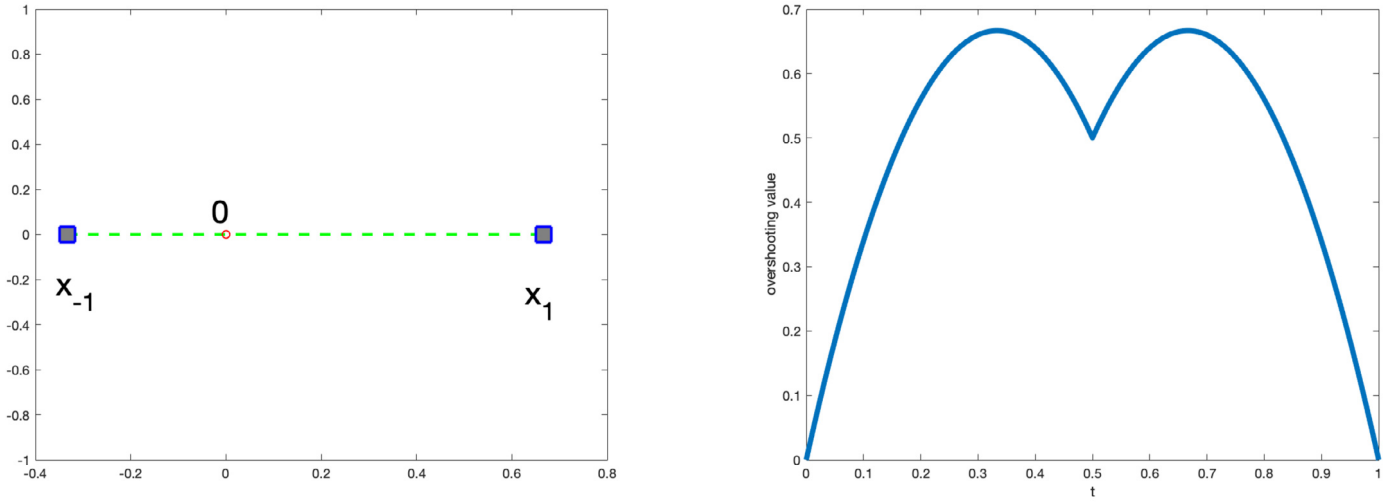
$$U_{-1} = 1-8t+6t^2 \quad \text{and} \quad U_1 = 1+4t-6t^2.$$

It is easy to show that

$$(U_1-1) > -(U_{-1}+1) \ge 0 \quad \text{if } 0<t<1/2,$$

$$-(U_{-1}+1) > (U_1-1) \ge 0 \quad \text{if } 1/2<t<1.$$

To quantify the overshoot phenomena, we define the following overshoot value function:



(a) the interval I_0 (b) overshoot values as $t \in (0, 1)$

Fig. 1. Discontinuous piecewise linear approximation on a discontinuity non-aligned mesh for a step function.

$$\texttt{os} = \max(U_1 - 1, -(U_{-1} + 1), 0).$$

As discussed in Remark 2.5, we do not use the $\|u - u_{h,1}\|_{L^\infty}$ here since we are only interested in the os (which is weaker than $\|u - u_{h,1}\|_{L^\infty}$).

We plot the value of os for $0 \le t \le 1$ on the right of Fig. 1. It is easy to see that only when the mesh is aligned with the discontinuity ($t = 0$ or 1), os is zero. For t away from 0 or 1, the overshoot phenomenon is severe. The maximum overshoot value $2/3$ appears at $t = 1/3$ or $2/3$.

We also notice that if the bisection of mesh is used, the relative position of the discontinuity normally will not converge to the left or the right point of the internal. For example, consider the interval $(-1/3, 2/3)$, i.e., the discontinuity is at the $1/3$ position of the internal. If we bisect the interval, the new interval containing the discontinuity is $(-1/3, 1/6)$, and the discontinuity is at the $2/3$ position. Keeping doing the bisection, we will find that the discontinuity will jump between the $1/3$ and $2/3$ positions. In this case, t will always be $1/3$ or $2/3$, and the overshoot value will not decrease. For other initial positions other than the aligned case, the overshoot values will oscillate between $[\delta, 2/3]$, for some $0 < \delta < 2/3$.

Computing the L^2-error in the interval I_0, we have,

$$\|u - u_{h,1}\|_{0,I_0} = 2\sqrt{t(1-t)(1-3t+3t^2)h}. \tag{3.7}$$

When $t = 0$ or 1, the error is zero. The largest error $\sqrt{h/3}$ appears at $t = 0.5 \pm \sqrt{3}/6$. Note that $1/4 \le 1 - 3t + 3t^2 \le 1$ for $0 \le t \le 1$, we do have

$$\|u - u_{h,1}\|_{0,I_0} \le \|u - u_{h,0}\|_{0,I_0}. \tag{3.8}$$

Thus, for discontinuous approximations, even though the L^2-norm of error of the P_1 approximation is better than the P_0 approximation, they are of the same order if the mesh is not aligned, and the overshoot cannot be avoided in most cases.

3.3. Continuous piecewise linear approximation

For the continuous piecewise linear approximation, the exact solution is not in the approximation space, even for the discontinuity-aligned mesh case. The projection is not local to one element anymore. Luckily, based on the numerical experiments, we found that the Gibbs phenomenon is most severe on the two nodes around the discontinuity point. The more nodes between the discontinuity and the node to be observed, the smaller the error. We assume the following **far-away assumption** for the continuous piecewise linear approximation.

Assumption 3.1. Approximation on nodes far-away from the discontinuity (far-away assumption): For a node x_k of the mesh $\mathcal{T}_h$, if there is at least two nodes between the discontinuity position and x_k, the error at x_k, $|u_{c,1}(x_k) - u(x_k)|$, is assumed to be less than $1/10$ the maximum error at the possible overshoot locations (the discontinuity point, and its two neighbor nodes).

For a node x_k of the mesh $\mathcal{T}_h$, if there is one node between the discontinuity position and x_k, the error at x_k, $|u_{c,1}(x_k) - u(x_k)|$, is assumed to be less than $1/2$ of the maximum error at the possible overshoot locations (the discontinuity point, and its two neighbor nodes).

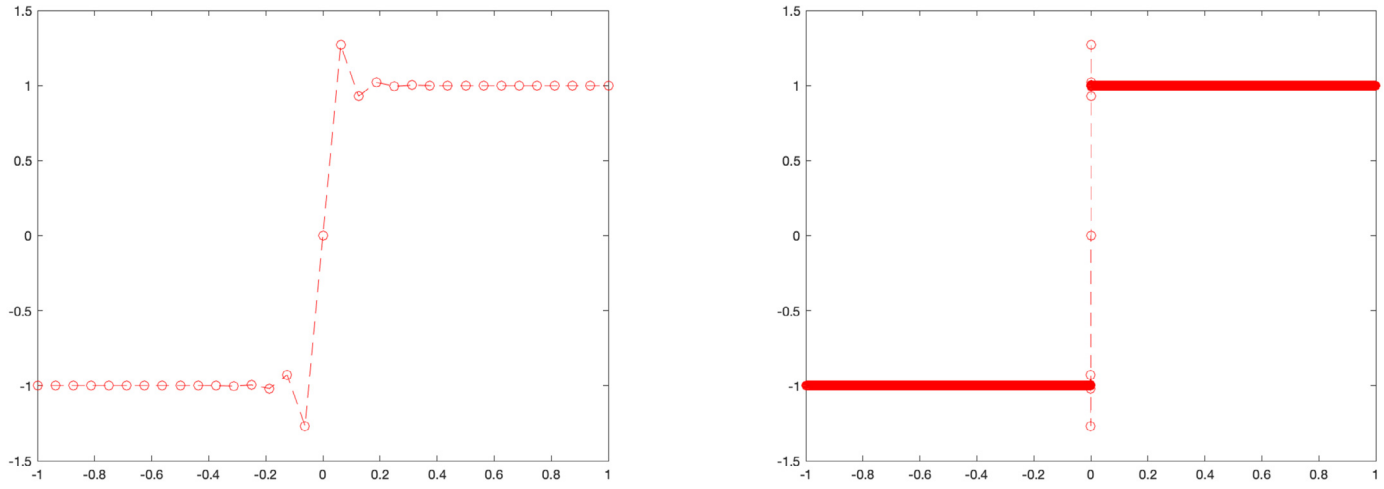
In our calculations, we plan to study the maximum error at the possible overshoot locations (the discontinuity point, and its two neighbor nodes) in our calculations. We assume that the error at x_k, which is two nodes away from the discontinuity, is neglectable. For those points only one node away from the discontinuity, we will either compute their values or assume it has a perturbation ϵ to be determined later.

3.3.1. Discontinuity-aligned mesh

We first consider a possible ideal case: the mesh is aligned with the discontinuity and the mesh is symmetric with respect to the discontinuity at 0. Let $\mathcal{T}_h$ be a symmetric mesh on $[-1, 1]$ with nodes: $x_{-N} < x_{-N+1} < \cdots < x_{-1} < x_0 < x_1 < \cdots < x_{N-1} < x_N$, where $x_{-i} = -x_i$ and $x_0 = 0$.

We first test on a uniform mesh with $h = 1/N$ and $x_{\pm i} = \pm ih$. From the numerical tests, for example, $h = 1/16$ and $h = 2^{-14}$ in Fig. 2, we note that the values of $u_{c,1}$ at the center nodes $x_{\pm i}$ are identical for different choice h for $i \le 4$. For i big enough, $u_{c,1}(x_{\pm i})$ is almost identical to u. The values $u_{c,1}(x_i)$ for $i = -4, -3, \cdots, 3, 4$ are -0.9948, -1.0192, -0.9282, -1.2679, 0, 1.2679, 0.9282, 1.0192, and 0.9948. The severe overshoots only appear on the two nodes around the discontinuity with a size of 0.2679. For the points $i = -3$ and $i = 3$, the errors are 0.0192, less than $1/10$ of the overshoot error of 0.2679. For the points $i = -2$ and $i = 2$, the errors are 0.0718, less than $1/3$ of the overshoot error 0.2679. The far-away assumption is true in this case.

With the far-away assumption, a simple calculation can be done to show why the values of $u_{c,1}$ at $x_{\pm 1}$ are close to ± 1.25 as observed. We simplify the calculation by assuming that $u_{c,1} = u$ at all $x_{\pm i}$ for $i > 2$. We assume the values of $u_{c,1}(x_{\pm 2}) = \pm(1 + \epsilon)$, where $|\epsilon|$ is a small number compared to 1. We have the rest three values at $x_{\pm 1}$ and x_0 to be computed by the projection. By the symmetry of nodes, it is easy to see that $u_{c,1}$ at $x_0 = 0$ is 0 and $u_{c,1}(x_i) = -u_{c,1}(x_{-i})$. Thus we only need to determine the value of $u_{c,1}(x_1) =: U_1$. Let λ_1 and λ_2 be the linear Lagrange



(a) $h = 2^{-4}$ (b) $h = 2^{-14}$

Fig. 2. Continuous piecewise linear approximations on a uniformly refined mesh for a step function.

basis functions on the mesh, with $\lambda_i(x_j) = \delta_{i,j}$, $i, j \in \{1, 2\}$. Then $u_{c,1}$ on (x_0, x_2) can be written as

$$u_{c,1} = U_1\lambda_1 + (1+\epsilon)\lambda_2, \quad x \in (x_0, x_2).$$

The value of U_1 can be obtained by a simple projection,

$$(U_1\lambda_1 + (1+\epsilon)\lambda_2, \lambda_1)_{(x_0,x_2)} = (1, \lambda_1)_{(x_0,x_2)}.$$

A simple calculation shows

$$U_1 = 1.25 - \epsilon/4. \tag{3.9}$$

Thus, the overshoot value is always close to $1/8$ of the discontinuity jump (2 in our example). The exact value of 1.2679 is due to the fact that ϵ is about -0.0718 in the example.

The L^2-errors in the two adjacent elements of 0 are also easy to be computed by choosing ϵ to be 0,

$$\|u - u_{c,1}\|_{0,(0,h)} \approx \|u - u_{c,1}\|_{0,(-h,0)} \approx \sqrt{13h/48} \approx 0.5204\sqrt{h}. \tag{3.10}$$

3.3.2. *Discontinuity-nonaligned mesh*

We consider the case where the mesh is not aligned with the discontinuity. From the above discussion, we notice that the size of each element is not essential (their ratio is important, see the example below). Three cases are considered: a local uniform mesh, a local adaptive mesh, and an adaptive mesh with coarsening.

A local uniform mesh Suppose $0 < t < 1$ and $h > 0$, we consider the following mesh of 5 elements with size h: $x_{-3} = -(t+2)h$, $x_{-2} = -(t+1)h$, $x_{-1} = -th$, $x_1 = (1-t)h$, $x_2 = (2-t)h$, and $x_3 = (3-t)h$, see the left of Fig. 3 for a test mesh with $t = 1/3$ and $h = 1$. The discontinuity cuts through the central interval $I_0 = (-th, (1-t)h)$ at 0. Let λ_i be the linear Lagrange basis function for $i = \pm1, \pm2, \pm3$. As before, we assume that $u_{c,1}$ is exact at $i = \pm3$ (we omit the small perturbation ϵ for simplicity). Then on the interval (x_{-3}, x_3),

$$u_{c,1} = -\lambda_{-3} + U_{-2}\lambda_{-2} + U_{-1}\lambda_{-1} + U_1\lambda_1 + U_2\lambda_2 + \lambda_3.$$

Solving the following projection problem,

$$(u_{c,1}, \lambda_i)_{(x_{-3},x_3)} = (u, \lambda_i)_{(x_{-3},x_3)}, \quad i = \pm1, \pm2,$$

we get

$$U_{-2} = -3(87 - 60t + 38t^2)/209, \quad U_{-1} = (-1 - 720t + 456t^2)/209,$$
$$U_1 = (265 + 192t - 456t^2)/209, \quad U_2 = 3(65 - 16t + 38t^2)/209.$$

We define the following overshoot value function:

$$\mathtt{os} = \max(-(U_{-2}+1), -(U_{-1}+1), (U_1 - 1), (U_2 - 1), 0).$$

We plot the overshoot value with respect to t on the right of Fig. 3. The overshoot values lie in the interval

$$[0.1818, 0.3646].$$

Also, when $0 < t < 1/2$, the maximum overshoot appears at U_1, and when $1/2 < t < 1$, the maximum overshoot happens at U_{-1}, that is, the maximum overshoot appears at one of the endpoints of I_0, which is farther from the discontinuity. The overshoot is non-trivial for any $0 < t < 1$.

We have the following L^2-error:

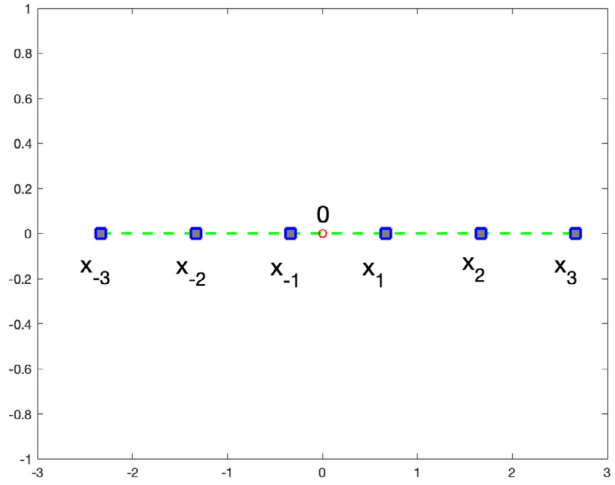
$$\|u - u_{c,1}\|_{0,I_0} = 2\sqrt{h(8869 + 60189t - 294117t^2 + 467856t^3 - 233928t^4)/131043}. \tag{3.11}$$

For $0 \le t \le 1$, we get $\|u - u_{c,1}\|_{0,I_0}$ is between $0.5203\sqrt{h}$ and $0.6236\sqrt{h}$. Note this error matches with the result of (3.10), which corresponds to $t = 0$ or 1.

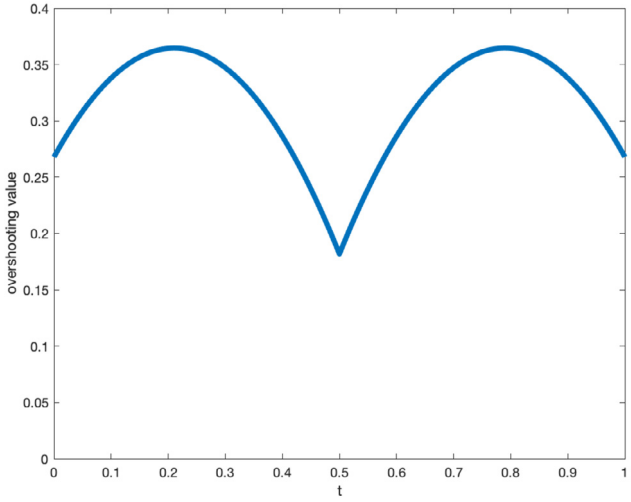
A local adaptive mesh We also test the following mesh: $x_{-3} = -(t+4)h$, $x_{-2} = -(t+2)h$, $x_{-1} = -th$, $x_1 = (1-t)h$, $x_2 = (2-t)h$, and $x_3 = (4-t)h$, with sizes $2h$, h, h, and $2h$, respectively. This kind of mesh can be viewed as an example of an adaptively refined mesh. With this mesh setting and assuming the values of $u_{c,1}$ at $x_{\pm3}$ are exact, by a similar procedure, we get

$$U_{-2} = 1/241(-281 + 138t - 87t^2),$$
$$U_{-1} = 3/241(-27 - 184t + 116t^2),$$
$$U_1 = 1/241(325 + 144t - 468t^2),$$
$$U_2 = 1/241(227 - 24t + 78t^2).$$

We show the result on the right of Fig. 4. The overshoot value lies in the interval

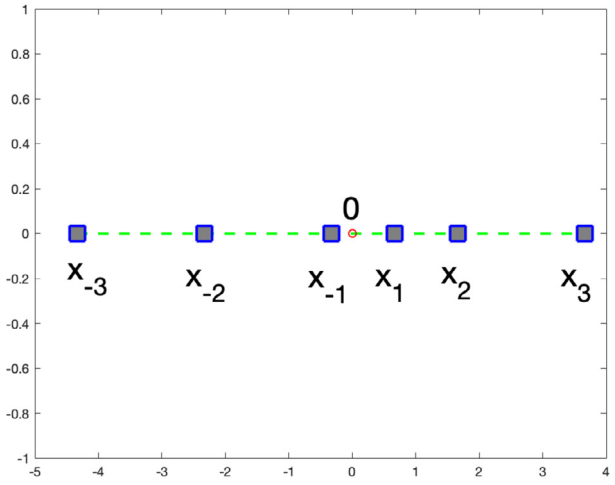


(a) a test mesh with $t = 1/3$ and $h = 1$

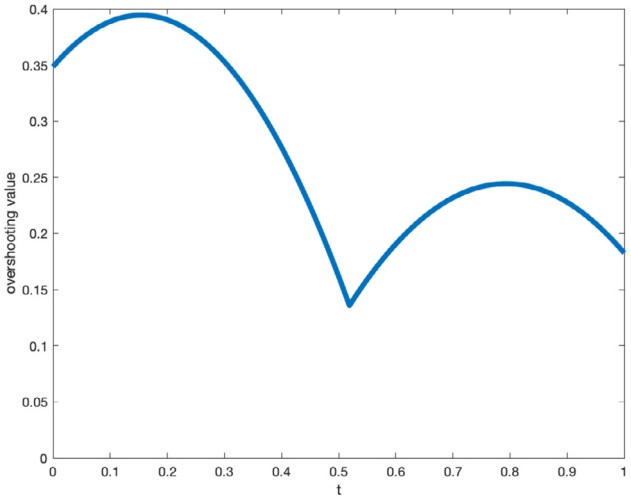


(b) overshoot values as $t \in (0, 1)$

Fig. 3. Continuous piecewise linear approximation on a discontinuity-nonaligned mesh for a step function.



(a) a test mesh



(b) overshoot values as $t \in (0, 1)$

Fig. 4. Continuous piecewise linear approximation on a discontinuity non-aligned nonuniform mesh.

$$[0.1358, 0.3954]. \tag{3.12}$$

Also, as before, we find that the maximum overshoot happens at U_1 when $0 < t < 1/2$, and it happens at U_{-1} when $1/2 < t < 1$. For any choice of $0 < t < 1$, the overshoot is non-trivial. We vary the element size of elements adjacent to I_0 from h, $2h$, $4h$, to $8h$, and get similar results of non-trivial overshoots. This shows that adaptive mesh refinements by bisection will not make the overshoot phenomenon disappear for a continuous linear approximation.

We have the following L^2-error:

$$\|u - u_{c,1}\|_{0,I_0} = 2\sqrt{h(20923 + 13173t - 221541t^2 + 450576t^3 - 250668t^4)/174243}. \tag{3.13}$$

For $0 \le t \le 1$, $\|u - u_{c,1}\|_{0,I_0}$ is between $0.5349\sqrt{h}$ and $0.6965\sqrt{h}$.

An adaptive mesh with coarsening For the one-dimensional problem, there is one special case that the overshoot phenomenon that can be eliminated.

Consider the following mesh $\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots$, with $x_0 = 0$, $x_{\pm 1} = \pm h$, and $x_{\pm 2} = \pm ch$, with $h > 0$ and $c > 1$. Similar to the uniform grid case, the projection $u_{c,1}$ on (x_0, x_2) can be written as

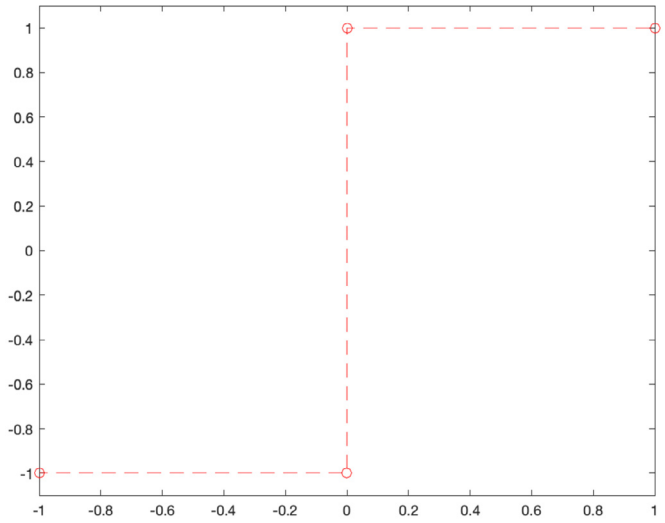
$$u_{c,1} = U_1\lambda_1 + (1 + \epsilon)\lambda_2, \quad x \in (x_0, x_2).$$

Computing by projection, we get

$$U_1 = 1 + \frac{1 - c\epsilon}{2(1 + c)}.$$

This means when the ratio $c > 1$ is big enough, the overshoot value can be reduced to a very small value. For a one-dimensional mesh, this is possible by combining the refining and coarsening, i.e., when an a posteriori error indicator on an element K is big, then the element is refined, while for those elements with small indicators, they are combined with the close elements to form bigger elements.

We test the case with the following example: the domain is $(-1, 1)$ and the initial mesh is $[-1; -1/3; 2/3; 1]$. We use the exact L^2 error as the error estimator, i.e., $\eta_K = \|u - u_{c,1}\|_{0,K}$, and set the refinement/coarsening criteria as "refine those elements whose $\eta_K > 0.6 \max_{\{T \in \mathcal{T}_h\}} \eta_T$ and coarsen those elements whose $\eta_K < 0.3 \max_{\{T \in \mathcal{T}_h\}} \eta_T$". For this very



(a) overshoot values as $t \in (0,1)$

Fig. 5. Continuous piecewise linear approximation with adaptive refinement/coarsening.

simple problem, the final mesh only contains three elements, a central very small element around 0, and two big elements on the left and right, receptively, see Fig. 5. After 14 iterations, the final mesh is $[-1; -0.0000814; 0.0000407; 1]$; the numerical solutions at four nodes are $[-1; -1.00008; 0.99998; 1.0]$. The overshoot phenomenon is almost invisible.

Note that such case will only happen in a one-dimensional setting. In two and three dimensions, to keep the mesh conforming and regular, the radio of mesh sizes between two adjacent elements sharing a same edge/face cannot be big, thus there will always be non-trivial overshoots. Even for one-dimensional problems, such a high ratio mesh is probably not possible due to other approximation errors to keep the mesh from coarsening, see a numerical test in Section 4.

4. Numerical test in 1D: a 1D singularly perturbed problem

Consider the following problem:

$$-\epsilon u'' + u = f, \quad \text{on } (0,1), \quad \text{and} \quad u(0) = u(1) = 0, \tag{4.1}$$

where

$$f = \begin{cases} 2x & x < 1/2, \\ 2x - 2 & x > 1/2. \end{cases}$$

We choose $\epsilon = 10^{-16}$. For ϵ this small, the solution is almost identical to f, and has an extremely sharp layer at $x = 1/2$. For the case that mesh size is not small enough to match with the layer, $x = 1/2$ can be viewed as a discontinuous location.

We test the problem with three numerical methods: the linear C^0-conforming finite element method, linear discontinuous Galerkin finite element method, and the lowest-order mixed finite element method.

Let $0 = x_0 < x_1 < \cdots < x_N = 1$ be a one-dimensional mesh $\mathcal{T}_h$ of $(0,1)$, with $K_i = (x_i, x_{i+1})$ and $h_i = x_{i+1} - x_i$.

Linear conforming finite element method (P1-conforming): Find $u_h \in S_{1,0}(\mathcal{T}_h)$, such that

$$(\epsilon u_h', v') + (u_h, v) = (f, v), \quad v \in S_{1,0}(\mathcal{T}_h), \tag{4.2}$$

where $S_{1,0}(\mathcal{T}_h) = S_1(\mathcal{T}_h) \cap H_0^1(\Omega)$. Note that for a very small ϵ, this is the L^2-projection to $S_{1,0}(\mathcal{T}_h)$ we discussed.

Linear discontinuous Galerkin finite element problem (P1-DG): Find $u_{dg} \in P_1(\mathcal{T}_h)$, such that

$$a_{dg}(u_{dg}, v) = (f, v), \quad \forall\, v \in P_1(\mathcal{T}_h), \tag{4.3}$$

where

$$\begin{aligned} a_{dg}(w, v) := {} & (\epsilon w', v')_{\mathcal{T}_h} + (w, v)_{\mathcal{T}_h} \\ & - \sum_{i=1}^{N-1} \left(\{\epsilon w'(x_i)\} [\![v(x_i)]\!] + \{\epsilon v'(x_i)\} [\![w(x_i)]\!] \right) \\ & + \sum_{i=1}^{N-1} \frac{\epsilon \mu}{h_i + h_{i-1}} [\![w(x_i)]\!] [\![v(x_i)]\!], \end{aligned}$$

where $\mu > 0$ is a big enough number (we can safely choose $\mu = 10$). For a very small ϵ, this can be viewed as the L^2-projection to $P_1(\mathcal{T}_h)$.

To introduce the mixed method, we let $\sigma = -\epsilon u' \in H^1(0,1)$, then $\sigma' + u = f$. The mixed variational formulation is: find $(\sigma, u) \in H^1(0,1) \times L^2(0,1)$, such that,

$$\begin{cases} (\epsilon^{-1}\sigma, \tau) - (\tau', u) = 0, & \forall \tau \in H^1(0,1), \\ -(\sigma', v) - (u, v) = -(f, v), & \forall v \in L^2(0,1). \end{cases} \tag{4.4}$$

We choose the approximation spaces to be S_1 and P_0 for σ and u, respectively.

Mixed finite element problem (P0-mixed): Find $(\sigma_h, u_h) \in S_1(\mathcal{T}_h) \times P_0(\mathcal{T}_h)$, s.t.,

$$\begin{cases} (\epsilon^{-1}\sigma_h, \tau) - (\tau', u_h) = 0, & \forall \tau \in S_1(\mathcal{T}_h), \\ -(\sigma_h', v) - (u, v) = -(f, v), & \forall v \in P_0(\mathcal{T}_h). \end{cases} \tag{4.5}$$

For a very small ϵ, this can be viewed as the L^2-projection to $P_0(\mathcal{T}_h)$.

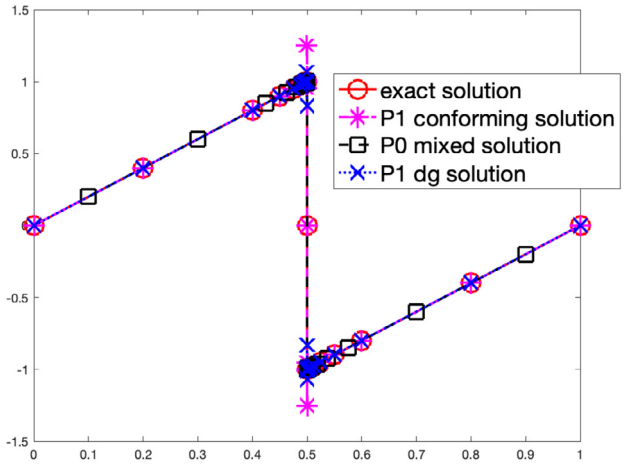
In the computation, we use the robust error estimator for the conforming finite element method in [34,21] to drive the adaptive mesh refinement and compute the conforming, DG, and mixed finite element solutions on the same adaptive mesh.

Discontinuity-aligned mesh case. An initial mesh $[0; 0.2; 0.4; 0.6; 0.8; 1]$ is chosen. We refine it 20 times with the maximum refinement strategy $\theta = 0.8$. The center nodes are $0.5 + 10^{-6} \times [-0.3815; -0.1907; 0; 0.1907; 0.3815]$. Numerical approximations with linear conforming, P0-mixed, and P1-DG approximations are shown in Fig. 6. The overshoot value for the linear conforming finite element approximation is $1.2546 - 1 = 0.2546$, matching the discontinuity match case. The P0-mixed solution has no overshoot because it is a piecewise constant approximation, and the P1-DGFEM has no overshoot since the discontinuity is matched with the mesh.

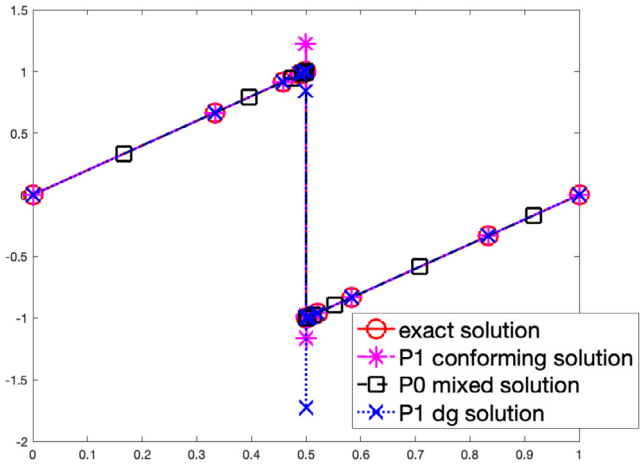
Discontinuity non-aligned mesh case. Next, we choose an initial mesh $[0; 1/3; 5/6; 1]$. With bisections, the discontinuity location will never be exactly matched. The center nodes are $0.5 + 10^{-5} \times [-0.2543; -0.0636; -0.0159; 0.0318; 0.1272; 0.5086]$. The values of linear conforming solution u_h at two nodes adjacent to $1/2$ are 1.2284 and -1.1647. This also matches the discussion in (3.12). A value of -1.7242 is found of P1-DGFEM approximation at the node $x = 0.0318$. The value is a little big larger than the discussion in the previous section, where a continuous piecewise linear function is used to approximate a discontinuous step function, while here it is a discontinuous piecewise linear function instead of a step function. The discontinuous mixed solution has no overshoot in this case.

For a similar problem, on the bottom page 165 of [10], the authors claimed that the DG method for this problem will have no oscillations. This is imprecise: the DG method will have oscillations unless the discontinuity is matched with the mesh. The numerical test in [10] uses a uniform mesh with $h = 1/N$, which is exactly the discontinuity-aligned case. Thus, the non-overshoot in [10] is due to the use of DGFEM and because of using DGFEM on a discontinuity-aligned mesh.

Adaptive mesh with both refinement and coarsening. Next, we test the problem with both refinement and coarsening. The initial mesh is chosen to be a uniform mesh with $h = 1/17$, and the nodes on the final mesh are $0, 0.49262, 0.49264, 0.9412, 1.0000$. The C^0 conforming finite element solution u_h at these nodes are $0, 0.95, -1.05, -0.10, 0$ with almost no overshoot, see Fig. 7.



(a) solutions with a discontinuity matched mesh



(b) solutions with a discontinuity non-matched mesh

Fig. 6. 1D reaction-diffusion problem with adaptive refinement on discontinuity aligned and non-aligned meshes.

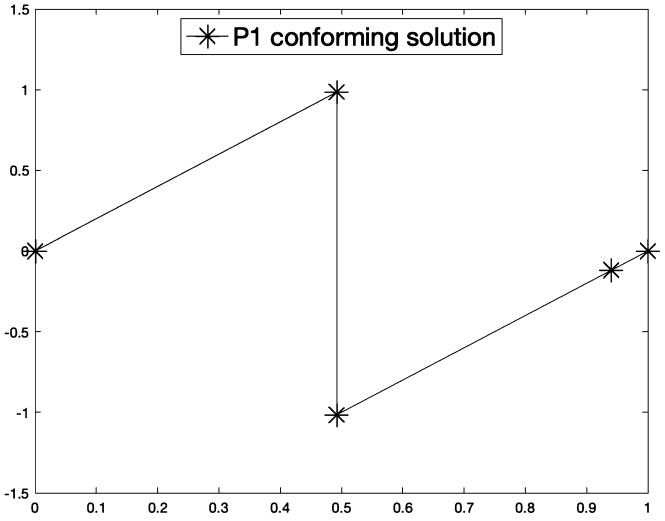


Fig. 7. 1D reaction-diffusion problem with adaptive refinement/coarsening.

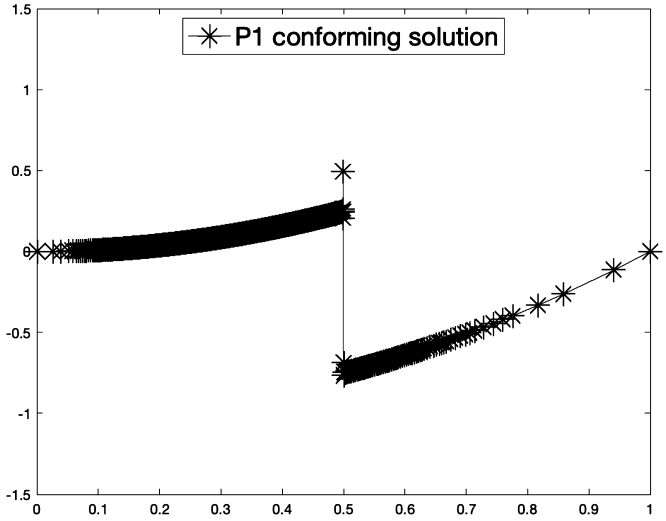


Fig. 8. 1D reaction-diffusion problem, adaptive refinement and coarsening with f_2.

We should point out that it is not always possible to get a high ratio mesh such as that in Fig. 7 with a combination of adaptive refinement and coarsening. We test the problem with the right-hand side:

$$f_2 = \begin{cases} x^2 & x < 1/2, \\ x^2 - 1 & x > 1/2. \end{cases}$$

Note that there is some approximation error on all elements for the linear finite element approximation. Thus these elements are not coarsened, we eventually have a standard adaptively refined mesh, and an overshoot phenomenon appears, see Fig. 8.

5. Numerical test in 2D: DG methods for a linear transport equation

Consider the following linear transport (advection) equation:

$$\begin{aligned} \nabla \cdot (\boldsymbol{\beta} u) + \gamma u &= f \quad \text{in } \Omega, \\ u &= g \quad \text{on } \Gamma_-, \end{aligned} \tag{5.1}$$

where the advective velocity field $\boldsymbol{\beta} = (\beta_1, \beta_2)^T \in [C^1(\overline{\Omega})]^2$ is a vector-valued function defined on $\bar{\Omega}$ and $\gamma \in L^\infty(\Omega)$. The inflow part of $\partial\Omega$, $\Gamma_- = \{x \in \partial\Omega : \boldsymbol{\beta}(x) \cdot \mathbf{n}(x) < 0\}$, is defined in the usual fashion, where $\mathbf{n}(x)$ denotes the unit outward normal vector to $\partial\Omega$ at $x \in \partial\Omega$.

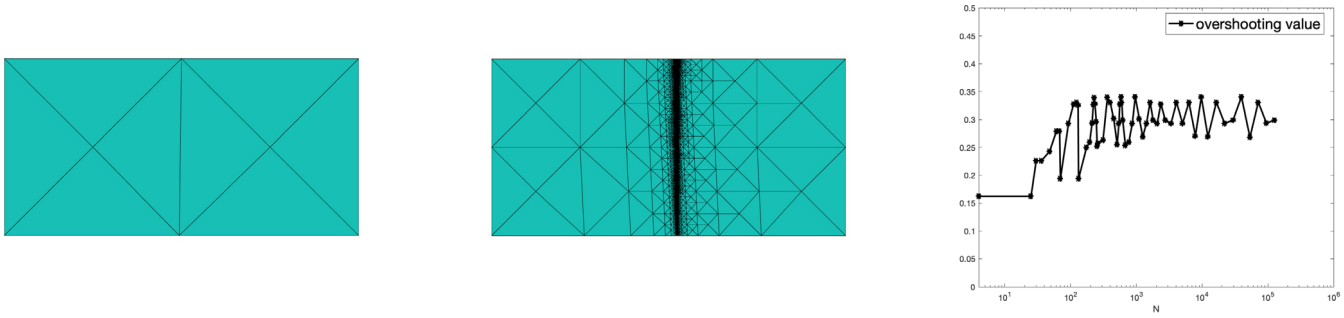
The DGFEM for the transport equation is [5]: find $u_h \in P_k$, $k = 0$ or 1 such that

$$a_{dg}(u_h, v_h) = f(v_h), \quad \forall v_h \in P_k, \tag{5.2}$$

where the bilinear form and linear form are defined by

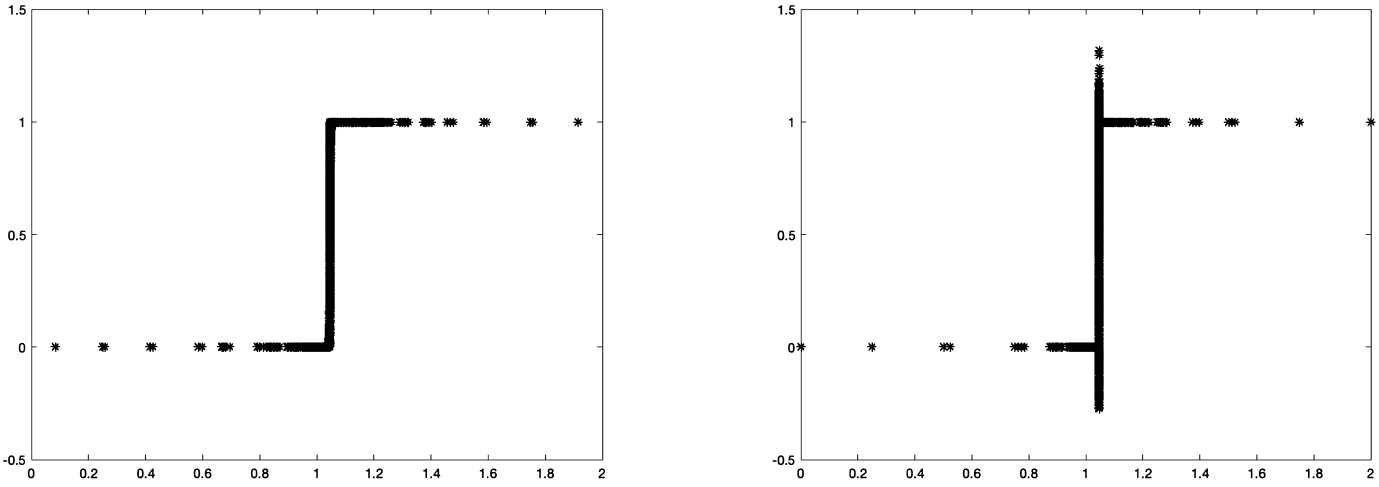
$$\begin{aligned} a_{dg}(w_h, v_h) &:= \sum_{K \in \mathcal{T}_h} (w_h, -\boldsymbol{\beta} \cdot \nabla v_h + \gamma v_h)_K + \sum_{F \in \mathcal{E}_I \cup \mathcal{E}_{out}} (\{\boldsymbol{\beta} \cdot \mathbf{n} w_h\}_{up}, [\![v_h]\!])_F, \\ f(v_h) &= (f, v_h) - \sum_{F \in \mathcal{E}_{in}} (\boldsymbol{\beta} \cdot \mathbf{n} g, v_h)_F, \end{aligned}$$

where the term $\{\boldsymbol{\beta} \cdot \mathbf{n} w_h\}_{up}$ is the usual upwind flux, meaning that $\{\boldsymbol{\beta} \cdot \mathbf{n} w_h\}_{up}$ takes the value on the inflow side of the edge/face F. The sets $\mathcal{E}_{in}$, $\mathcal{E}_I$, and $\mathcal{E}_{out}$ are the sets of inflow, interior, and outflow edges, respectively. We call the methods P0-DGFEM and P1-DGFEM for $k = 0$ and 1. We use a residual type of error estimator similar to that developed in [6] to drive the adaptive mesh refinement.



(a) initial mesh (b) final adaptive mesh (c) overshoots on different levels of refined meshes with P1-DGFEM

Fig. 9. Piecewise constant solution transport problem with a non-matching grid test problem.



(a) P0-DGFEM (b) P1-DGFEM

Fig. 10. Projected solutions with P0- and P1-DGFEMs for piecewise constant solution transport problem on a non-matching mesh.

5.1. A discontinuous solution on a non-matching mesh

Consider the problem: $\Omega = (0,2)\times(0,1)$ with $\boldsymbol{\beta} = (0,1)^T$. The inflow boundary is $\{x \in (0,1), y = 0\}$. Let $\gamma = 0$ and $f = 0$. Choose the inflow boundary condition such that the exact solution is

$$u(x,y) = \begin{cases} 0 & \text{if} \quad x < \pi/3, \\ 1 & \text{if} \quad x > \pi/3. \end{cases}$$

The initial mesh is shown on the left of Fig. 9. The point $(\pi/3, 0)$ and $(1,1)$ are the bottom central node and the top central node, respectively. Although the inflow boundary condition is matched, the mesh and its subsequent bisection mesh are not aligned with the discontinuity.

In the center of Fig. 9, a final adaptive mesh is shown with many refinements along the discontinuity since the mesh is not matched. On the right of Fig. 9, we show the overshoot values of the P1-DGFEM on different levels of refined meshes, with the overshoot value defined as:

$$\text{os} = \max_{\mathbf{x}\in\mathcal{N}} \{u_h(\mathbf{x}) - 1, -u_h(\mathbf{x}), 0\}, \tag{5.3}$$

$\mathcal{N}$ is the collection of nodes of the finite element mesh.

From the figure, we see that the overshoot values oscillate between 0.15 and 0.35 and never go to zero when the mesh is adaptively refined, which is matched with our analysis. For the P0-DGFEM, the overshoot value is in the order of machine accuracy, 10^{-16}, so there is no overshoot at all.

Since the exact solution of the problem is not changing with respect to the y-coordinate, we plot a projected solution by plotting the numerical solutions of the element center (P_0)/ nodes (P_1) suppressing the y-coordinate. In Fig. 10, we show the P0-DGFEM projected solution computed on the final adaptive mesh on the left and the P1-DGFEM one on the right. We clearly see the overshoot of P1-DGFEM.

5.2. Curved transport problem 1

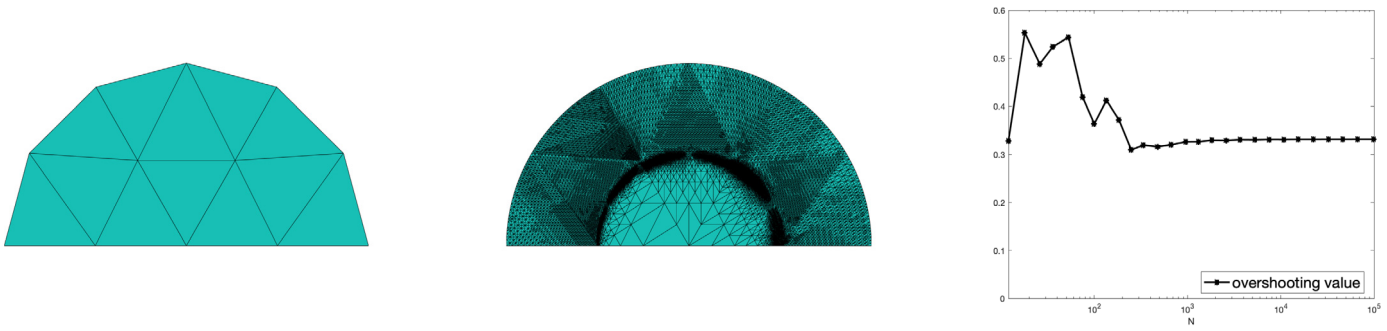
Consider the problem on the half disk $\Omega = \{(x,y) : x^2 + y^2 < 1; y > 0\}$. The inflow boundary is $\{-1 < x < 0; y = 0\}$. Let $\boldsymbol{\beta} = (\sin\theta, -\cos\theta)^T = (y/\sqrt{x^2+y^2}, -x/\sqrt{x^2+y^2})^T$, with θ is the polar angle. Let $\gamma = 0$, $f = 0$, and the inflow condition and the exact solution be

$$g = \begin{cases} 1 & \text{if} \quad -1 < x < -0.5, \\ 0 & \text{if} \quad -0.5 < x < 0, \end{cases} \quad \text{and} \quad u = \begin{cases} 1 & \text{if } x^2 + y^2 > 0.25, \\ 0 & \text{otherwise.} \end{cases}$$

The initial mesh is shown on the left of Fig. 11. On the center of Fig. 11, a final adaptive mesh is shown. On the right of Fig. 9, we show the overshoot values of the P1-DGFEM, with the overshoot value (5.3).

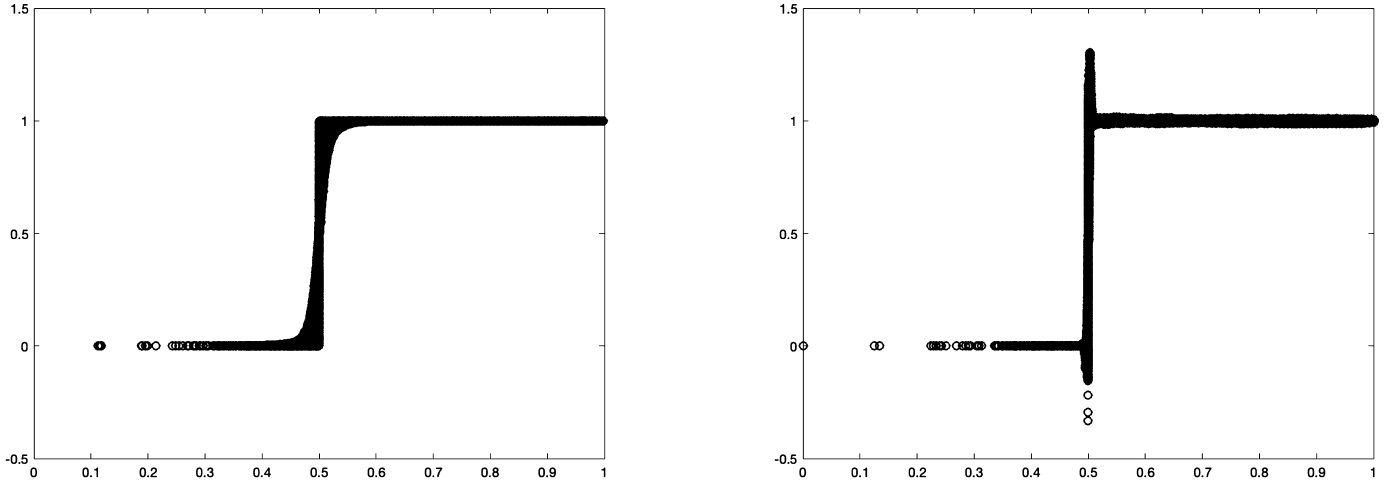
From the figure, we see that the overshoot values oscillate between 0.3 and 0.6 and never go to zero with the mesh refinements.

We plot projected solutions by plotting the numerical solutions of the element center (P_0)/ nodes (P_1) with respect to the radius in Fig. 12, we show the P0-DGFEM projected solution computed on the final adaptive mesh on the left and the P1-DGFEM one on the right. We clearly



(a) initial mesh (b) adaptive mesh (c) os by P1-DGFEM

Fig. 11. P1-DGFEMs for the curved transport problem 1.



(a) P0-DGFEM (b) P1-DGFEM

Fig. 12. Projected solutions with P0 and P1-DGFEMs for curved transport problem 1.

see the overshoot of the P1-DG approximation but the non-overshoot of the P0-DG approximation.

5.3. Curved transport problem 2

Consider the following problem: $\Omega = (0,1)^2$ with $\boldsymbol{\beta} = (y+1,-x)^T/\sqrt{x^2+(y+1)^2}$, $\gamma = 0.1$, and $f = 0$. The inflow boundary is $\{x = 1, y \in (0,1)\} \cup \{x \in (0,1), y = 0\}$, i.e., the west and north boundaries of the domain. Choose g such that the exact solution u is

$$u = \frac{1}{4}\exp\left(\gamma r \arcsin\left(\frac{y+1}{r}\right)\right)\arctan\left(\frac{r-1.5}{\epsilon}\right),$$

with $r = \sqrt{x^2+(y+1)^2}$.

When $\epsilon = 10^{-10}$, the layer is never fully resolved in our experiments and can be viewed as discontinuous. In Fig. 13, we show numerical solutions by P0- and P1-DGFEMs. It is easy to see that there is no overshoot for the P0-DGFEM solution, and a non-trivial overshoot can be found for the P1-DGFEM solution.

Similar results are available for these three numerical tests with $H(\mathrm{div})$ flux-based least-squares finite element formulations can be found in [29,28].

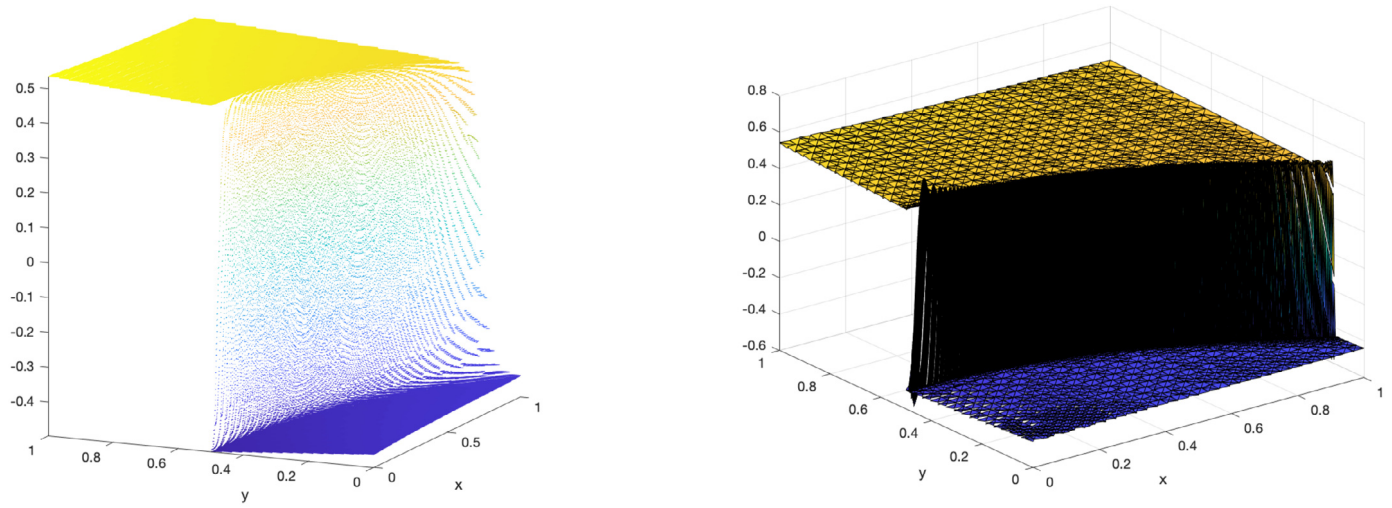
6. Some discussions of numerical tests in the literature

In [3], various numerical methods are tested for the convection-dominated convection-diffusion equations Hemker problem, which contains boundary and interior layers. The interior layers are not matched with the mesh. The authors also conclude that the classic finite volume method has minimal overshoots. The classic finite volume method uses piecewise constant approximation. Combined with adaptive methods, it has a hope of getting an accurate and spurious oscillations-free solution.

In [30], a least-squares-based weak Galerkin finite element method is applied to reaction-diffusion problems. The discontinuous finite element space is used to approximate the primal variable in the method. In a one-dimensional example (Example 7) with an almost discontinuity at the center $x = 0.5$. From Figures 7.6-7.8 of the paper, we can see a grid point on $x = 0.5$. This is the discontinuous approximation in a discontinuity matched case, we should not observe the Gibbs phenomenon, and it matches the numerical test in Figures 7.6-7.8 of [30].

In [27], a weak Galerkin method is introduced to solve convection-dominated convection-diffusion-reaction equations. Its numerical example 4 is $-\epsilon\Delta u + \nabla\cdot(\boldsymbol{\beta}u) + \gamma u = 0$ with discontinuous boundary conditions. For a very small ϵ and a mesh of 1024 uniform squares, it is the linear advection equation (5.1). In Figs. 7-8, the Gibbs phenomenon is observed near the discontinuity. The authors "*believe that a better numerical approximation can be obtained if the Dirichlet boundary conditions are enforced weakly*", while it is most likely due to that the mesh is not aligned with the discontinuity as opposed to [30].

In [7], for the singularly perturbed reaction-diffusion problem, a two-step finite element method is developed. First, the flux is computed independently, then the numerical solution u_h is recovered in P_k. In Theorem 6.5 of [7], $\|u-u_h\|_{L^\infty(K)} \le \|u - P_h u_h\|_{L^\infty(K)}$ + extra terms are proved, where P_h is the L^2-projection. The numerical results show that when $k = 0$ is chosen, then u_h does not show a significant numerical



(a) P0-DGFEM (b) P1-DGFEM

Fig. 13. Numerical solutions with adaptive P0- and P1-DGFEMs for Curved Transport Problem 2.

oscillation. This matches our discussion. On the other hand, if $k \geq 1$ is chosen and the mesh is not aligned with discontinuity, then the term $\|u - P_h u_h\|_{L^\infty(K)}$ will be big, and the Gibbs phenomenon will appear.

7. A short discussion of L^1-based methods

7.1. Discontinuous approximations

For the discontinuous function u defined in (3.1) and an interval I, consider the following L^1 minimization problem

$$\text{Find} \quad w_h \in V_h, \quad \text{such that } \|u - w_h\|_{L^1(I)} \to \min! \tag{7.4}$$

7.1.1. Discontinuous piecewise constant approximations

In this simple case, we only need to compute a constant c on the interval I_0. We should discuss two cases, $|c| > 1$ and $|c| \leq 1$. It is easy to see that the minimizer should happen when $|c| \leq 1$.

Then

$$\|u - c\|_{L^1(I_0)} = \int_{-th}^{(1-t)h} |u - c| dx = h(1 - c + 2tc). \tag{7.5}$$

After a simple discussion of t in two cases, we have:

$$\text{if } 1/2 \leq t \leq 1, \quad \text{then } c = -1 \text{ and } \|u - c\|_{L^1(I_0)} = 2(1-t)h,$$

$$\text{if } 0 \leq t \leq 1/2, \quad \text{then } c = 1 \text{ and } \|u - c\|_{L^1(I_0)} = 2th.$$

Since c is 1 or -1, then there is no overshoot.

7.1.2. Discontinuous piecewise linear approximations

Since the L^1-minimization problem is more complicated than a simple L^2-projection, we only discuss a simple case to show that L^1-minimization can cause an overshoot with a P_1 approximation.

For the interval $I_0 = [-th, (1-t)h]$, we consider the simplest and symmetric case $t = 1/2$. As discussed in Section 3.2.2, let $w_h = W_{-1}\lambda_{-1} + W_1\lambda_1$. Due to symmetry, we can safely assume that $w_h(0) = 0$ and $W_1 = -W_{-1} \geq 0$; thus, we only need to consider the following L^1-minimization problem on the positive half of I_0 (with $\lambda_1 = 2x/h$):

$$\text{Find } W_1 \geq 0 : \|u - w_h\|_{L^1(I_0)} = 2\int_0^{h/2} |u - W_1\lambda_1| dx = h\int_0^1 |1 - W_1 x| dx \to \min! \tag{7.6}$$

The problem can be explicitly solved by discussing two cases: $W_1 \geq 1$ or $0 \leq W_1 \leq 1$:

$$J(W_1) := \int_0^1 |1 - W_1\lambda_1| dx = \begin{cases} 1/W_1 + W_1/2 - 1 & \text{if } W_1 \geq 1, \\ 1 - W_1/2 & \text{if } 0 \leq W_1 \leq 1,. \end{cases} \tag{7.7}$$

A simple calculation shows that

$$\min_{0 \leq W_1} J(W_1) = \sqrt{2} - 1 \quad \text{when} \quad W_1 = \sqrt{2}. \tag{7.8}$$

This shows a non-trivial overshoot appears for $t = 1/2$. For other cases $t \in [0, 1]$, the Gibbs phenomenon still can happen (at least for t close to $1/2$). We skip the detailed computation since we usually cannot control discontinuity location for non-matched cases.

7.2. Continuous approximations

7.2.1. Continuous piecewise linear approximations in 1D

The discussion of L^1-best approximation on 1D can be found in Theorem 1.1 (a boundary discontinuity case) and Theorem 1.4 (a jump discontinuity case) of [18]. The theorem concludes that, under special cases, for example, a uniform grid (or a less restrictive case that the elements cannot be too small compared to their neighboring element closer to the discontinuity) with a **discontinuity matched mesh**, the Gibbs phenomenon can be eliminated. In other cases, the Gibbs phenomenon will happen.

7.2.2. Continuous piecewise linear approximations in 2D

The discussion of L^1-best approximation on 2D can be found in Theorem 1.5 (boundary discontinuity case) and numerical tests of [18]. It is shown in [18] that even for the boundary discontinuity case, which is a discontinuity matched mesh, the L^1-best approximation can only eliminate the Gibbs phenomenon on certain very special meshes.

For the general case of discontinuity non-aligned meshes, the situation is more complicated. From observations of the 1D cases and 2D discontinuity aligned mesh cases in [18], the Gibbs phenomenon still appears for some L^1-best approximations. It is also observed in some numerical tests of [18], that the mesh adaptivity might be helpful in reducing the overshoot, but the underlying mechanism is unclear.

8. A table of results

In Table 1, we list the Gibbs phenomenon in different situations. The computation of an L^2-projection method is relatively easy, a lin-

Table 1
Gibbs Phenomenon in Different Situations (O: overshoot, NO: no overshoot).

mesh	L^2-projection			L^1-best approximation		
	P_0	P_k $(k \geq 1)$	S_k $(k \geq 1)$	P_0	P_k $(k \geq 1)$	S_k $(k \geq 1)$
Discontinuity-aligned (include boundary layer)	NO	NO	O	NO	NO	both
Discontinuity-nonaligned	NO	O	O	NO	O	O

ear equation needs to be solved in an L^2 projection. On the other hand, methods based on L^1-best approximation methods are more complicated, and iterative optimization algorithms are needed; see for example some algorithms suggested in [13]. From the comparisons of the table, we do not recommend such L^1-based methods due to their similar behavior to L^2-based methods and more complicated implementations.

For L^2-minimization-based methods, we have the following conclusions. For the simple discontinuity-aligned mesh case, piecewise discontinuous approximations are always good. For the general non-matched case, the piecewise discontinuous constant approximation combined with adaptive mesh refinements is a good choice to achieve accuracy without overshoots. For discontinuous piecewise linear approximations, non-trivial overshoots will be observed unless the mesh is matched with discontinuity. For continuous piecewise linear approximations, non-trivial overshoots will always be observed under regular meshes.

9. Concluding remarks

In this paper, we study the behavior of using adaptive continuous and discontinuous finite elements to approximate discontinuous and nearly discontinuous PDE solutions from an approximation point of view.

For a singularly perturbed problem with a small diffusion or other equations with a transient layer, the numerical solution can accurately approximate the PDE solution when the mesh is fine enough around the transient layer. This is the case that the belief that the adaptive finite element works come from.

For a pre-asymptotic mesh of such layer problems, the solution can be viewed as discontinuous, and we can treat it as a genuinely discontinuous case.

To approximate a discontinuous solution, the global continuous function is always a bad choice since it will always have non-trivial oscillations. It is also hard to impose a discontinuous boundary condition in a continuous approximation space. If the exact location of the discontinuity is known, then a high-order discontinuous approximation on a discontinuity-matched mesh is a good choice. For singularly perturbed problems with small diffusion, this mesh-matched case is reduced to that the mesh interface lies within the small transient layer. If the location of the discontinuity is unknown, the only way to avoid the overshoot is to use piecewise constant approximations for those elements where a discontinuity cuts through. Discontinuous piecewise linear function approximations will have a non-trivial overshoot for most cases. The adaptive finite element method with a good a posteriori error estimator can reduce the L^2-norms (or H^1 and other integration-based norms) of the error. However, it cannot reduce the non-trivial oscillations unless a piecewise constant approximation is used in the discontinuity-crossing elements.

The non-trivial overshoot we find in the paper is of the same magnitude as the Gibbs phenomenon in the Fourier analysis, where a number around 0.17 of the jump gap is found, see [16].

With the above results in mind, an hp adaptivity approach is preferred when designing finite element methods for problems with unknown location discontinuous solutions. We can use high-order finite elements (continuous or discontinuous) in the smooth region and use piecewise constant approximations in the discontinuity-crossing elements. It is also worth mentioning that the method proposed by Lax [24] is precisely such a method in the context of the finite difference method.

It is also interesting to compare the result to various *order barrier theorems* in the finite difference/finite volume schemes, see [37]. These theorems are developed in a different context, with the finite difference like method, and are proved by methods very different from the discussion in this paper. On the other hand, the discussion in this paper is from the approximation point of view without a specific partial differential equation. It is ongoing research to find the connections between these order barrier theorems and the results of this paper.

Acknowledgements

The author sincerely thanks the anonymous referees for the helpful comments and valuable suggestions, which considerably improved the exposition of this work.

References

[1] M. Ainsworth, J.T. Oden, A Posteriori Error Estimation in Finite Element Analysis, Wiley, 2000.
[2] D.N. Arnold, F. Brezzi, B. Cockburn, L.D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM J. Numer. Anal. (2002) 1749–1779.
[3] M. Augustin, A. Caiazzo, A. Fiebach, J. Fuhrmann, V. John, A. Linke, R. Umla, An assessment of discretizations for convection-dominated convection–diffusion equations, Comput. Methods Appl. Mech. Eng. 200 (2011) 3395–3409.
[4] D. Braess, Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics, Cambridge University Press, 2007.
[5] F. Brezzi, L.D. Marini, E. Süli, Discontinuous Galerkin methods for first-order hyperbolic problems, Math. Models Methods Appl. Sci. 14 (2004) 1893–1903.
[6] E. Burman, A posteriori error estimation for interior penalty finite element approximations of the advection-reaction equation, SIAM J. Numer. Anal. 47 (2009) 3584–3607.
[7] Z. Cai, J. Ku, A dual finite element method for a singularly perturbed reaction-diffusion problem, SIAM J. Numer. Anal. 58 (2020) 1654–1673.
[8] P.G. Ciarlet, Basic error estimates for elliptic problems, in: P. Ciarlet, J.-L. Lions (Eds.), Handbook of Numerical Analysis, vol. II, North-Holland, 1991, pp. 17–351, Ch. Finite Element Methods (Part I).
[9] A. Ern, J.-L. Guermond, Theory and Practice of Finite Elements, Springer, 2004.
[10] S. Franz, H.-G. Roos, The capriciousness of numerical methods for singular perturbations, SIAM Rev. 53 (2011) 157–173.
[11] S.K. Godunov, A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations, Math. Sb. 47 (1959) 271–306.
[12] D. Gottlieb, C.-W. Shu, The Gibbs phenomenon and its resolution, SIAM Rev. 39 (1997) 644–668.
[13] J.-L. Guermond, A finite element technique for solving first order PDE's in L^p, SIAM J. Numer. Anal. 47 (2004) 714–737.
[14] J.S. Hesthaven, Numerical Methods for Conservation Laws: From Analysis to Algorithms, SIAM, 2018.
[15] J.S. Hesthaven, T. Waburton, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, Texts in Applied Mathematics, vol. 54, Springer, 2008.
[16] E. Hewitt, R. Hewitt, The Gibbs-Wilbraham phenomenon: an episode in Fourier analysis, Arch. Hist. Exact Sci. 21 (1979) 129–160.
[17] P. Houston, S. Roggendorf, K.G. van der Zee, Eliminating Gibbs phenomena: a non-linear Petrov-Galerkin method for the convection-diffusion-reaction equation, Comput. Math. Appl. 80 (2020) 851–873, https://doi.org/10.1016/j.camwa.2020.03.025.
[18] P. Houston, S. Roggendorf, K.G. van der Zee, Gibbs phenomena for L^q-best approximation in finite element spaces, ESAIM: M2AN 56 (2022) 177–211.
[19] B.-n. Jiang, Non-oscillatory and non-diffusive solution of convection problems by the iteratively reweighted least-squares finite element method, J. Comput. Phys. 105 (1993) 108–121.
[20] C. Johnson, J. Pitkäranta, An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation, Math. Comput. 46 (1986) 1–26.
[21] G. Kunert, A note on the energy norm for a singularly perturbed model problem, Computing 69 (2002) 265–272.
[22] J.E. Lavery, Non-oscillatory solution of the steady-state inviscid Burgers equation by mathematical programming, J. Comput. Phys. 79 (1988) 436–448.

[23] J.E. Lavery, Non-oscillatory solution of the steady-state inviscid Burgers equation by mathematical programming, SIAM J. Numer. Anal. 26 (1989) 1081–1089.
[24] P.D. Lax, Gibbs phenomena, J. Sci. Comput. 28 (2006) 445–449.
[25] J. Li, L. Demkowicz, An L^p-DPG method for the convection–diffusion problem, Comput. Math. Appl. 95 (2021) 172–185, https://doi.org/10.1016/j.camwa.2020.08.013.
[26] R. Lin, Discontinuous discretization for least-squares formulation of singularly perturbed reaction-diffusion problems in one and two dimensions, SIAM J. Numer. Anal. 47 (2008) 89–108.
[27] R. Lin, X. Ye, S. Zhang, P. Zhu, A weak Galerkin finite element method for singularly perturbed convection-diffusion–reaction problems, SIAM J. Numer. Anal. 56 (2018) 1482–1497.
[28] Q. Liu, S. Zhang, Adaptive flux-only least-squares finite element methods for linear transport equations, J. Sci. Comput. 84 (26) (2020), https://doi.org/10.1007/s10915-020-01269-y.
[29] Q. Liu, S. Zhang, Adaptive least-squares finite element methods for linear transport equations based on an H(div) flux reformulation, Comput. Methods Appl. Mech. Eng. 366 (2020) 113041, https://doi.org/10.1016/j.cma.2020.113041.
[30] L. Mu, J. Wang, X. Ye, A least-squares-based weak Galerkin finite element method for second order elliptic equations, SIAM J. Sci. Comput. 39 (2017) A1531–A1557.
[31] J. Nitsche, Uber ein variationsprinzip zur losung von dirichlet-problemen bei verwendung von teilraumen, die keinen Randbedingungen unterworfen sind, Abh. Math. Semin. Univ. Hamb. 36 (1971) 9–15.
[32] H.-G. Roos, M. Stynes, L. Tobiska, Robust Numerical Methods for Singularly Perturbed Differential Equations, 2 ed., Springer Ser. Comput. Math., Springer, 2008.
[33] E.B. Saff, S. Tashev, Gibbs phenomenon for best lp approximation by polygonal lines, East J. Approx. 5 (1999) 235–251.
[34] R. Verfürth, Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation, Numer. Math. 78 (1998) 479–493.
[35] R. Verfürth, A Posteriori Error Estimation Techniques for Finite Element Methods, Oxford University Press, 2013.
[36] L.B. Wahlbin, Local behavior in finite element methods, in: P.G. Ciarlet, J.-L. Lions (Eds.), Handbook of Numerical Analysis, vol. 2, North-Holland, 1991, pp. 353–522, Ch. Finite Element Methods (Part I).
[37] P. Wesseling, Principles of Computational Fluid Dynamics, Springer Series in Computational Mathematics, vol. 29, Springer, 2001.