

IMPROVED LEAST-SQUARES ERROR ESTIMATES FOR SCALAR HYPERBOLIC PROBLEMS,¹

P.B. BOCHEV

E-mail: bochev@uta.edu

J. CHOI

Department of mathematics, The University of Texas at Arlington
Box 19408, Arlington, TX 76019-0408

Abstract. We consider an L^2 -norm least-squares principle for a scalar hyperbolic problem. A proper variational framework for the associated finite element method is developed and studied. Analysis of the discretization error based on the least-squares projection property shows a gap (see [8]) of one. This number cannot be improved with a standard duality argument because the least-squares dual does not possess full elliptic regularity. Using a perturbed dual problem we are able to show that the actual gap of the least-squares method in the constant convection case is not worse than $2/3$.

2000 Mathematics Subject Classification: 65N30, 65N12.

Key words: finite elements, convection problems, least-squares..

1. Introduction

The error of a finite element approximation depends on two factors: the regularity of the exact solution and the polynomial degree of the finite element space. An *optimal* error estimate can be stated symbolically as

$$\|\text{error}\|_0 \leq Ch^r \|\text{exact solution}\|_r; \quad (1.1)$$

where k is the polynomial degree of the finite element space and $\|\cdot\|_s$ denotes the usual Sobolev space norm². An optimal error estimate establishes the fact that a finite element approximation of a sufficiently regular solution converges at the same asymptotic rate as the interpolant of this solution. Ability to generate optimally convergent approximations in various Sobolev space norms is not an intrinsic property of the finite element method. There are instances when even under ideal circumstances the approximation will converge at a slower rate than the solution interpolant. In such cases the relevant error estimate becomes

$$\|\text{error}\|_0 \leq Ch^s \|\text{exact solution}\|_r. \quad (1.2)$$

where $s < r$. The notion of the *gap* as a measure of a method's deficiency relative to the interpolation error has been introduced in [8] by Johnson et. al.. Specifically, the gap equals the difference $r - s$ between the potential convergence rate afforded by the solution regularity and the actual rate of the method.

The first kind of error estimates, i.e., optimal errors with zero gap, typically arise in the context of elliptic equations, while finite element methods for hyperbolic problems tend to produce approximations with a positive gap; see [8]. A typical example of such a problem is the scalar convection equation (often referred to as the *reduced problem*)

$$u_{\beta} + cu = f \quad \text{in } \Omega, \quad (1.3)$$

$$u = g \quad \text{on } \Gamma_-, \quad (1.4)$$

¹This work was sponsored by the National Science Foundation under grant number DMS-0073698.

²Throughout the paper we adhere to standard Sobolev space notations such as (\cdot, \cdot) and $|\cdot|_k$ for the L^2 inner product and $H^k(\Omega)$ seminorm, respectively. Bold face symbols will denote vector valued functions.

that will be in the focus of the present work. In (1.3)-(1.4) Ω is a bounded convex region in R^2 , $\boldsymbol{\beta} = (\beta_1(x), \beta_2(x))$ is a given convection vector, $c(x)$ is a bounded measurable function on Ω , $f \in L^2(\Omega)$ is the source term, and $u_{\boldsymbol{\beta}} = \boldsymbol{\beta} \cdot \nabla u$ is the derivative of u in the direction of $\boldsymbol{\beta}$. As usual,

$$\Gamma_- = \{x \in \Gamma | \mathbf{n}(x) \cdot \boldsymbol{\beta}(x) < 0\} \quad (1.5)$$

denotes the inflow portion of $\Gamma = \partial\Omega$. It is well-known that the gap of the standard Galerkin method for (1.3)-(1.4) is one, while the gap of SUPG, see [7], is $1/2$. In section 2 we carry analysis of a least-squares method for the model problem using standard arguments based on the projection property of the least-squares principle. The gap revealed by this analysis equals one, i.e., the same as for the Galerkin method. On the other hand, there is computational evidence that for smooth solutions the least-squares method will perform at a level comparable with that of SUPG; see [1], [5] and [3]. Moreover, in [9] Lazarov et. al. have shown that the gap of a least-squares method for the full convection-diffusion problem is the same as in SUPG. This prompts the question whether the theoretical gap of the least-squares method for the reduced problem (1.3)-(1.4) can't be improved³. A standard way of doing this would be to pursue a duality argument (Nitsche's trick). However, in section 3 we show that the least-squares dual problem is in fact a coupled system of two hyperbolic problems. As a result, this problem does not possess the elliptic regularity necessary to carry out the usual duality argument. To circumvent this problem we proceed to define a regularized dual problem obtained by adding a small amount of crosswind diffusion and a Neumann boundary condition.

For rectangular domains and constant advection in one of the coordinate directions this enables us to improve the gap of the least-squares method to $2/3$. While numerical evidence strongly suggests that such a result may actually be valid for more general domains and advection fields, at present we are not aware whether or not the technique used in this paper can be extended to such cases.

We also point out that while the main result in this paper is established for rectangular domains, it remains valid for any regular triangulation of this domain into finite elements. It is also clear that regardless of the grids and elements employed, e.g., triangles, rectangles or isoparametric bricks, the least-squares scheme developed here always leads to symmetric and positive algebraic systems for the unknown coefficients, which is one of the valuable computational advantages of least-squares compared with other discretization schemes.

2. The least-squares method

In this paper we focus on an L^2 -norm least-squares method for the reduced problem (1.3)-(1.4) in which the boundary condition is imposed on the trial space. In this form the method has been introduced by Jiang; see [5]. Even though the method itself is not new, it still lacks a proper variational formulation which, as we show below, differs from the setting used for the full problem; see e.g., [9]. Thus, we first proceed to establish the proper mathematical framework for the least-squares principle⁴, that is, the relevant function spaces and norms. For simplicity let $c \equiv 1$ and consider the space

$$\mathcal{V} = \{v \in L^2(\Omega) | v_{\boldsymbol{\beta}} + v \in L^2(\Omega)\}. \quad (2.1)$$

\mathcal{V} is Hilbert space when equipped with the graph norm

$$\|v\|_{\mathcal{G}} = \left(\|v\|_0^2 + \|v_{\boldsymbol{\beta}} + v\|_0^2 \right)^{1/2}. \quad (2.2)$$

It is also known that functions in \mathcal{V} have well-defined traces on $L^2(\Gamma_-)$ and $L^2(\Gamma_+)$, where $\Gamma_+ = \Gamma - \Gamma_-$; see, e.g., [6], so that

$$\mathcal{V}_0 = \{v \in \mathcal{V} | v = 0 \text{ on } \Gamma_-\}. \quad (2.3)$$

is a well-defined subspace of \mathcal{V} . Then, a proper least-squares principle for the model problem is to seek a minimizer of the quadratic functional

$$\mathcal{J}(u; f, g) = \frac{1}{2} \|u_{\boldsymbol{\beta}} + u - f\|_0^2 \quad (2.4)$$

out of the space \mathcal{V}_0 . It is implicitly assumed that inhomogeneous boundary data has been removed in the usual manner by writing $u = u_0 + u_g$ where $u_0 \in \mathcal{V}_0$ is the unknown part of the solution that vanishes on Γ_- and $u_g = g$ on Γ_- is a given function in \mathcal{V} . This amounts to solving (1.3)-(1.4) with homogeneous boundary

³The gap of the Galerkin method is genuine in the sense that the suboptimal convergence rate is observed in actual computations.

⁴For analysis of an alternative method with weakly imposed boundary condition we refer to [3].

data and a modified source term. In a finite element implementation the role of u_g is played by the boundary interpolant of g . For simplicity we retain the symbols u and f to denote the redefined unknown and source term, respectively.

The necessary minimum condition for (2.4) is to seek $u \in \mathcal{V}_0$ such that

$$\mathcal{B}_L(u; v) = \mathcal{F}_L(v) \quad \text{for all } v \in \mathcal{V}_0. \quad (2.5)$$

The bilinear form and right hand side functional in (2.5) are

$$\begin{aligned} \mathcal{B}_L(u; v) &= (u_{\boldsymbol{\beta}} + u, v_{\boldsymbol{\beta}} + v), \\ \mathcal{F}_L(v) &= (f, v_{\boldsymbol{\beta}} + v). \end{aligned}$$

Our next task is to show that the least-squares principle gives rise to a *problem-dependent* inner product and norm that are equivalent to the graph norm and inner product on \mathcal{V}_0 . As usual, we refer to these as the *energy* inner product and norm.

Theorem 2.1. *Assume that $1 - \nabla \cdot \boldsymbol{\beta} \geq \sigma > 0$. Then, the energy norm*

$$|||v|||_E = \mathcal{B}_L(v; v)^{1/2} \equiv \|v_{\boldsymbol{\beta}} + v\|_0 \quad (2.6)$$

is equivalent to the graph norm, that is, there exist positive constants C_1 and C_2 such that

$$C_1 |||v|||_E \leq |||v|||_G \leq C_2 |||v|||_E \quad \text{for all } v \in \mathcal{V}_0. \quad (2.7)$$

Proof. The lower bound is a simple consequence of the energy norm definition (2.6). To prove the upper bound it suffices to show that

$$C(\|u_{\boldsymbol{\beta}}\|_0^2 + \|u\|_0^2) \leq \|u_{\boldsymbol{\beta}} + u\|_0^2 = \|u_{\boldsymbol{\beta}}\|_0^2 + \|u\|_0^2 + 2(u, u_{\boldsymbol{\beta}}).$$

This easily follows from the Green's formula,

$$2 \int_{\Omega} u_{\boldsymbol{\beta}} u \, dx = - \int_{\Omega} u^2 \nabla \cdot \boldsymbol{\beta} \, d\Gamma + \int_{\Gamma_+} u^2 \boldsymbol{\beta} \cdot \mathbf{n} \, d\Gamma,$$

the assumption that $1 - \nabla \cdot \boldsymbol{\beta} \geq \sigma$, and the fact that $\boldsymbol{\beta} \cdot \mathbf{n} \geq 0$ on Γ_+ . \square

Corollary 2.1. *The weak least-squares problem has a unique solution $u \in \mathcal{V}_0$. Moreover,*

$$|||u|||_E \leq \|f\|_0.$$

Proof. The norm equivalence (2.7) implies that $\mathcal{B}_L(\cdot; \cdot)$ is coercive and continuous on $\mathcal{V}_0 \times \mathcal{V}_0$. It also easy to see that $\mathcal{F}_L(\cdot)$ is a bounded linear functional $\mathcal{V}_0 \mapsto \mathbf{R}$. Thus, existence and uniqueness of a weak solution follows by virtue of Riesz representation theorem. Stability of the solution follows from (2.5) with $v \equiv u$ and the Cauchy's inequality. \square

To define a least-squares finite element method for the model problem we shall make use of the standard Lagrangian finite element spaces

$$P_k = \{u^h \in C^0(\Omega) \mid u^h|_K \in \mathcal{P}_k, \forall K \in \mathcal{T}_h\}, \quad (2.8)$$

where \mathcal{P}_k denotes the space of algebraic polynomials in x and y whose degree does not exceed k . It is assumed that P_k are defined with respect to a uniformly regular partition \mathcal{T}_h of Ω into triangles K . Next, we introduce a finite element subspace of \mathcal{V}_0

$$\mathcal{V}^h = \{u^h \in P_k \mid u^h = 0 \quad \text{on } \Gamma_-\}. \quad (2.9)$$

The discrete problem then is to seek $u^h \in \mathcal{V}^h$ such that

$$\mathcal{B}_L(u^h; v^h) = \mathcal{F}_L(v^h) \quad \text{for all } v^h \in \mathcal{V}^h. \quad (2.10)$$

Thanks to the inclusion $\mathcal{V}^h \subset \mathcal{V}_0$ the lower bound in (2.7) remains valid when $\mathcal{B}_L(\cdot; \cdot)$ is restricted to $\mathcal{V}^h \times \mathcal{V}^h$. Consequently, (2.10) is a linear algebraic system with a symmetric and positive definite matrix and the weak discrete problem will have a unique solution. As least-squares methods share similar variational foundations with classical Rayleigh-Ritz principles, optimal error estimates in the energy norm can be obtained in a completely standard manner.

Table 1: Convergence rate for $u = \exp(x) \sin(x) \sin(y)$: P2 elements.

	L^2 error	H^1 error	Streamline	Energy
Galerkin	2.0005	1.0073	1.0008	2.0005
SUPG	3.0000	2.0024	1.9970	2.5098
LSFEM	3.0055	2.0087	2.0377	2.0026

Theorem 2.2. *Assume that $u \in \mathcal{V}_0$ solves (1.3)-(1.4) and let $u^h \in \mathcal{V}^h$ denote the least-squares approximate solution. Then*

$$\| \|u - u^h\| \|_E = \inf_{v^h \in \mathcal{V}^h} \| \|u - v^h\| \|_E. \quad (2.11)$$

If, in addition $u \in H^{k+1}(\Omega) \cap \mathcal{V}_0$, then

$$\| \|u - u^h\| \|_E \leq Ch^k |u|_{k+1}. \quad (2.12)$$

Proof. A fundamental property of the least-squares method is that u^h is the orthogonal projection of the exact solution with respect to the problem-dependent energy inner product $\mathcal{B}_L(\cdot; \cdot)$. This immediately implies the first assertion of the theorem. To prove the second statement consider the interpolant $\tilde{u}^h \in \mathcal{V}^h$ of the exact solution. Then, (see e.g., [2])

$$\| \|u - \tilde{u}^h\| \|_0 + h|u - \tilde{u}^h|_1 \leq Ch^{k+1} |u|_{k+1}.$$

This, (2.11) and the triangle's inequality establish (2.12). \square

From the proof of Theorem 2.1 we know that the least-squares energy norm controls the L^2 -norms of the streamline derivative u_β and the solution. Thus, (2.12) translates into the optimal error bound

$$\| \|u_\beta - u_\beta^h\| \|_0 = O(h^k),$$

and the suboptimal estimate

$$\| \|u - u^h\| \|_0 = O(h^k).$$

Therefore, based on Theorem 2.2 we can only infer that the gap of the least-squares method equals one. This gap equals the gap of the Galerkin method and is half order worse than the gap of the SUPG method where; see [8],

$$\| \|u - u_{SUPG}^h\| \|_0 \leq Ch^{k+1/2} |u|_{k+1}.$$

At the same time, numerical studies in [1], [3] strongly suggest that the actual gap of the least-squares method is less than one. Table 1 contains a typical example of the convergence rates that we have observed and reported in [1]. The gap of the Galerkin method is clearly seen, while least-squares and SUPG errors are essentially the same. Jiang; see [5], also notes that, as a rule, for smooth solutions the least-squares L^2 error is almost always optimal.

3. Improved L^2 -norm error estimates

To improve the L^2 error estimate we will use a duality argument. However, the seemingly natural choice of the dual problem: seek $v \in \mathcal{V}_0$ such that

$$\mathcal{B}_L(u; v) = (\psi, u) \quad \text{for all } u \in \mathcal{V}_0, \quad (3.1)$$

is not appropriate for this purpose. To see this define

$$\Gamma_{++} = \{x \in \Gamma | \mathbf{n}(x) \cdot \beta(x) > 0\},$$

and let $\Gamma_0 = \Gamma - \Gamma_- \cup \Gamma_{++}$. Also, let $\boldsymbol{\eta}$ denote a vector field that is orthogonal to β . Then, the normal direction on Γ_0 coincides with $\boldsymbol{\eta}$. After formal integration by parts in (3.1) we find that the strong dual problem is

$$-\nabla \cdot (\beta(v_\beta + v)) + (v_\beta + v) = \psi \quad \text{in } \Omega \quad (3.2)$$

$$v = 0 \quad \text{on } \Gamma_- \quad (3.3)$$

$$v_\beta + v = 0 \quad \text{on } \Gamma_{++}. \quad (3.4)$$

Eqs. (3.2)-(3.4) are a degenerate elliptic problem. As such, the differential operator in (3.2) can only control the second derivatives along the streamlines but not over the whole region. In fact, it is easy to see that (3.2)-(3.4) can be written as a coupled system of two hyperbolic PDE's (see [6] for a similar situation arising in the context of SUPG)

$$\begin{aligned} -\nabla \cdot (\boldsymbol{\beta}z) + z &= \psi \quad \text{in } \Omega \\ z &= 0 \quad \text{on } \Gamma_{++}, \end{aligned}$$

$$\begin{aligned} \boldsymbol{\beta} \cdot \nabla v + v &= z \quad \text{in } \Omega \\ v &= 0 \quad \text{on } \Gamma_-. \end{aligned}$$

Then, using Rauch's differentiability theorem [10] one can show that for a smooth ψ ,

$$\|v\|_k \leq \|\psi\|_k.$$

This result is sharp and indeed means that the least-squares dual problem does not possess the necessary elliptic regularity. To remedy this situation let us perturb (3.2)-(3.4) by adding the crosswind diffusion term $-\varepsilon v \boldsymbol{\eta} \boldsymbol{\eta}$ and, provided Γ_0 is not empty, the Neumann condition $\varepsilon v \boldsymbol{\eta} = 0$ on Γ_0 . After some manipulations the regularized dual problem reads

$$-\varepsilon v \boldsymbol{\eta} \boldsymbol{\eta} - v \boldsymbol{\beta} \boldsymbol{\beta} - v \boldsymbol{\beta} \nabla \cdot \boldsymbol{\beta} + (1 - \nabla \cdot \boldsymbol{\beta})v = \psi \quad \text{in } \Omega \quad (3.5)$$

$$v = 0 \quad \text{on } \Gamma_- \quad (3.6)$$

$$v \boldsymbol{\beta} + v = 0 \quad \text{on } \Gamma_{++} \quad (3.7)$$

$$v \boldsymbol{\eta} = 0 \quad \text{on } \Gamma_0. \quad (3.8)$$

We now turn to develop an improved L^2 -error estimate. To establish the elliptic regularity of (3.5)-(3.8) we adopt the same setting as in [4, p.471]. Specifically, we consider $\boldsymbol{\beta} = (1, 0)^T$, $\Omega = (0, 1) \times (0, 1)$ so that

$$\boldsymbol{\beta} \cdot \nabla u = u_x$$

and

$$\begin{aligned} \Gamma_- &= \Gamma_1, \\ \Gamma_{++} &= \Gamma_3 \\ \Gamma_0 &= \Gamma_2 \cup \Gamma_4, \end{aligned}$$

where $\Gamma_1, \Gamma_3, \Gamma_2$ and Γ_4 denote the left and right and the bottom and top walls of Ω . With these assumptions our model problem specializes to

$$u_x + u = f \quad \text{in } \Omega, \quad (3.9)$$

$$u = g \quad \text{on } \Gamma_-, \quad (3.10)$$

while the least-squares dual (3.2)-(3.4) takes the form

$$\begin{aligned} -v_{xx} + v &= \psi \quad \text{in } \Omega \\ v &= 0 \quad \text{on } \Gamma_1 \\ v_x + v &= 0 \quad \text{on } \Gamma_3. \end{aligned}$$

Likewise, the regularized dual is

$$-\varepsilon v_{yy} - v_{xx} + v = \psi \quad \text{in } \Omega \quad (3.11)$$

$$v = 0 \quad \text{on } \Gamma_1 \quad (3.12)$$

$$v_y = 0 \quad \text{on } \Gamma_2 \cup \Gamma_4 \quad (3.13)$$

$$v_x + v = 0 \quad \text{on } \Gamma_3. \quad (3.14)$$

Next, the full elliptic regularity of (3.11)-(3.14) will be established.

Lemma 3.1. *Assume that v is a smooth function such that (3.12)-(3.14) hold. Then*

$$|v|_2 \leq \|\Delta v\|_0. \quad (3.15)$$

Proof. To prove the lemma we need to show that

$$\|\Delta v\|_0^2 - |v|_2^2 = 2 \int_{\Omega} (v_{xx}v_{yy} - v_{xy}^2) dx \geq 0.$$

Integrating by parts twice and accounting for the fact that $n_2 = 0$ on $\Gamma_1 \cup \Gamma_3$ and $n_1 = 0$ on $\Gamma_2 \cup \Gamma_4$ yields the identity

$$\int_{\Omega} v_{xx}v_{yy} dx = \int_{\Omega} v_{xy}^2 dx + \int_{\Gamma_2 \cup \Gamma_4} v_{xx}v_y n_2 d\Gamma - \int_{\Gamma_1 \cup \Gamma_3} v_{xy}v_y n_1 d\Gamma.$$

The first boundary integral vanishes thanks to the Neumann boundary condition (3.13). The integral along Γ_1 is zero thanks to the Dirichlet condition (3.12) as v_y is the tangential derivative of $v(0, y) \equiv 0$. Lastly, the Robin condition (3.14) means that $v = -v_x$ along Γ_3 and

$$- \int_{\Gamma_3} v_{xy}v_y d\Gamma = \int_{\Gamma_3} (-v_x)_y v_y d\Gamma = \int_{\Gamma_3} v_y^2 d\Gamma \geq 0.$$

As a result,

$$\int_{\Omega} (v_{xx}v_{yy} - v_{xy}^2) dx = \int_{\Gamma_3} v_y^2 d\Gamma \geq 0$$

which proves the lemma. \square

Theorem 3.1. *Solution of the dual problem (3.11)-(3.14) satisfies the stability estimates*

$$\sqrt{\varepsilon} \|\nabla v\|_0 \leq \|\psi\|_0, \quad (3.16)$$

$$\|\Delta_{\varepsilon} v\|_0 \leq \|\psi\|_0, \quad (3.17)$$

where $\Delta_{\varepsilon} v = v_{xx} + \varepsilon v_{yy}$

Proof. Multiplying (3.11) by v , integrating by parts and noting that $n_2 = 0$ on $\Gamma_1 \cup \Gamma_3$ and $n_1 = 0$ on $\Gamma_2 \cup \Gamma_4$ gives the identity

$$\|v_x\|_0^2 + \varepsilon \|v_y\|_0^2 + \|v\|_0^2 - \int_{\Gamma_2 \cup \Gamma_4} v v_y n_2 d\Gamma - \int_{\Gamma_1 \cup \Gamma_3} v v_x n_1 d\Gamma = \int_{\Omega} v \psi dx.$$

The integral along $\Gamma_2 \cup \Gamma_4$ vanishes thanks to the Neumann condition (3.13); the integral along Γ_1 is zero because of the Dirichlet condition (3.12) and the remaining integral along Γ_3 gives a nonnegative contribution thanks to the Robin condition $v_x = -v$. As a result,

$$\varepsilon \|\nabla v\|_0 \leq \|v_x\|_0^2 + \varepsilon \|v_y\|_0^2 + \|v\|_0^2 \leq \|\psi\|_0 \|v\|_0$$

from where (3.16) easily follows.

To prove the second inequality, multiply (3.11) by $-\Delta_{\varepsilon} v$ and integrate over Ω to obtain

$$\|\Delta_{\varepsilon} v\|_0^2 + \int_{\Omega} v(-v_{xx} - \varepsilon v_{yy}) dx = \int_{\Omega} (-\Delta_{\varepsilon} v) \psi dx.$$

Consider the term $\int_{\Omega} v(-\varepsilon v_{yy}) dx$. Integrating by parts, using the Neumann condition (3.13) and the fact that $n_2 = 0$ on $\Gamma_1 \cup \Gamma_3$ gives

$$\int_{\Omega} v(-\varepsilon v_{yy}) dx = \varepsilon \|v_y\|_0^2 - \varepsilon \int_{\Gamma_2 \cup \Gamma_4} v v_y d\Gamma \varepsilon \|v_y\|_0^2.$$

For the term $\int_{\Omega} v(-v_{xx}) dx$ integration by parts and $n_1 = 0$ on $\Gamma_2 \cup \Gamma_4$ gives

$$\begin{aligned} \int_{\Omega} v(-v_{xx}) dx &= \|v_x\|_0^2 + \int_{\Gamma_1} v v_x d\Gamma - \int_{\Gamma_3} v v_x d\Gamma \\ &= \|v_x\|_0^2 + \int_{\Gamma_3} v^2 d\Gamma. \end{aligned}$$

Again, the integral along Γ_1 vanishes thanks to the Dirichlet condition (3.12), while the integral along Γ_3 turns out to be nonnegative thanks to the Robin condition (3.14). The desired bound (3.17) now can be easily obtained from the resulting identity

$$\|\Delta_{\varepsilon} v\|_0^2 + \|v_x\|_0^2 + \varepsilon \|v_y\|_0^2 + \int_{\Gamma_3} v^2 d\Gamma = \int_{\Omega} (-\Delta_{\varepsilon} v) \psi dx.$$

\square

Corollary 3.1. For smooth solutions of (3.11)-(3.14)

$$|v|_2 \leq \frac{1}{\varepsilon} \|\Delta_\varepsilon v\|_0. \quad (3.18)$$

The next lemma is a version of the duality argument, (*Nitsche's trick*), specialized to our needs.

Lemma 3.2. Assume that $u \in H^{k+1}(\Omega)$ solves (3.9)-(3.10), and that u^h is the least-squares approximation. Then, there exists a constant C such that

$$\|u - u^h\|_0 \leq Ch^{\frac{1}{3}} |u - u^h|_1. \quad (3.19)$$

Proof. Let $e = u - u^h \in \mathcal{V}_0$ denote the least-squares approximation error and consider a solution v of the regularized dual problem with $\psi = e$ so that $e = -\Delta_\varepsilon v + v$. Taking under consideration the boundary conditions on v and e yields the identity

$$\begin{aligned} \|e\|_0^2 &= (e, -\Delta_\varepsilon v + v) \\ &= -\varepsilon(e, v_{yy}) + (e, -(v_x + v)_x + (v_x + v)) \\ &= \varepsilon(e_y, v_y) + (e_x + e, v_x + v) - \int_\Gamma e v_y n_2 \, d\Gamma - \int_\Gamma e (v_x + v) n_1 \, d\Gamma \\ &= \varepsilon(e_y, v_y) + \mathcal{B}_L(e; v). \end{aligned}$$

The projection property of the least-squares weak problem means that $\mathcal{B}_L(e; v) = \mathcal{B}_L(e; v - z^h)$ for any $z^h \in \mathcal{V}^h$. As a result,

$$\begin{aligned} \|e\|_0^2 &= \varepsilon(e_y, v_y) + \mathcal{B}_L(e; v - z^h) \\ &\leq \varepsilon \|e_y\|_0 \|v_y\|_0 + \|e\|_E \|v - z^h\|_E \quad \forall z^h \in \mathcal{V}^h. \end{aligned}$$

Since both $v - z^h$ and e vanish on a part of Γ with a positive measure, the Friedrichs-Poincare inequality (see [2]) can be used to show that

$$\|e\|_E \leq |e|_1 \quad \text{and} \quad \|v - z^h\|_E \leq |v - z^h|_1,$$

and therefore,

$$\|e\|_0^2 \leq \sqrt{\varepsilon} |e|_1 \sqrt{\varepsilon} \|v_y\|_0 + |e|_1 |v - z^h|_1.$$

To estimate the second term in this inequality we set z^h equal to the interpolant \tilde{v}^h of v . The definition of v combined with (3.18) yields

$$|v - \tilde{v}^h|_1 \leq Ch |v|_2 \leq C \frac{h}{\varepsilon} \|\Delta_\varepsilon v\|_0 \leq C \frac{h}{\varepsilon} \|e\|_0.$$

Likewise, from (3.16) we see that

$$\sqrt{\varepsilon} \|v_y\|_0 \leq \sqrt{\varepsilon} \|\nabla v\|_0 \leq \|\psi\|_0 = \|e\|_0.$$

Therefore,

$$\|e\|_0 \leq C \left(\sqrt{\varepsilon} + \frac{h}{\varepsilon} \right) |e|_1.$$

The minimum of $(\sqrt{\varepsilon} + \frac{h}{\varepsilon})$ occurs when $\varepsilon = (2h)^{\frac{2}{3}}$ which establishes the final result. \square

Combining the Lemma 3.2 with the *a priori* error estimate in Theorem 2.2, gives the improved L^2 -norm error estimate.

Theorem 3.2. Assume that the condition of the Lemma 3.2 holds. Then, there exists a constant C such that

$$\|u - u^h\|_0 \leq Ch^{k+\frac{1}{3}} |u|_{k+1}. \quad (3.20)$$

Acknowledgement

The authors wish to thank the anonymous referees for their suggestions and recommendations.

References

- [1] P. Bochev and J. Choi. A Comparative Numerical Study of Least-Squares, SUPG and Galerkin Methods for Convection Problems. *Int. Jour. Comp. Fluid Dynamics*, to appear..
- [2] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, 1994.
- [3] J. M. Choi. *The least-squares method for hyperbolic problems*. Doctoral Thesis, The University of Texas at Arlington, 2000.
- [4] K. Eriksson, D. Estep, P. Hansbo and C. Johnson. *Computational Differential Equations*. Cambridge University Press, 1996.
- [5] B. N. Jiang. Non-oscillatory and non-diffusive solution of convection problems by the iteratively reweighted least-squares finite element method. *J. Comp. Phys.*, **105** (1), (1993), 108–121.
- [6] P. Houston, R. Rannacher and E. Suli. A posteriori error analysis for stabilized finite element approximations of transport problems. *Report 99/04, Oxford University Computing Laboratory, Oxford, England*, 1999.
- [7] T.J.R. Hughes and A. Brooks. A theoretical framework for Petrov-Galerkin methods with discontinuous weighting functions: Application to the streamline-upwind procedure. In: *Finite Elements in Fluids 4*, edited by R.H. Gallagher et al, J. Willey & Sons , , 1982, 47–65..
- [8] C. Johnson, U. Navert and J. Pitkaranta. Finite element methods for linear hyperbolic problems. *Comput. Meth. Appl. Mech. Engrg.*, **45** 1984, 285–312.
- [9] R.D. Lazarov, L. Tobiska, and P. S. Vassilevski. Streamline-diffusion least-squares mixed finite element methods for convection-diffusion problems. *East-West J. Numer. Math*, **5** (4), 1997, 249-264.
- [10] J. Rauch. L_2 is a continuable initial condition for Kreiss' mixed problems. *Comm. Pure Appl. Math.*, **25** 1972, 265–285.