



Balanced-Norm and Energy-Norm Error Analyses for a Backward Euler/FEM Solving a Singularly Perturbed Parabolic Reaction-Diffusion Problem

Xiangyun Meng¹ · Martin Stynes²

Received: 17 November 2021 / Revised: 9 May 2022 / Accepted: 29 June 2022 /
Published online: 8 July 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

In the derivation of error bounds, uniformly in the singular perturbation parameter ε , for finite element methods (FEMs) applied to elliptic singularly perturbed linear reaction-diffusion problems, the usual energy norm is unsatisfactory since it is essentially no stronger than the L^2 norm. Consequently various researchers have analysed errors in FEM solutions, uniformly in ε , using balanced norms whose H^1 component is weighted correctly to maintain its influence. But the derivation of energy and balanced-norm error bounds for FEM solutions of singularly perturbed reaction-diffusion problems is confined almost entirely to steady-state elliptic problems — little has been proved for time-dependent parabolic singularly perturbed problems. The present paper addresses this gap in the literature: the backward Euler method in time, combined with a bilinear FEM on a spatial Shishkin mesh, is applied to solve a parabolic singularly perturbed reaction-diffusion problem, and energy-norm and balanced-norm error estimates, which are uniform in the singular perturbation parameter ε , are derived — these results are stronger than any previous results of the same type. Furthermore, numerical experiments demonstrate the sharpness of our error bounds.

Keywords Energy norm · Balanced norm · Singularly perturbed · Reaction-diffusion · Parabolic

Mathematics Subject Classification 65M15 · 65M60

The research of Xiangyun Meng is supported in part by the Fundamental Research Funds for the Central Universities under grant 2020RC101 and the National Natural Science Foundation of China under grants 12101039. The research of Martin Stynes is supported in part by the National Natural Science Foundation of China under grants 12171025 and NSAF-U1930402.

✉ Martin Stynes
m.stynes@csrc.ac.cn

Xiangyun Meng
xymeng1@bjtu.edu.cn

¹ School of Mathematics and Statistics, Beijing Jiaotong University, Beijing 100044, China

² Applied and Computational Mathematics Division, Beijing Computational Science Research Center, Beijing 100193, China

1 Introduction

The numerical solution of elliptic singularly perturbed linear reaction-diffusion problems by finite element methods (FEMs) has been extensively researched; see [14, 20, 22]. In particular, the FEM solution of such problems using Shishkin meshes on the unit square in \mathbb{R}^2 has been well understood for some time; for example, an optimal-order energy-norm error analysis that is uniform in the singular perturbation parameter ε is given in [16]. But although the energy norm H_e^1 (an ε -weighted H^1 norm — see Sect. 5) seems a natural choice for FEM error analysis, it was pointed out in [13] that it is scaled incorrectly when the problem is singularly perturbed: typically, its H^1 component of the error is dominated by its L^2 component, so an energy-norm error bound is in practice only an L^2 error bound.

As a consequence, in the last 10 years several papers (see [3, 7, 8, 13, 17–19] and their references) have given more sophisticated FEM analyses for elliptic singularly perturbed reaction-diffusion problems, deriving error bounds (uniformly in the singular perturbation parameter) in *balanced norms* where each component (H^1 and L^2) of the error has the same order of magnitude for typical solutions. (In Sect. 6 we shall define a balanced norm $\|\cdot\|_{Bal}$ that fits the problem that we study here.)

Despite these successes with the elliptic singularly perturbed linear reaction-diffusion problem, FEM energy-norm analysis of the corresponding time-dependent parabolic problem has lagged behind. To get a sense of the difficulty that arises in the parabolic problem, consider the error analysis of a semidiscretisation of the classical heat equation $u_t - \Delta u = f$ when it is discretised only in space using a standard FEM. In [23, Theorem 1.2] the $L^\infty(L^2)$ error of this method is analysed, and one sees easily that the same argument will work in the singularly perturbed case $u_t - \varepsilon^2 \Delta u = f$, where $0 < \varepsilon \ll 1$ (of course then one has to choose a suitable spatial mesh, such as a Shishkin mesh, to obtain a satisfactory result). A related argument in [23, Theorem 1.3] bounds the $L^\infty(H^1)$ error for the heat equation semidiscretisation — but if one attempts to apply this argument to the singularly perturbed problem to obtain a bound in H_e^1 , the final error estimate is unsatisfactory because it contains a multiplicative factor ε^{-1} .

The only papers we know of that give H_e^1 error estimates for a FEM applied in space to a singularly perturbed parabolic problem are [5, 10], who consider a convection-diffusion problem. One can modify their analyses by setting the convection term equal to zero to address the reaction-diffusion problem $u_t - \varepsilon^2 \Delta u + bu = f$, but the bound obtained is only in $L^2(H_e^1)$ instead of the stronger $L^\infty(H_e^1)$ norm. Moreover, as we pointed out earlier, the ε -weighting in the H_e^1 norm is unbalanced (i.e., too strong) in the sense of Sect. 6. We do not know of any $L^\infty(H_e^1)$ norm error analysis for a FEM in space for this singularly perturbed problem, nor are we aware of a balanced $L^p(H^1)$ norm error bound for any $p > 1$. (Indeed, the only balanced-norm result appears to be the balanced $L^1(H^1)$ error bound in the preprint [1], which appeared after our paper was submitted for publication, and which unlike our paper uses a discontinuous Galerkin time discretisation.)

The current paper will fill both these gaps in the literature. It presents $L^\infty(H_e^1)$ and $L^\infty(\|\cdot\|_{Bal})$ error bounds when a spatial FEM is used to solve a parabolic singularly perturbed reaction-diffusion problem. The derivation of these error bounds requires, as one would expect, the introduction of some novel techniques.

We shall consider a singularly perturbed parabolic PDE where the spatial domain Ω is the unit square in \mathbb{R}^2 . The corresponding steady-state problem has been extensively studied; see [14, 20, 22]. Any typical solution of this class of parabolic problems exhibits boundary layers on all sides of Ω at all positive times. To solve the problem numerically, we use a uniform

mesh in time with backward Euler differencing, and in space a piecewise bilinear FEM on a Shishkin mesh (as is usually done in the steady-state case).

The numerical method is not particularly original, but our error analysis of it is very new — it differs substantially from previous error analyses of singularly perturbed parabolic reaction-diffusion problems. For example, Lemmas 3.1 and 3.2 for the backward Euler scheme are inspired by work on fractional-derivative parabolic problems; these inequalities have the advantage of simplicity but their usefulness in a singularly perturbed context has not previously been recognised. In contrast, the analysis in [1] depends on much deeper results from [6] for the discontinuous Galerkin time discretisation.

Our analysis leads to a $L^\infty(H^1_e)$ error bound in Theorem 5.4, and a $L^\infty(\|\cdot\|_{Bal})$ error bound in Theorem 6.4, both of which are novel — no analogous results have previously appeared in the literature for any numerical method that uses a FEM in space to solve this parabolic problem — and are of optimal order, as our numerical experiments will show.

The paper is structured as follows. In Sect. 2 we describe the singularly perturbed initial-boundary value problem that we study and the boundary layer behaviour of typical solutions. Some properties of the backward Euler scheme are derived in Sect. 3. The full numerical method (backward Euler in time on a uniform temporal mesh; piecewise bilinear FEM in space on a spatial Shishkin mesh) for solving our initial–boundary value problem is defined in Sect. 4. The energy-norm and balanced-norm error analyses for this method are carried out in Sects. 5 and 6 respectively. Finally, numerical experiments in Sect. 7 confirm the sharpness of our theoretical results.

Notation. We use $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ for the norm and inner product in $L^2(\Omega)$, while $\|\cdot\|_k$ and $|\cdot|_k$ denote the Sobolev norm and seminorm on $H^k(\Omega)$ for $k = 1, 2$. The generic constant C is independent of the singular perturbation parameter ε and of the mesh, so that *throughout our analysis any dependence on ε is stated explicitly.*

2 Statement of the Problem

Consider the parabolic singularly perturbed problem

$$\frac{\partial u}{\partial t}(x, t) - \varepsilon^2 \Delta u(x, t) + b(x)u(x, t) = f(x, t) \text{ for } (x, t) \in Q := \Omega \times (0, T], \quad (2.1)$$

where $\Omega = (0, 1)^2$ and T is a positive constant, with initial condition $u(x, 0) = u_0(x)$ for $x = (x_1, x_2) \in \Omega$, and boundary conditions $u(x, t) = 0$ for $(x, t) \in \partial\Omega \times (0, T]$. We assume that $u_0 \in C(\bar{\Omega})$ where $\bar{\Omega} := [0, 1]^2$, with $u_0(x) = 0$ for $x \in \partial\Omega$. We also assume that f and b are smooth functions (more precise hypotheses will be given later), and without loss of generality we take $b > \beta^2$ on $\bar{\Omega}$, where $\beta > 0$ is a constant — this can always be achieved by a change of variable of the form $u(x, t) = e^{kt}v(x, t)$ for some suitable constant k .

Error bounds for our numerical method will be derived in two distinct norms: the energy norm of Sect. 5 and the balanced norm of Sect. 6. In these analyses, Sect. 6 requires more regularity of the solution u than Sect. 5.

Set $\bar{Q} := \bar{\Omega} \times [0, T]$. We use the Hölder spaces $C^{\beta, \beta/2}(\bar{Q})$, with $\beta > 0$, that are standard in the analysis of parabolic PDEs. Let $\sigma \in (0, 1)$ be arbitrary but fixed. From [12, pp. 319, 320] (see also [2, Section 2] and [21, Section 5.2]), sufficient conditions for $u \in C^{k+\sigma, (k+\sigma)/2}(\bar{Q})$ with $k = 5$ (needed in Sect. 5) and $k = 6$ (needed in Sect. 6) are that

$f \in C^{k-2+\sigma, (k-2+\sigma)/2}(\bar{Q})$, $b \in C^{k-2+\sigma}(\bar{\Omega})$, and $u_0 \in C^{k+\sigma}(\bar{\Omega})$, and that the following compatibility conditions are satisfied: setting $\mathcal{L}_\varepsilon w := -\varepsilon^2 \Delta w + bw$, for all $x \in \partial\Omega$ one has

$$u_0(x) = 0, \tag{2.2a}$$

$$-\mathcal{L}_\varepsilon u_0(x) + f(x, 0) = 0, \tag{2.2b}$$

$$(\mathcal{L}_\varepsilon)^2 u_0(x) - \mathcal{L}_\varepsilon f(x, 0) + \frac{\partial f}{\partial t}(x, 0) = 0, \tag{2.2c}$$

$$-(\mathcal{L}_\varepsilon)^3 u_0(x) + (\mathcal{L}_\varepsilon)^2 f(x, 0) - \mathcal{L}_\varepsilon \frac{\partial f}{\partial t}(x, 0) + \frac{\partial^2 f}{\partial t^2}(x, 0) = 0, \tag{2.2d}$$

where (2.2a)–(2.2c) are required when $k = 5$ and (2.2a)–(2.2d) are required when $k = 6$.

Then from [2,Section 2] and [21,Section 5.2], the solution u can be decomposed as $u = U + \sum_{i=1}^4 v_i + \sum_{i=1}^4 w_i$, where U is a smooth component, the v_i ($i = 1, 2, 3, 4$) are edge boundary layer functions associated with the four sides of the unit square and the w_i ($i = 1, 2, 3, 4$) are corner layer terms. (This terminology is standard in this research area, although the corner layers are located not at the corners of \bar{Q} but along the 4 line segments (x_1, x_2, t) with (x_1, x_2) a corner of $\bar{\Omega}$ and $0 < t \leq T$; a similar statement can be made for the edge layers.) Furthermore, these components satisfy the following bounds for all $(x, t) \in Q$ and $k_1 + k_2 + 2k_t \leq k$: there exists a constant $C > 0$ such that

$$\left| \frac{\partial^{k_1+k_2+k_t} U(x, t)}{\partial x_1^{k_1} \partial x_2^{k_2} \partial t^{k_t}} \right| \leq C, \tag{2.3a}$$

$$\left| \frac{\partial^{k_1+k_2+k_t} v_1(x, t)}{\partial x_1^{k_1} \partial x_2^{k_2} \partial t^{k_t}} \right| \leq C\varepsilon^{-k_1} e^{-\beta x_1/\varepsilon}, \tag{2.3b}$$

$$\left| \frac{\partial^{k_1+k_2+k_t} w_1(x, t)}{\partial x_1^{k_1} \partial x_2^{k_2} \partial t^{k_t}} \right| \leq C\varepsilon^{-k_1-k_2} \min \{ e^{-\beta x_1/\varepsilon}, e^{-\beta x_2/\varepsilon} \}, \tag{2.3c}$$

with analogous bounds for the other layer terms.

3 Stability of the Backward Euler Scheme

Throughout the paper, we use the uniform temporal mesh $\{t_m := m\tau\}_{m=0}^M$ where M is a positive integer and $\tau = T/M$. Let δ_t denote the standard backward Euler operator defined by $\delta_t V^m = (V^m - V^{m-1})/\tau$ for each mesh function $\{V^m\}_{m=0}^M$.

The following lemma is related to [11,Theorem 2.1], which is a stability result for a discretisation of a Caputo fractional derivative.

- Lemma 3.1** (i) Suppose that the mesh function $\{V^m\}_{m=0}^M$ satisfies $V^0 = 0$ and $|\delta_t V^m| \leq K$ for $m = 1, 2, \dots, M$, where $K \geq 0$ is some quantity that is independent of m . Then $|V^m| \leq Km\tau$ for $m = 0, 1, \dots, M$.
- (ii) The conclusion of part (i) still holds if the hypothesis $|\delta_t V^m| \leq K$ is replaced by $\delta_t |V^m| \leq K$.

Proof Part (i): For $m = 1, \dots, M$, from $|\delta_t V^m| \leq K$ we get $|V^m| \leq |V^{m-1}| + K\tau$. An easy induction argument using $V^0 = 0$ then gives $|V^m| \leq Km\tau$, as desired.

Part (ii): Like [11, Theorem 2.1], define the mesh function $\{W^m\}_{m=0}^M$ by $W^0 = 0$ and $\delta_t W^m = \max\{0, \delta_t |V^m|\}$ for $m = 1, 2, \dots, M$. Then $0 \leq |V^m| \leq W^m$ for all m since δ_t is associated with an M-matrix, and the result follows from applying Part (i) to W^m . \square

The backward Euler scheme also enjoys the following properties (a related inequality for the L1 discretisation of the Caputo fractional derivative is proved in [9, Lemma 4.3]).

Lemma 3.2 *Let $v^m \in L^2(\Omega)$ for $m = 0, 1, \dots, M$. Then*

$$\langle \delta_t v^m, v^m \rangle \geq (\delta_t \|v^m\|) \|v^m\| \quad \text{and} \quad \langle \delta_t v^m, v^m \rangle \geq \frac{1}{2} \delta_t (\|v^m\|^2)$$

for each m .

Proof The definition of $\delta_t v^m$ and a Cauchy-Schwarz inequality give

$$\tau \langle \delta_t v^m, v^m \rangle = \langle v^m - v^{m-1}, v^m \rangle \geq \|v^m\|^2 - \|v^{m-1}\| \|v^m\| = \tau (\delta_t \|v^m\|) \|v^m\|,$$

which proves the first inequality. For the second inequality, we have similarly

$$\begin{aligned} 2\tau \langle \delta_t v^m, v^m \rangle &\geq 2 \|v^m\|^2 - 2 \|v^{m-1}\| \|v^m\| \\ &\geq 2 \|v^m\|^2 - \|v^{m-1}\|^2 - \|v^m\|^2 = \tau \delta_t (\|v^m\|^2). \end{aligned}$$

\square

In Sect. 5 the first inequality of Lemma 3.2 will be used to bound the $L^2(\Omega)$ norm of the error, while the second inequality will bound its $H^1(\Omega)$ seminorm.

4 The Numerical Method

To discretise (2.1) we use the backward Euler scheme in time, and in space a bilinear FEM on a Shishkin mesh (to deal with the boundary layers in the solution). We now define the Shishkin mesh and the FEM space.

Let N be an even positive integer. Let λ be a mesh transition parameter that specifies where the piecewise-uniform mesh changes from coarse to fine: it is defined by $\lambda = \min\{1/4, 2\varepsilon\beta^{-1} \ln N\}$. Without loss of generality one can assume that ε is so small that $\lambda = 2\varepsilon\beta^{-1} \ln N$. Divide each of the intervals $[0, \lambda]$ and $[1 - \lambda, 1]$ into $N/4$ equidistant subintervals and divide $[\lambda, 1 - \lambda]$ into $N/2$ equidistant subintervals. This gives a 1D Shishkin mesh that is coarse on $[\lambda, 1 - \lambda]$ and fine elsewhere in $[0, 1]$. Then take a tensor product of two 1D Shishkin meshes to construct the 2D Shishkin mesh on Ω ; Fig. 1 displays an example of this mesh for the case $N = 8$. (See [22] for further discussion of Shishkin meshes.) Finally, let $V_{h0} \subset H_0^1(\Omega)$ be the piecewise bilinear finite element space defined on the Shishkin mesh Ω_h .

For any suitable function g , set $g^m(x) = g(x, t_m)$, $\partial g^m(x)/\partial t = [\partial g(x, t)/\partial t]_{t=t_m}$ and $\delta_t g^m(x) = [g(x, t_m) - g(x, t_{m-1})]/\tau$.

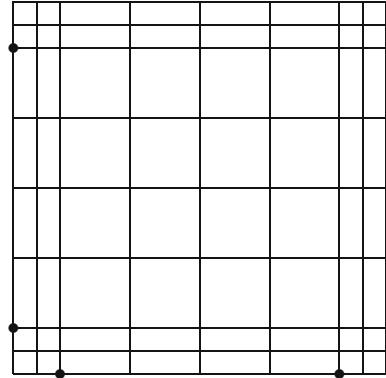
Define the $L^2(\Omega)$ projector $P_h : L^2(\Omega) \rightarrow V_{h0}$ by $\langle P_h w, v_h \rangle = \langle w, v_h \rangle$ for all $v_h \in V_{h0}$. Clearly $\|P_h w\| \leq \|w\|$ for all $w \in L^2(\Omega)$.

Define the Ritz projector $\mathcal{R}_h : H_0^1(\Omega) \rightarrow V_{h0}$ by

$$\varepsilon^2 \langle \nabla \mathcal{R}_h v, \nabla w_h \rangle + \langle b \mathcal{R}_h v, w_h \rangle = \varepsilon^2 \langle \nabla v, \nabla w_h \rangle + \langle bv, w_h \rangle$$

for all $w_h \in V_{h0}$.

Fig. 1 Shishkin mesh for reaction-diffusion



Define the discrete Laplacian $\Delta_h : V_{h0} \rightarrow V_{h0}$ by $\langle \Delta_h v_h, w_h \rangle = -\langle \nabla v_h, \nabla w_h \rangle$ for all $v_h, w_h \in V_{h0}$.

Our numerical method for solving (2.1) is: for $m = 1, \dots, M$, find $u_h^m := u_h^m(\cdot, t_m) \in V_{h0}$ satisfying

$$\langle \delta_t u_h^m, v_h \rangle + \varepsilon^2 \langle \nabla u_h^m, \nabla v_h \rangle + \langle bu_h^m, v_h \rangle = \langle f^m, v_h \rangle \quad \forall v_h \in V_{h0}, \tag{4.1}$$

with $u_h^0 := \mathcal{R}_h u^0$. One can write (4.1) as

$$\langle \delta_t u_h^m, v_h \rangle - \varepsilon^2 \langle \Delta_h u_h^m, v_h \rangle + \langle P_h (bu_h^m), v_h \rangle = \langle P_h f^m, v_h \rangle \quad \forall v_h \in V_{h0},$$

which is equivalent to

$$\delta_t u_h^m - \varepsilon^2 \Delta_h u_h^m + P_h (bu_h^m) = P_h f^m, \tag{4.2}$$

since each of these terms lies in V_{h0} .

5 Energy Norm Error Analysis

We begin our error analysis with some preliminary estimates involving \mathcal{R}_h .

For $i = 0, 1, 2$, set

$$\rho_i(x, t) := \frac{\partial^i [\mathcal{R}_h u(x, t) - u(x, t)]}{\partial t^i} = \left[\mathcal{R}_h \left(\frac{\partial^i u}{\partial t^i} \right) \right](x, t) - \frac{\partial^i u(x, t)}{\partial t^i} \tag{5.1}$$

since \mathcal{R}_h acts only in the spatial variables. Also, set $\rho_i^m(x) = \rho_i(x, t_m)$.

Lemma 5.1 *Assume that the derivative bounds (2.3) hold true for $k = 4$. Then there exist constants C such that for $i = 0, 1$ and all $t \in (0, T)$, one has*

$$\varepsilon |\rho_i(\cdot, t)|_1 + \|\rho_i(\cdot, t)\| \leq C (\varepsilon^{1/2} N^{-1} \ln N + N^{-2}) \tag{5.2}$$

and

$$\varepsilon^{1/2} |\rho_i(\cdot, t)|_1 + \|\rho_i(\cdot, t)\| \leq CN^{-1} \ln N. \tag{5.3}$$

If the derivative bounds (2.3) hold true for $k = 6$, then (5.2) and (5.3) are also true when $i = 2$.

Proof From its definition, $\mathcal{R}_h u(\cdot, t)$ is the Galerkin solution of the steady-state problem got by deleting $\partial u / \partial t$ from (2.1) and taking $f = f(\cdot, t)$. Hence in the case $i = 0$, one gets (5.2) from an inspection of the proof of [16, Theorem 3.1], while (5.3) follows from the proof of [7, Theorem 2.6]; note that both of these arguments use the bounds (2.3) only for $k_1 + k_2 \leq 2$ and $k_t = 0$.

The case $i = 1$ is proved in a similar way, replacing u by $\partial u / \partial t$ and using (5.1) and (2.3) for $k_1 + k_2 \leq 2$ and $k_t = 1$; and if the derivative bounds (2.3) hold true for $k = 6$, then the same argument applied to $\partial^2 u / \partial t^2$ with $k_1 + k_2 \leq 2$ and $k_t = 2$ proves the case $i = 2$. \square

For $m = 1, 2, \dots, M$, set $r_1^m := (\delta_t - \partial / \partial t) \mathcal{R}_h u^m$.

Lemma 5.2 *Assume that the derivative bounds (2.3) hold true for $k = 5$. Then there exists a constant C such that $\|r_1^m\| + \varepsilon |r_1^m|_1 \leq CM^{-1}$ for $m = 1, \dots, M$.*

Proof The definition of \mathcal{R}_h implies that for $t \in (0, T]$ one has

$$\begin{aligned} \|(\partial / \partial t)^2 \mathcal{R}_h u(\cdot, t)\| + \varepsilon |(\partial / \partial t)^2 \mathcal{R}_h u(\cdot, t)|_1 &= \|\mathcal{R}_h(\partial / \partial t)^2 u(\cdot, t)\| + \varepsilon |\mathcal{R}_h(\partial / \partial t)^2 u(\cdot, t)|_1 \\ &\leq \|(\partial / \partial t)^2 u(\cdot, t)\| + \varepsilon |(\partial / \partial t)^2 u(\cdot, t)|_1 \\ &\leq C, \end{aligned}$$

where we used $k = 5$ in (2.3) to derive the final inequality. The result follows. \square

Set $e_h^m := \mathcal{R}_h u^m - u_h^m$ and

$$r^m := \rho_1^m + r_1^m = \frac{\partial [\mathcal{R}_h u^m - u^m]}{\partial t} + (\delta_t - \partial / \partial t) \mathcal{R}_h u^m = \delta_t \mathcal{R}_h u^m - \frac{\partial u^m}{\partial t}.$$

In the next lemma, we derive some preliminary bounds on $e_h^m(x)$.

Lemma 5.3 *There exist constants C such that for $m = 1, \dots, M$ one has*

$$\max_{m=1, \dots, M} \|e_h^m\| \leq C \max_{m=1, \dots, M} \|r^m\| \tag{5.4}$$

and

$$\varepsilon \max_{m=1, \dots, M} |e_h^m|_1 \leq C \max_{m=1, \dots, M} \|r^m\|. \tag{5.5}$$

Proof The definition of e_h^m and (4.2) give

$$\delta_t e_h^m - \varepsilon^2 \Delta_h e_h^m + P_h (b e_h^m) = \delta_t \mathcal{R}_h u^m - \varepsilon^2 \Delta_h \mathcal{R}_h u^m + P_h (b \mathcal{R}_h u^m) - P_h f^m. \tag{5.6}$$

Take $v = u^m$ in the definition of \mathcal{R}_h , then recall the definitions of Δ_h and P_h to get

$$\begin{aligned} \langle -\varepsilon^2 \Delta_h \mathcal{R}_h u^m + P_h (b \mathcal{R}_h u^m), w_h \rangle &= \langle -\varepsilon^2 \Delta u^m + b u^m, w_h \rangle \\ &= \left\langle -\frac{\partial u^m}{\partial t} + f^m, w_h \right\rangle \end{aligned}$$

using (2.1), for all $w_h \in V_{h0}$. Thus $-\varepsilon^2 \Delta_h \mathcal{R}_h u^m + P_h (b \mathcal{R}_h u^m) = P_h \left(-\frac{\partial u^m}{\partial t} + f^m \right)$. Hence (5.6) simplifies to

$$\delta_t e_h^m - \varepsilon^2 \Delta_h e_h^m + P_h (b e_h^m) = \delta_t \mathcal{R}_h u^m - P_h \left(\frac{\partial u^m}{\partial t} \right) = P_h r^m. \tag{5.7}$$

Invoking the first inequality of Lemma 3.2, we have

$$\begin{aligned}
 (\delta_t \|e_h^m\|) \|e_h^m\| &\leq \langle \delta_t e_h^m, e_h^m \rangle \\
 &\leq \langle \delta_t e_h^m, e_h^m \rangle + \varepsilon^2 \langle \nabla e_h^m, \nabla e_h^m \rangle + \langle b e_h^m, e_h^m \rangle \\
 &= \langle \delta_t e_h^m, e_h^m \rangle - \varepsilon^2 \langle \Delta_h e_h^m, e_h^m \rangle + \langle P_h (b e_h^m), e_h^m \rangle \\
 &= \langle P_h r^m, e_h^m \rangle \\
 &\leq \|P_h r^m\| \|e_h^m\|,
 \end{aligned}$$

where we used (5.7). Thus, either $\|e_h^m\| = 0$ or $\delta_t \|e_h^m\| \leq \|P_h r^m\| \leq \max_{m=1, \dots, M} \|r^m\|$. Now Lemma 3.1 yields (5.4), since one can start its inductive proof at each value of m for which $\|e_h^m\| = 0$ (note that $\|e_h^0\| = 0$).

Invoking the second inequality of Lemma 3.2, we obtain

$$\begin{aligned}
 \|\delta_t e_h^m\|^2 + \frac{1}{2} \varepsilon^2 \delta_t (\|\nabla e_h^m\|^2) &\leq \langle \delta_t e_h^m, \delta_t e_h^m \rangle + \varepsilon^2 \langle \nabla e_h^m, \delta_t \nabla e_h^m \rangle \\
 &= \langle \delta_t e_h^m, \delta_t e_h^m \rangle - \varepsilon^2 \langle \Delta_h e_h^m, \delta_t e_h^m \rangle \\
 &= \langle P_h r^m, \delta_t e_h^m \rangle - \langle P_h (b e_h^m), \delta_t e_h^m \rangle \\
 &\leq \frac{1}{2} (\|P_h r^m\|^2 + \|\delta_t e_h^m\|^2 + \|P_h (b e_h^m)\|^2 + \|\delta_t e_h^m\|^2),
 \end{aligned}$$

where we again used (5.7). Hence

$$\begin{aligned}
 \varepsilon^2 \delta_t (\|\nabla e_h^m\|^2) &\leq \|P_h r^m\|^2 + \|P_h (b e_h^m)\|^2 \\
 &\leq C \left(\max_{m=1, \dots, M} \|r^m\| \right)^2,
 \end{aligned}$$

by (5.4). Then Lemma 3.1 gives us

$$(\varepsilon \|\nabla e_h^m\|)^2 \leq C m \tau \left(\max_{m=1, \dots, M} \|r^m\| \right)^2 \leq C \left(\max_{m=1, \dots, M} \|r^m\| \right)^2,$$

which is (5.5). □

For any function $w \in H^1(\Omega)$, define its energy norm $\|w\|_{1,e}$ by

$$\|w\|_{1,e} = \{\varepsilon^2 |w|_1^2 + \|w\|^2\}^{1/2}$$

This norm was referred to as the H_e^1 norm in Sect. 1.

We now derive a $L^\infty(H_e^1)$ error bound for our method.

Theorem 5.4 (Energy norm error bound) *Assume that the derivative bounds (2.3) hold true for $k = 5$. Then there exists a constant C such that the solution u of (2.1) and the solution u_h^m of (4.1) satisfy*

$$\max_{m=1, \dots, M} \|u^m - u_h^m\|_{1,e} \leq C (\varepsilon^{1/2} N^{-1} \ln N + N^{-2} + M^{-1}).$$

Proof By Lemma 5.3, Lemma 5.2, and (5.2), one has the energy norm error estimate

$$\begin{aligned}
 \max_{m=1, \dots, M} \|e_h^m\|_{1,e} &\leq C \max_{m=1, \dots, M} \|r^m\| \\
 &\leq C \left(\max_{m=1, \dots, M} \|\rho_1^m\| + \max_{m=1, \dots, M} \|r_1^m\| \right) \\
 &\leq C (\varepsilon^{1/2} N^{-1} \ln N + N^{-2} + M^{-1}).
 \end{aligned}$$

But $u - u_h^m = e_h^m - \rho_0^m$, so we can combine the above estimate and the bound (5.2) of Lemma 5.1 for ρ_0^m to finish the proof. \square

6 Balanced Norm Error Analysis

In this section we use the derivative bounds (2.3) with $k = 6$.

As we pointed out in Sect. 1, the energy norm of Sect. 5 is weaker than it looks — for solutions of typical singularly perturbed problems, it is dominated by its L^2 component. Thus we now derive an error bound in a *balanced norm* where the ε -dependent weighting of the H^1 component is such that the H^1 and L^2 components of the error have similar orders of magnitude. This result will be proved under the additional assumption that *the reaction term coefficient b is a positive constant*.

Remark 6.1 To extend our balanced-norm analysis to variable $b(x)$ seems not to be straightforward, essentially because the $L^2(\Omega)$ projector P_h is not $H^1(\Omega)$ -stable on a Shishkin mesh (this follows from [4,p.527]).

First, we sharpen the result of Lemma 5.2 under a stronger hypothesis on the derivative bounds (2.3).

Lemma 6.2 *Assume that the derivative bounds (2.3) hold true for $k = 6$. Then*

$$\|r_1^m\| + \varepsilon^{1/2} |r_1^m|_1 \leq CM^{-1} \text{ for } m = 1, \dots, M.$$

Proof From (2.3) with $k = 6$, one sees that $\|(\partial/\partial t)^2 u(\cdot, t)\| + \varepsilon^{1/2} |(\partial/\partial t)^2 u(\cdot, t)|_1 \leq C$ for all $t \in (0, T]$. This inequality, Lemma 5.1 with $i = 2$, and a triangle inequality yield the desired result. \square

Lemma 6.3 *There exists a constant C such that*

$$\max_{m=1, \dots, M} |e_h^m|_1 \leq C \left(\max_{m=1, \dots, M} |P_h \rho_1^m|_1 + \varepsilon^{-1/2} M^{-1} \right). \tag{6.1}$$

Proof From Lemma 3.2 and b constant it follows that

$$\begin{aligned} [\delta_t \|\nabla e_h^m\|] \|\nabla e_h^m\| &\leq \langle \delta_t \nabla e_h^m, \nabla e_h^m \rangle \\ &\leq \langle \delta_t e_h^m, -\Delta_h e_h^m \rangle - \varepsilon^2 \langle \Delta_h e_h^m, -\Delta_h e_h^m \rangle + \langle b \nabla e_h^m, \nabla e_h^m \rangle \\ &= \langle \delta_t e_h^m, -\Delta_h e_h^m \rangle - \varepsilon^2 \langle \Delta_h e_h^m, -\Delta_h e_h^m \rangle + \langle b e_h^m, -\Delta_h e_h^m \rangle \\ &= \langle P_h r^m, -\Delta_h e_h^m \rangle \\ &\leq \|\nabla P_h r^m\| \|\nabla e_h^m\|, \end{aligned}$$

where we used (5.7). If $\|\nabla e_h^m\| = 0$ we are done; thus we assume that $\|\nabla e_h^m\| \neq 0$ and deduce that

$$\delta_t |e_h^m|_1 \leq |P_h r^m|_1 \leq |P_h \rho_1^m|_1 + |P_h r_1^m|_1 \leq \max_{m=1, \dots, M} |P_h \rho_1^m|_1 + C\varepsilon^{-1/2} M^{-1},$$

by Lemma 6.2 and $P_h r_1^m = r_1^m$. Now an appeal to Lemma 3.1 gives (6.1). \square

From the discussion in [13] and the bounds (2.3), one sees that the norm

$$\|w\|_{Bal} := \{\varepsilon |w|_1^2 + \|w\|^2\}^{1/2} \quad \forall w \in H^1(\Omega)$$

defines a *balanced norm* for the class of problems that we are studying. The next result is an error bound for our numerical method in $L^\infty(\|\cdot\|_{Bal})$. Note that, unlike Theorem 5.4, the bound does not contain a factor that vanishes as $\varepsilon \rightarrow 0$ (for fixed N); this is precisely because of the balanced nature of the result.

Theorem 6.4 (Balanced norm error bound) *Assume that the derivative bounds (2.3) hold true for $k = 6$. Recall that u is the solution of (2.1) and u_h^m is the solution of (4.1). Then there exists a constant C such that*

$$\max_{m=1,\dots,M} \|u^m - u_h^m\|_{Bal} \leq C (N^{-1} (\ln N)^{3/2} + M^{-1}).$$

Proof From (5.1) we have

$$\begin{aligned} |P_h \rho_1^m(\cdot)|_1 &= |\mathcal{R}_h [\partial u^m(\cdot)/\partial t] - P_h [\partial u^m(\cdot)/\partial t]|_1 \\ &\leq |\mathcal{R}_h [\partial u^m(\cdot)/\partial t] - [\partial u^m(\cdot)/\partial t]|_1 + |[\partial u^m(\cdot)/\partial t] - P_h [\partial u^m(\cdot)/\partial t]|_1 \\ &\leq |\rho_1^m|_1 + C\varepsilon^{-1/2} N^{-1} (\ln N)^{3/2}, \end{aligned} \tag{6.2}$$

by [19,(13)] applied to the function $\partial u/\partial t$, since the spatial derivative bounds for $\partial u/\partial t$ in (2.3) are the same as for the elliptic problem studied in [19].

Observe that $\|w\|_{Bal}$ is equivalent to $\varepsilon^{1/2} \|w\|_1 + \|w\|$ for all $w \in H^1(\Omega)$. Then Lemma 6.3 and inequality (5.4) in Lemma 5.3 yield

$$\begin{aligned} \max_{m=1,\dots,M} \|e_h^m\|_{Bal} &\leq C\varepsilon^{1/2} \left(\max_{m=1,\dots,M} |P_h \rho_1^m|_1 + \varepsilon^{-1/2} M^{-1} \right) + \max_{m=1,\dots,M} \|r^m\| \\ &\leq C \left(\max_{m=1,\dots,M} \varepsilon^{1/2} |P_h \rho_1^m|_1 + M^{-1} \right) \\ &\quad + \max_{m=1,\dots,M} \|\rho_1^m\| + \max_{m=1,\dots,M} \|r_1^m\|, \end{aligned}$$

since $r^m = \rho_1^m + r_1^m$. Hence

$$\max_{m=1,\dots,M} \|e_h^m\|_{Bal} \leq C (N^{-1} (\ln N)^{3/2} + M^{-1})$$

by (6.2), inequality (5.3) of Lemma 5.1, and Lemma 5.2. As $u^m - u_h^m = e_h^m - \rho_0^m$, the desired result now follows from Lemma 5.1 and a triangle inequality. \square

7 Numerical Results

Our numerical experiments will use the same test problem as [15,Example 1]. That is, we take $b = 1$ and $T = 1$, and choose the exact solution

$$u(x, y, t) = (1 - e^{-t}) \left(\cos \frac{\pi x}{2} - \frac{e^{-x/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}} \right) \left(1 - y - \frac{e^{-y/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}} \right).$$

Then f is chosen so that (2.1) is satisfied, and we take $u_0(x, y) = u(x, y, 0) \equiv 0$. The derivatives of u have exactly the form of the bounds (2.3), so it is a valid solution on which to test our theory. Unlike in (2.1), the function f depends on ε , but in a harmless way — one can verify easily that our error analysis is unaffected by this deviation from the form of (2.1).

Numerical errors will be measured in the energy norm and balanced norm that were used in Theorems 5.4 and 6.4 respectively to bound the error in the computed solution.

Table 1 Energy-norm errors and convergence rates

	N=16,M=64	N=32,M=182	N=64,M=512	N=128,M=1449
$\varepsilon = 10^{-2}$	2.0165e-2 0.7657	1.1860e-2 0.8179	6.7278e-3 0.8386	3.7619e-3
$\varepsilon = 10^{-4}$	2.8471e-3 0.9603	1.4632e-3 0.9201	7.7322e-4 0.8999	4.1436e-4
$\varepsilon = 10^{-6}$	1.3766e-3 1.4000	5.2163e-4 1.3184	2.0915e-4 1.2524	8.7790e-5
$\varepsilon = 10^{-8}$	1.2389e-3 1.5203	4.3190e-4 1.4837	1.5442e-4 1.4714	5.5691e-5

Table 2 Balanced-norm errors and convergence rates

	N=16,M=64	N=32,M=182	N=64,M=512	N=128,M=1449
$\varepsilon = 10^{-2}$	1.5443e-1 0.6341	9.9505e-2 0.7228	6.0290e-2 0.7754	3.5223e-2
$\varepsilon = 10^{-4}$	1.5308e-1 0.6204	9.9575e-2 0.7130	6.0744e-2 0.7691	3.5642e-2
$\varepsilon = 10^{-6}$	1.5300e-1 0.6201	9.9543e-2 0.7127	6.0736e-2 0.7689	3.5642e-2
$\varepsilon = 10^{-8}$	1.5300e-1 0.6201	9.9543e-2 0.7127	6.0736e-2 0.7689	3.5642e-2

We concentrate first on the spatial error, since this is where the effect of the boundary layers is felt. In Tables 1 (energy-norm errors) and 2 (balanced-norm errors) we take $(N, M) = (16, 64), (32, 182), (64, 512), (128, 1449)$, i.e., $M \approx N^{3/2}$ in each case, so that the spatial component of the error should dominate the total error. To see the rates of convergence more easily, we graph these energy-norm and balanced-norm errors in Figs. 2 and 3 respectively, where we take $\varepsilon = 10^{-2}$ so that each figure encompasses the regimes $N \ll \varepsilon^{-1}, N \approx \varepsilon^{-1}, N \gg \varepsilon^{-1}$.

Theorem 5.4 predicts an energy-norm error of $O(\varepsilon^{1/2}N^{-1} \ln N + N^{-3/2})$; Table 1 and Fig. 2 agree with this. In particular, in Fig. 2 where $\varepsilon = 10^{-2}$, the convergence of the energy-norm error is $O(N^{-1} \ln N)$, while when $\varepsilon = 10^{-8} \ll N^{-1}$ in Table 1, then the energy-norm error is $O(N^{-3/2})$.

Theorem 6.4 predicts balanced-norm errors that are $O(N^{-1}(\ln N)^{3/2})$, and this agrees with the numerical results in Table 2 and Fig. 3 (these errors may be $O(N^{-1} \ln N)$, which is slightly better). Note that for each fixed N in Table 2 the balanced-norm errors are essentially independent of ε when ε is small, as our theory predicts.

Of course the errors in the balanced norm are larger than those in the energy norm; compare Tables 1 and 2.

Next, we consider the temporal error by taking $(N, M) = (16, 8), (32, 14), (64, 23), (128, 39)$, i.e., $M \approx N^{3/4}$, so that the temporal error dominates the total error. Now Theorems 5.4 and 6.4 predict both the energy-norm error and balanced-norm error to be $O(N^{-3/4})$. Tables 3 and 4 evidently agree with this prediction.

Fig. 2 log-log graph of the energy-norm error when $\varepsilon = 10^{-2}$

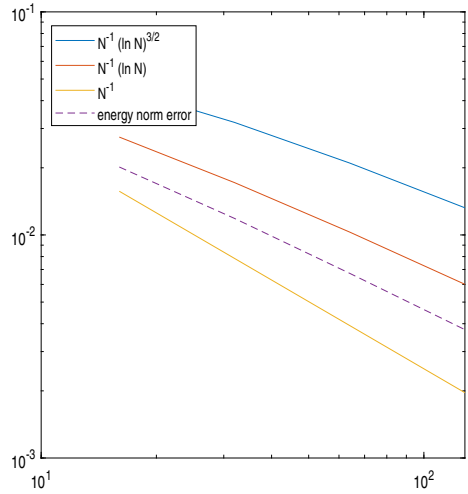


Fig. 3 log-log graph of the balanced-norm error when $\varepsilon = 10^{-2}$

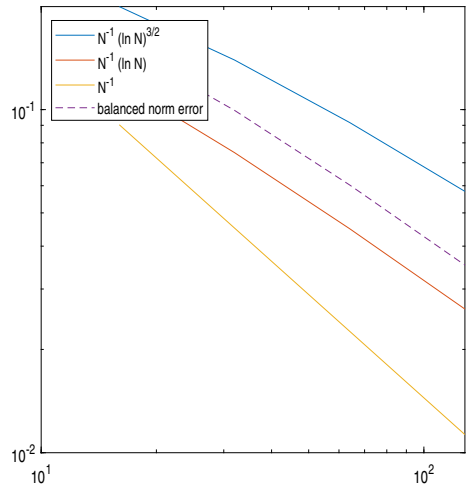


Table 3 Energy-norm errors and convergence rates

	N=16,M=8	N=32,M=14	N=64,M=23	N=128,M=39
$\varepsilon = 10^{-2}$	2.4863e-2	1.5139e-2	9.1139e-3	5.3379e-3
	0.7157	0.7321	0.7717	
$\varepsilon = 10^{-4}$	1.0458e-2	6.2024e-3	3.8125e-3	2.2607e-3
	0.7537	0.7020	0.7539	
$\varepsilon = 10^{-6}$	9.0854e-3	5.3094e-3	3.2677e-3	1.9407e-3
	0.7749	0.7002	0.7516	
$\varepsilon = 10^{-8}$	8.9491e-3	5.2203e-3	3.2131e-3	1.9087e-3
	0.7776	0.7001	0.7513	

Table 4 Balanced-norm errors and convergence rates

	N=16,M=8	N=32,M=14	N=64,M=23	N=128,M=39
$\varepsilon = 10^{-2}$	1.5867e-1 0.6280	1.0267e-1 0.7118	6.2686e-2 0.7670	3.6835e-2
$\varepsilon = 10^{-4}$	1.6017e-1 0.6204	1.0418e-1 0.7076	6.3796e-2 0.7653	3.7530e-2
$\varepsilon = 10^{-6}$	1.6019e-1 0.6203	1.0420e-1 0.7076	6.3807e-2 0.7653	3.7537e-2
$\varepsilon = 10^{-8}$	1.6019e-1 0.6203	1.0420e-1 0.7076	6.3807e-2 0.7653	3.7537e-2

Acknowledgements We are grateful to two unknown reviewers who provided several perceptive and helpful comments that guided us in improving the clarity of the paper.

Funding The research of Xiangyun Meng is supported in part by the National Natural Science Foundation of China under grant 12101039 and by the Fundamental Research Funds for the Central Universities under grant 2020RC101. The research of Martin Stynes is supported in part by the National Natural Science Foundation of China under grants 12171025 and NSAF-U1930402.

Availability of data and materials Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Brdar, M., Franz, S., Ludwig, L., Roos, H.-G.: A time dependent singularly perturbed problem with shift in space, (2022). [arXiv:2202.01601](https://arxiv.org/abs/2202.01601)
2. Bujanda, B., Clavero, C., Gracia, J.L., Jorge, J.C.: A high order uniformly convergent alternating direction scheme for time dependent reaction-diffusion singularly perturbed problems. *Numer. Math.* **107**(1), 1–25 (2007)
3. Cai, Z., Ku, J.: A dual finite element method for a singularly perturbed reaction-diffusion problem. *SIAM J. Numer. Anal.* **58**(3), 1654–1673 (2020)
4. Crouzeix, M., Thomée, V.: The stability in L_p and W_p^1 of the L_2 -projection onto finite element function spaces. *Math. Comp.* **48**(178), 521–532 (1987)
5. Dolejší, V., Roos, H.: BDF-FEM for parabolic singularly perturbed problems with exponential layers on layers-adapted meshes in space. *Neural Parallel Sci. Comput.* **18**(2), 221–235 (2010)
6. Franz, S., Matthies, G.: A unified framework for time-dependent singularly perturbed problems with discontinuous Galerkin methods in time. *Math. Comp.* **87**(313), 2113–2132 (2018)
7. Franz, S., Roos, H.-G.: Error estimates in balanced norms of finite element methods for higher order reaction-diffusion problems. *Int. J. Numer. Anal. Model.* **17**(4), 532–542 (2020)
8. Heuer, N., Karkulik, M.: A robust DPG method for singularly perturbed reaction-diffusion problems. *SIAM J. Numer. Anal.* **55**(3), 1218–1242 (2017)
9. Huang, C., Stynes, M.: A direct discontinuous Galerkin method for a time-fractional diffusion equation with a Robin boundary condition. *Appl. Numer. Math.* **135**, 15–29 (2019)
10. Kaland, L., Roos, H.-G.: Parabolic singularly perturbed problems with exponential layers: robust discretizations using finite elements in space on Shishkin meshes. *Int. J. Numer. Anal. Model.* **7**(3), 593–606 (2010)

11. Kopteva, N., Meng, X.: Error analysis for a fractional-derivative parabolic problem on quasi-graded meshes using barrier functions. *SIAM J. Numer. Anal.* **58**(2), 1217–1238 (2020)
12. Ladyzenskaja, O. A., Solonnikov, V. A., Ural'tseva, N. N.: *Linear and quasilinear equations of parabolic type*. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968. Translated from the Russian by S. Smith
13. Lin, R., Stynes, M.: A balanced finite element method for singularly perturbed reaction-diffusion problems. *SIAM J. Numer. Anal.* **50**(5), 2729–2743 (2012)
14. Linß, T.: Layer-adapted meshes for reaction-convection-diffusion problems, volume 1985 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, (2010)
15. Linss, T., Madden, N.: Analysis of an alternating direction method applied to singularly perturbed reaction-diffusion problems. *Int. J. Numer. Anal. Model.* **7**(3), 507–519 (2010)
16. Liu, F., Madden, N., Stynes, M., Zhou, A.: A two-scale sparse grid method for a singularly perturbed reaction-diffusion problem in two dimensions. *IMA J. Numer. Anal.* **29**(4), 986–1007 (2009)
17. Liu, X., Yang, M.: Error estimations in the balanced norm of finite element method on Bakhvalov-Shishkin triangular mesh for reaction-diffusion problems. *Appl. Math. Lett.* **123**, 1075,237 (2022)
18. Madden, N., Stynes, M.: A weighted and balanced FEM for singularly perturbed reaction-diffusion problems. *Calcolo* **58**(2), 28,16 (2021)
19. Roos, H.-G., Schopf, M.: Convergence and stability in balanced norms of finite element methods on Shishkin meshes for reaction-diffusion problems. *ZAMM Z. Angew. Math. Mech.* **95**(6), 551–565 (2015)
20. Roos, H.-G., Stynes, M., Tobiska, L.: Robust numerical methods for singularly perturbed differential equations, volume 24 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, second edition, Convection-diffusion-reaction and flow problems (2008)
21. Shishkin, G.I., Shishkina, L.P.: Difference methods for singular perturbation problems, Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, vol. 140. CRC Press, Boca Raton, FL (2009)
22. Stynes, M., Stynes, D.: Convection-diffusion problems, volume 196 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI; Atlantic Association for Research in the Mathematical Sciences (AARMS), Halifax, NS, An introduction to their analysis and numerical solution (2018)
23. Thomée, V.: Galerkin finite element methods for parabolic problems, volume 25 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, second edition (2006)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.