Computational Methods in Applied Mathematics
Vol. 13 (2013), No. 3, pp. 251–279

Doi: 10.1515/cmam-2012-0006

DE GRUYTER

Some Error Estimates for the Finite Volume Element Method for a Parabolic Problem

Panagiotis Chatzipantelidis · Raytcho Lazarov · Vidar Thomée

Abstract — We study spatially semidiscrete and fully discrete finite volume element methods for the homogeneous heat equation with homogeneous Dirichlet boundary conditions and derive error estimates for smooth and nonsmooth initial data. We show that the results of our earlier work [Math. Comp. 81 (2012), 1–20] for the lumped mass method carry over to the present situation. In particular, in order for error estimates for initial data only in L_2 to be of optimal second order for positive time, a special condition is required, which is satisfied for symmetric triangulations. Without any such condition, only first order convergence can be shown, which is illustrated by a counterexample. Improvements hold for triangulations that are almost symmetric and piecewise almost symmetric.

2010 Mathematical subject classification: 65M60, 65M15.

Keywords: Finite Volume Method, Parabolic Partial Differential Equations, Nonsmooth Initial Data, Error Estimates.

1. Introduction

We consider the model initial-boundary value problem

$$u_t - \Delta u = 0, \text{ in } \Omega, \quad u = 0, \text{ on } \partial\Omega, \quad \text{for } t \geqslant 0, \quad \text{with } u(0) = v, \text{ in } \Omega, \tag{1.1}$$

where Ω is a bounded convex polygonal domain in $\mathbb{R}^2$. We restrict ourselves to the homogeneous heat equation, thus without a forcing term, so that the initial values v are the only data of the problem. This problem has a unique solution $u(t)$, under appropriate assumptions on v, and this solution is smooth for $t > 0$, even if v is not.

To express the smoothness properties of the solution of (1.1), let, for $q \geqslant 0$, $\dot{H}^q \subset L_2(\Omega)$ be the Hilbert space defined by the norm

$$|w|_q = \Big(\sum_{j=1}^{\infty} \lambda_j^q (w, \phi_j)^2\Big)^{1/2}, \quad \text{where } (w, \varphi) = \int_\Omega w\varphi \, dx, \tag{1.2}$$

Panagiotis Chatzipantelidis
Department of Mathematics, University of Crete, 71409 Heraklion, Greece
E-mail: `chatzipa@math.uoc.gr`.

Raytcho Lazarov
Department of Mathematics, Texas A&M University, College Station, TX 77843, USA; and Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 8, 1113 Sofia, Bulgaria
E-mail: `lazarov@math.tamu.edu`.

Vidar Thomée
Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, SE-41296 Göteborg, Sweden; and Institute of Applied and Computational Mathematics, FORTH, 71110 Heraklion, Greece
E-mail: `thomee@chalmers.se`.

and where $\{\lambda_j\}_{j=1}^{\infty}$, $\{\phi_j\}_{j=1}^{\infty}$ are the eigenvalues, in increasing order, and orthonormal eigenfunctions of $-\Delta$ in Ω, with homogeneous Dirichlet boundary conditions on $\partial\Omega$. Thus $|w|_0 = \|w\| = (w,w)^{1/2}$ is the norm in $L_2 = L_2(\Omega)$, $|w|_1 = \|\nabla w\|$ the norm in $H_0^1 = H_0^1(\Omega)$ and $|w|_2 = \|\Delta w\|$ is equivalent to the norm in $H^2(\Omega)$ when $w = 0$ on $\partial\Omega$. Eigenfunction expansion and Parseval's relation show for the solution $u(t) = E(t)v$ of (1.1) the stability and smoothing estimate

$$|E(t)v|_p \leqslant Ct^{-(p-q)/2}|v|_q, \quad \text{for } 0 \leqslant q \leqslant p, \quad \text{and} \quad t > 0. \tag{1.3}$$

In fact, since the smallest eigenvalue is positive, a factor of e^{-ct}, with $c > 0$, may be included in the right-hand side, and this holds for all our stability, smoothing and error estimates throughout our paper. Since our interest here is in small time we shall not keep track of this decay for large time below. We shall also use the norm $\|w\|_{\mathcal{C}^k} = \sum_{|\gamma|\leqslant k} \sup_{x\in\Omega}|D_x^\gamma w(x)|$ in $\mathcal{C}^k = \mathcal{C}^k(\bar{\Omega})$, with $\mathcal{C} = \mathcal{C}^0$, the space of continuous functions on $\bar{\Omega}$. Here for $\gamma = (\gamma_1, \gamma_2)$, $D_x^\gamma = (\partial/\partial x_1)^{\gamma_1}(\partial/\partial x_2)^{\gamma_2}$ and $|\gamma| = \gamma_1 + \gamma_2$.

We first recall some facts about the spatially semidiscrete standard Galerkin finite element method for (1.1) in the space of piecewise linear functions

$$S_h = \{\chi \in \mathcal{C} : \chi|_\tau \text{ linear}, \forall\, \tau \in \mathcal{T}_h;\ \chi|_{\partial\Omega} = 0\},$$

where $\{\mathcal{T}_h\}$ is a family of regular triangulations $\mathcal{T}_h = \{\tau\}$ of Ω, with h denoting the maximum diameter of the triangles $\tau \in \mathcal{T}_h$. This method defines an approximation $u_h(t) \in S_h$ of $u(t)$, for $t \geqslant 0$, from

$$(u_{h,t}, \chi) + (\nabla u_h, \nabla\chi) = 0, \quad \forall\, \chi \in S_h, \quad \text{for } t \geqslant 0, \quad \text{with } u_h(0) = v_h, \tag{1.4}$$

where $v_h \in S_h$ is an approximation of v. It is well known that we have the smooth data error estimate, valid uniformly down to $t = 0$, see, e.g., [12],

$$\|u_h(t) - u(t)\| \leqslant Ch^2|v|_2, \quad \text{if } \|v_h - v\| \leqslant Ch^2|v|_2, \quad \text{for } t \geqslant 0. \tag{1.5}$$

We also have a nonsmooth data error estimate, for v only assumed to be in L_2, which is of optimal order $O(h^2)$ for t bounded away from zero, but deteriorates as $t \to 0$,

$$\|u_h(t) - u(t)\| \leqslant Ch^2t^{-1}\|v\|, \quad \text{if } v_h = P_h v, \quad \text{for } t > 0, \tag{1.6}$$

where P_h denotes the orthogonal L_2-projection onto S_h. Note that the choice of discrete initial data is not as general in this case as in (1.5). We emphasize that the triangulations $\mathcal{T}_h$ are assumed to be independent of t, and thus the use of finer $\mathcal{T}_h$ for t small is not considered here.

We note that a possible choice in (1.5) is $v_h = P_h v$, and hence, by interpolation, we have the intermediate result between (1.5) and (1.6),

$$\|u_h(t) - u(t)\| \leqslant Ch^2t^{-1/2}|v|_1, \quad \text{if } v_h = P_h v, \quad \text{for } t > 0. \tag{1.7}$$

Recently, in [4], we showed results similar to (1.5)–(1.7) for the lumped mass finite element method, which may be defined by replacing the L_2-inner product in the first term in (1.4) by the quadrature approximation $(u_{h,t}, \chi)_h$, where, with $I_h : \mathcal{C} \to S_h$ being the interpolant defined by $I_h v(z) = v(z)$ for any vertex z of $\mathcal{T}_h$,

$$(\chi, \psi)_h = \int_\Omega I_h(\chi\,\psi)\,dx, \quad \forall\, \chi, \psi \in S_h.$$

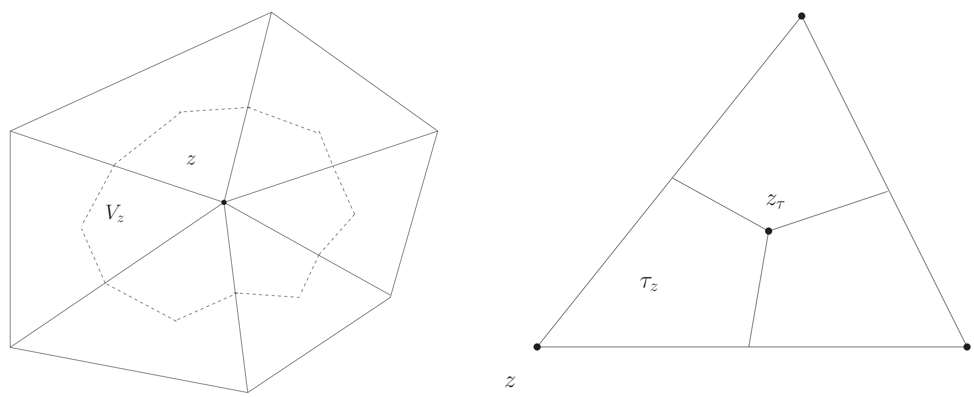

Figure 1. *Left:* A union of triangles that have a common vertex z; the dotted line shows the boundary of the corresponding control volume V_z. *Right:* A triangle τ partitioned into the three subregions τ_z.

Improving earlier results, we demonstrated that (1.5) remains valid for the lumped mass method, but that (1.6) requires restrictive conditions on $\{\mathcal{T}_h\}$, caused by the use of quadrature in (1.4), and satisfied, in particular, for symmetric triangulations. We remark that the choice of discrete initial data in the analogue of (1.7) was incorrectly stated in [4], see Section 3 below.

In the present paper our purpose is to carry over the analysis in [4] to the finite volume element method for problem (1.1). This method is based on a local conservation property associated with the differential equation. Namely, integrating (1.1) over any region $V \subset \Omega$ and using Green's formula, we obtain

$$\int_V u_t\,dx - \int_{\partial V} \nabla u \cdot n\,d\sigma = 0, \quad \text{for } t \geqslant 0, \tag{1.8}$$

where n denotes the unit exterior normal vector to ∂V. The semidiscrete finite volume element approximation $\tilde{u}_h(t) \in S_h$ will satisfy (1.8) for V in a finite collection of subregions of Ω called control volumes, the number of which will be equal to the dimension of the finite element space S_h. These control volumes are constructed in the following way. Let z_τ be the barycenter of $\tau \in \mathcal{T}_h$. We connect z_τ by line segments to the midpoints of the edges of τ, thus partitioning τ into three quadrilaterals τ_z, $z \in Z_h(\tau)$, where $Z_h(\tau)$ are the vertices of τ. Then with each vertex $z \in Z_h = \bigcup_{\tau\in\mathcal{T}_h} Z_h(\tau)$ we associate a control volume V_z, which consists of the union of the subregions τ_z, sharing the vertex z (see Figure 1, left). We denote the set of interior vertices of Z_h by Z_h^0. The semidiscrete finite volume element method for (1.1) is then to find $\tilde{u}_h(t) \in S_h$ such that

$$\int_{V_z} \tilde{u}_{h,t}\,dx - \int_{\partial V_z} \nabla\tilde{u}_h \cdot n\,d\sigma = 0, \quad \forall\, z \in Z_h^0, \quad \text{for } t \geqslant 0, \quad \text{with } \tilde{u}_h(0) = v_h, \tag{1.9}$$

where $v_h \in S_h$ is an approximation of v.

This problem may also be expressed in a weak form. For this purpose we introduce the finite-dimensional space of piecewise constant functions

$$Y_h = \{\eta \in L_2 : \eta|_{V_z} = \text{constant}, \forall\, z \in Z_h^0;\ \eta|_{V_z} = 0, \forall\, z \in Z_h \setminus Z_h^0\}.$$

We now multiply (1.9) by $\eta(z)$ for an arbitrary $\eta \in Y_h$, and sum over $z \in Z_h^0$ to obtain the Petrov–Galerkin formulation

$$(\tilde{u}_{h,t}, \eta) + a_h(\tilde{u}_h, \eta) = 0, \quad \forall\, \eta \in Y_h, \quad \text{for } t \geqslant 0, \quad \text{with } \tilde{u}_h(0) = v_h, \tag{1.10}$$

where the bilinear form $a_h(\cdot,\cdot) : S_h \times Y_h \to \mathbb{R}$ is defined by

$$a_h(\chi,\eta) = -\sum_{z\in Z_h^0} \eta(z) \int_{\partial V_z} \nabla\chi \cdot n \, d\sigma, \quad \forall\, \chi \in S_h,\ \eta \in Y_h. \tag{1.11}$$

Obviously, we can define $a_h(\cdot,\cdot)$ also for χ replaced by $w \in H^2$, and using Green's formula we then easily see that

$$a_h(w,\eta) = -(\Delta w,\eta), \quad \forall\, w \in H^2,\ \eta \in Y_h.$$

We shall now rewrite the Petrov–Galerkin method (1.10) as a Galerkin method in S_h. For this purpose, we introduce the interpolation operator $J_h : \mathcal{C} \mapsto Y_h$ by

$$J_h u = \sum_{z\in Z_h^0} u(z)\Psi_z,$$

where Ψ_z is the characteristic function of the control volume V_z. It is known that J_h is selfadjoint and positive definite, see [5], and hence the following defines an inner product $\langle\cdot,\cdot\rangle$ on S_h,

$$\langle\chi,\psi\rangle = (\chi, J_h\psi), \quad \forall\, \chi,\psi \in S_h. \tag{1.12}$$

Also, the corresponding discrete norm is equivalent to the L_2-norm, uniformly in h, i.e., with $C \geqslant c > 0$,

$$c\|\chi\| \leqslant |||\chi||| \leqslant C\|\chi\|, \quad \forall\, \chi \in S_h, \quad \text{where } |||\chi||| \equiv \langle\chi,\chi\rangle^{1/2},$$

see [5]. Further, in [2], it is shown that

$$a_h(\chi, J_h\psi) = (\nabla\chi, \nabla\psi), \quad \forall\, \chi,\psi \in S_h,$$

and therefore, $a_h(\cdot,\cdot)$ is symmetric and $a_h(\chi, J_h\chi) = \|\nabla\chi\|^2$, for $\chi \in S_h$.

With this notation, (1.10) may equivalently be written in Galerkin form as

$$\langle\tilde u_{h,t}, \chi\rangle + (\nabla\tilde u_h, \nabla\chi) = 0, \quad \forall\, \chi \in S_h, \quad \text{for } t \geqslant 0, \quad \text{with } \tilde u_h(0) = v_h. \tag{1.13}$$

Our aim is thus to show analogues of (1.5)–(1.7) for the solution of (1.13), with the appropriate choices of v_h, i.e.,

$$\|\tilde u_h(t) - u(t)\| \leqslant Ch^2 t^{-1+q/2} |v|_q, \quad \text{for } t > 0, \quad q = 0,1,2. \tag{1.14}$$

This will be done below for $q = 2$, and in the case $q = 1$ under the additional assumption that $\{\mathcal{T}_h\}$ is quasi-uniform. However, for $q = 0$, as in [4], we are only able to show (1.14) under an additional hypothesis, expressed in terms of the quadrature error operator $Q_h : S_h \to S_h$, defined by

$$(\nabla Q_h\psi, \nabla\chi) = \varepsilon_h(\psi,\chi), \quad \forall\, \chi,\psi \in S_h, \tag{1.15}$$

where $\varepsilon_h(\cdot,\cdot)$ is the quadrature error defined here by

$$\varepsilon_h(f,\chi) = (f, J_h\chi) - (f,\chi), \quad \forall\, f \in L_2,\ \chi \in S_h, \tag{1.16}$$

and requiring

$$\|Q_h\psi\| \leqslant Ch^2\|\psi\|, \quad \forall\, \psi \in S_h. \tag{1.17}$$

We will show that this assumption is satisfied for *symmetric* triangulations $\mathcal{T}_h$. Symmetry of $\mathcal{T}_h$, however, is a severe restriction which can only hold for special shapes of Ω. For this reason we will also consider less restrictive families $\{\mathcal{T}_h\}$. We will demonstrate that (1.17) holds for *almost symmetric* families (discussed in Section 4), with the addition of a logarithmic factor; we also show that this logarithmic factor is not needed in one space dimension. Further, for *piecewise almost symmetric* families of triangulations, see Section 4, the inequality (1.17) holds with an $O(h^{3/2})$ bound.

We then give two examples of nonsymmetric triangulations such that (1.14) does not hold for $q = 0$. In the first example we construct $\{\mathcal{T}_h\}$ such that the convergence factor is at most of order $O(h)$ for $t > 0$, and in the second example, with nonsymmetry only along a line, of order $O(h^{3/2})$. Without any additional condition on $\mathcal{T}_h$ we are only able to show the nonoptimal order error estimate

$$\|\tilde{u}_h(t) - u(t)\| \leqslant Cht^{-1/2}\|v\|, \quad \text{if } v_h = P_h v, \quad \text{for } t > 0.$$

We remark that in [11], in the more general case of a parabolic integro-differential equation, the nonsmooth data error estimate (1.14), for $q = 0$, with an extra factor $|\log h|$, was stated for any quasi-uniform family $\{\mathcal{T}_h\}$. Unfortunately, this result is in contradiction to our above counterexamples, and its proof incorrect.

We also discuss optimal order $O(h)$ error estimates for the gradient of $\tilde{u}_h - u$, under various assumptions on the smoothness of v and choices of v_h. Further, in a separate section, we consider briefly the extension of our results for the spatially semidiscrete problem to the fully discrete backward Euler and Crank–Nicolson finite volume methods.

As for the lumped mass method in [4], our analysis yields improvements of earlier results, in [3], where it was shown that, for smooth initial data and $v_h = R_h v$,

$$\|\tilde{u}_h(t) - u(t)\| \leqslant Ch^2|v|_3, \quad \text{for } t > 0,$$

and

$$\|\nabla(\tilde{u}_h(t) - u(t))\| \leqslant Ch\epsilon^{-1}|v|_{2+\epsilon}, \quad \text{for } t > 0,\ \epsilon > 0 \text{ small}.$$

As in the case of the lumped mass method in [4], these improvements are made possible by combining the error estimates (1.5)–(1.7) for the standard Galerkin finite element method with bounds for the difference $\delta = \tilde{u}_h - u_h$, which, by (1.13) and (1.4), satisfies

$$\langle \delta_t, \chi \rangle + (\nabla\delta, \nabla\chi) = -\varepsilon_h(u_{h,t}, \chi), \quad \forall\, \chi \in S_h, \quad \text{for } t \geqslant 0. \tag{1.18}$$

In the final section we sketch the extension of the theory developed above to more general parabolic equations, considering the initial-boundary value problem

$$u_t + Au = 0, \text{ in } \Omega, \quad u = 0, \text{ on } \partial\Omega, \quad \text{for } t \geqslant 0, \quad \text{with } u(0) = v, \text{ in } \Omega, \tag{1.19}$$

where $Au = -\nabla \cdot (\alpha \nabla u) + \beta u$, with α a smooth symmetric, positive definite 2×2 matrix function on $\bar{\Omega}$ and β a non-negative smooth function.

Here, let $u_h(t) \in S_h$ denote the standard Galerkin finite element approximation of $u(t)$, defined by

$$(u_{h,t}, \chi) + a(u_h, \chi) = 0, \quad \forall\, \chi \in S_h, \quad \text{for } t \geqslant 0, \quad \text{with } u_h(0) = v_h, \tag{1.20}$$

where $v_h \in S_h$ is an approximation of v and

$$a(w, \varphi) = (\alpha\nabla w, \nabla\varphi) + (\beta w, \varphi), \quad \text{for } w, \varphi \in H_0^1. \tag{1.21}$$

In a straightforward way the estimates (1.5)–(1.7) extend to the solution of (1.20).

The natural generalization of the finite volume method (1.10) would now be to find $\tilde{u}_h(t) \in S_h$ such that

$$\langle \tilde{u}_{h,t}, \chi \rangle + a_h(\tilde{u}_h, J_h\chi) = 0, \quad \forall\, \chi \in S_h, \quad \text{for } t \geqslant 0, \quad \text{with } \tilde{u}_h(0) = v_h, \tag{1.22}$$

where, instead of (1.11), one uses the bilinear defined by

$$a_h(\psi, \eta) = \sum_{z \in Z_h^0} \eta(z)\Big(-\int_{\partial V_z} (\alpha \nabla \psi) \cdot n \, d\sigma + \int_{V_z} \beta \psi \, dx\Big), \quad \forall\, \psi \in S_h, \; \eta \in Y_h. \tag{1.23}$$

It is known that, in general, the bilinear form $a_h(\psi, J_h\chi)$ is nonsymmetric on S_h but it is not far from being symmetric, or $|a_h(\chi, J_h\psi) - a_h(\psi, J_h\chi)| \leqslant Ch\|\nabla \chi\| \, \|\nabla \psi\|$, cf. [5]. Also, if α and β are constants over each $\tau \in \mathcal{T}_h$, then, see, e.g., [2, 6],

$$a_h(\psi, J_h\chi) = (\alpha \nabla \psi, \nabla \chi) + (\beta \psi, J_h\chi), \quad \forall\, \psi, \chi \in S_h, \tag{1.24}$$

and thus $a_h(\psi, J_h\chi)$ is symmetric, since as we shall show $(\beta\psi, J_h\chi) = (\beta\chi, J_h\psi)$. Therefore, since symmetry is important in our analysis, we introduce the modified bilinear form

$$\tilde{a}_h(\psi, \eta) = \sum_{z \in Z_h^0} \eta(z)\Big(-\int_{\partial V_z} (\tilde{\alpha} \nabla \psi) \cdot n \, d\sigma + \int_{V_z} \tilde{\beta} \psi \, dx\Big), \quad \forall\, \psi \in S_h, \; \eta \in Y_h, \tag{1.25}$$

where, for $z \in \tau$, $\tau \in \mathcal{T}_h$, $\tilde{\alpha}(z) = \alpha(z_\tau)$ and $\tilde{\beta}(z) = \beta(z_\tau)$, with z_τ the barycenter of τ. This choice of $\tilde{a}_h(\cdot,\cdot)$ leads to the finite volume element method, to find $\tilde{u}_h(t) \in S_h$ such that

$$\langle \tilde{u}_{h,t}, \chi \rangle + \tilde{a}_h(\tilde{u}_h, J_h\chi) = 0, \quad \forall\, \chi \in S_h, \quad \text{for } t \geqslant 0, \quad \text{with } \tilde{u}_h(0) = v_h, \tag{1.26}$$

and for this the desired analogues of the estimates (1.14) are established in Theorems 7.1–7.3.

The following is an outline of the paper. In Section 2, we introduce notation and give some preliminary material needed for the analysis of the finite volume element method. Further, we derive smooth and nonsmooth initial data estimates for the gradient of the error in the standard Galerkin method. In Section 3 we derive the error estimates (1.14) discussed above under the different assumptions on smoothness of data and the triangulations $\{\mathcal{T}_h\}$. In Section 4 we show that assumption (1.17) is valid for symmetric meshes, and discuss the corresponding properties for almost symmetric and piecewise almost symmetric meshes. In Section 5 we present two nonsymmetric triangulations in two space dimensions for which optimal order L_2-convergence for nonsmooth data does not hold. In Section 6 we consider briefly the application to the fully discrete backward Euler and Crank–Nicolson finite volume methods. Finally, Section 7 contains the extension of Section 3 to more general parabolic equations.

2. Preliminaries

In this section we show a smoothing property for the finite volume element method, and discuss the quadrature associated with this method. We also derive some estimates for the gradient of the error in the standard Galerkin finite element method which will be needed later.

We first recall that for the standard Galerkin method, one may introduce the discrete Laplacian $\Delta_h : S_h \to S_h$ by

$$-(\Delta_h \psi, \chi) = (\nabla \psi, \nabla \chi), \quad \forall\, \psi, \chi \in S_h,$$

and write the problem (1.4) as

$$u_{h,t} - \Delta_h u_h = 0, \quad \text{for } t \geqslant 0, \quad \text{with } u_h(0) = v_h. \tag{2.1}$$

Let $\{\lambda_j^h\}_{j=1}^{N_h}$ denote the eigenvalues, in increasing order, and $\{\phi_j^h\}_{j=1}^{N_h}$ the corresponding eigenfunctions of $-\Delta_h$, orthonormal with respect to $(\cdot,\cdot)$, where $N_h = \dim S_h$. Then we have for the solution operator $E_h(t) = e^{\Delta_h t}$ of (2.1), by eigenfunction expansion,

$$u_h(t) = E_h(t)v_h = \sum_{j=1}^{N_h} e^{-\lambda_j^h t}(v_h, \phi_j^h)\phi_j^h, \quad \text{for } t \geqslant 0.$$

The following smoothing property analogous to (1.3) holds for $v_h \in S_h$ and $t > 0$,

$$\|\nabla^p D_t^\ell E_h(t) v_h\| \leqslant C t^{-\ell-(p-q)/2} \|\nabla^q v_h\|, \quad \ell \geqslant 0,\ p, q = 0, 1,\ 2\ell + p \geqslant q, \tag{2.2}$$

with $D_t = \partial/\partial t$.

Turning to the finite volume method (1.13), we now introduce the discrete Laplacian $\widetilde{\Delta}_h : S_h \to S_h$, corresponding to the inner product $\langle \cdot, \cdot \rangle$ in (1.12), by

$$-\langle \widetilde{\Delta}_h \psi, \chi \rangle = (\nabla \psi, \nabla \chi), \quad \forall\, \psi, \chi \in S_h. \tag{2.3}$$

The finite volume method (1.13) can then be written in operator form as

$$\tilde{u}_{h,t} - \widetilde{\Delta}_h \tilde{u}_h = 0, \quad \text{for } t \geqslant 0, \quad \text{with } \tilde{u}_h(0) = v_h. \tag{2.4}$$

For the solution operator $\widetilde{E}_h(t) = e^{\widetilde{\Delta}_h t}$ of (2.4) we have

$$\tilde{u}_h(t) = \widetilde{E}_h(t) v_h = \sum_{j=1}^{N_h} e^{-\tilde{\lambda}_j^h t} \langle v_h, \tilde{\phi}_j^h \rangle \tilde{\phi}_j^h, \quad \text{for } t \geqslant 0, \tag{2.5}$$

where $\{\tilde{\lambda}_j^h\}_{j=1}^{N_h}$ and $\{\tilde{\phi}_j^h\}_{j=1}^{N_h}$ are the eigenvalues, in increasing order, and the corresponding eigenfunctions, orthonormal with respect to $\langle \cdot, \cdot \rangle$, of the positive definite operator $-\widetilde{\Delta}_h$. For $\widetilde{E}_h(t)$ the following analogue of (2.2) holds, cf. [4, Lemma 2.1].

Lemma 2.1. *For $\widetilde{E}_h$ defined by* (2.5) *we have, for $v_h \in S_h$ and $t > 0$,*

$$\|\nabla^p D_t^\ell \widetilde{E}_h(t) v_h\| \leqslant C t^{-\ell-(p-q)/2} \|\nabla^q v_h\|, \quad \ell \geqslant 0,\ p, q = 0, 1,\ 2\ell + p \geqslant q.$$

Proof. Introducing the square root $\widetilde{G}_h = (-\widetilde{\Delta}_h)^{1/2} : S_h \to S_h$ of $-\widetilde{\Delta}_h$, we get

$$\|\nabla v_h\|^2 = \langle (-\widetilde{\Delta}_h) v_h, v_h \rangle = \sum_{j=1}^{N_h} \tilde{\lambda}_j^h \langle v_h, \tilde{\phi}_j^h \rangle^2 = |||\widetilde{G}_h v_h|||^2.$$

Since the norms $|||\cdot|||$ and $\|\cdot\|$ are equivalent on S_h, we find, for $t > 0$,

$$\begin{aligned} \|\nabla^p D_t^\ell \widetilde{E}_h(t) v_h\|^2 &\leqslant C |||\widetilde{G}_h^p D_t^\ell \widetilde{E}_h(t) v_h|||^2 = C \sum_{j=1}^{N_h} (\tilde{\lambda}_j^h)^{2\ell+p-q} e^{-2\tilde{\lambda}_j^h t} (\tilde{\lambda}_j^h)^q \langle v_h, \tilde{\phi}_j^h \rangle^2 \\ &\leqslant C\, t^{-(2\ell+p-q)} |||\widetilde{G}_h^q v_h|||^2 \leqslant C\, t^{-(2\ell+p-q)} \|\nabla^q v_h\|^2. \quad \square \end{aligned}$$

The quadrature error functional $\varepsilon_h(\cdot,\cdot)$ defined by (1.16) has an important role in our analysis below. For this reason we recall the following lemma, cf. [3].

Lemma 2.2. *For the error functional* ε_h, *defined by* (1.16), *we have*

$$|\varepsilon_h(f,\psi)| \leqslant Ch^{p+q}\|\nabla^p f\|\,\|\nabla^q\psi\|, \quad \forall\, f\in H^1,\, \psi\in S_h, \quad \textit{and} \quad p,q=0,1.$$

Proof. Since $\int_\tau (J_h\psi-\psi)\,dx = 0$ for ψ linear in τ, for any $\tau\in\mathcal{T}_h$, see [5], we have that $J_h\psi-\psi$ is orthogonal to $\bar S_h$, the set of piecewise constants on $\mathcal{T}_h$. Hence

$$\varepsilon_h(f,\psi) = (f, J_h\psi-\psi) = (f-\bar P_h f, J_h\psi-\psi),$$

where $\bar P_h$ is the orthogonal projection onto $\bar S_h$. The lemma now easily follows since we have $\|J_h\psi-\psi\|\leqslant Ch\|\nabla\psi\|$ and $\|\bar P_h f - f\|\leqslant Ch\|\nabla f\|$. □

The following estimate holds for the quadrature error operator Q_h in (1.15).

Lemma 2.3. *Let* $\widetilde\Delta_h$ *and* Q_h *be the operators defined by* (2.3) *and* (1.15). *Then*

$$\|\nabla Q_h\chi\| + h\|\widetilde\Delta_h Q_h\chi\| \leqslant Ch^{p+1}\|\nabla^p\chi\|, \quad \forall\,\chi\in S_h,\ p=0,1.$$

Proof. By (1.15) and Lemma 2.2, with $\psi=Q_h\chi$ and $q=1$, it follows easily that

$$\|\nabla Q_h\chi\|^2 = \varepsilon_h(\chi, Q_h\chi) \leqslant Ch^{p+1}\|\nabla^p\chi\|\,\|\nabla Q_h\chi\|, \quad \text{for } p=0,1,$$

which shows the desired estimate for $\|\nabla Q_h\chi\|$. Also, by the definition of $\widetilde\Delta_h$, Lemma 2.2 with $q=0$ shows, for $p=0,1$,

$$|||\widetilde\Delta_h Q_h\chi|||^2 = -(\nabla Q_h\chi, \nabla\widetilde\Delta_h Q_h\chi) = -\varepsilon_h(\chi, \widetilde\Delta_h Q_h\chi) \leqslant Ch^p\|\nabla^p\chi\|\,\|\widetilde\Delta_h Q_h\chi\|.$$

Since the norms $|||\cdot|||$ and $\|\cdot\|$ are equivalent on S_h, this implies the bound for the remaining term $\|\widetilde\Delta_h Q_h\chi\|$. □

In addition to the orthogonal L_2-projection P_h, our error analysis will use the Ritz projection $R_h: H_0^1\to S_h$ defined by

$$(\nabla R_h w, \nabla\chi) = (\nabla w, \nabla\chi), \quad \forall\,\chi\in S_h.$$

It is well known that R_h satisfies

$$\|R_h w - w\| + h\|\nabla(R_h w - w)\| \leqslant Ch^q|w|_q, \quad \text{for } w\in \dot H^q,\ q=1,2. \tag{2.6}$$

We close with some estimates for the gradient of the error, slightly generalizing those of [4, Theorem 2.1].

Theorem 2.1. *Let* u *and* u_h *be the solutions of* (1.1) *and* (2.1). *Then, for* $t>0$,

$$\|\nabla(u_h(t)-u(t))\| \leqslant \begin{cases} Ch|v|_2, & \textit{if } \|\nabla(v_h-v)\|\leqslant Ch|v|_2,\\ Cht^{-1/2}|v|_1, & \textit{if } \|v_h-v\|\leqslant Ch|v|_1,\\ Cht^{-1}\|v\|, & \textit{if } v_h=P_h v.\end{cases}$$

Proof. In [4, Theorem 2.1] this was shown with $v_h=R_h v$ in the first two estimates, and thus it remains to bound $\nabla E_h(t)(v_h-R_h v)$. With $\vartheta := v_h - R_h v$ we find easily, by Lemma 2.1, for smooth data, $\|\nabla E_h(t)\vartheta(0)\|\leqslant\|\nabla\vartheta(0)\|\leqslant Ch|v|_2$, and for mildly nonsmooth data, $\|\nabla E_h(t)\vartheta(0)\|\leqslant Ct^{-1/2}\|\vartheta(0)\|\leqslant Ct^{-1/2}h|v|_1$. □

3. Smooth and Nonsmooth Initial Data Error Estimates

In this section we derive optimal order error estimates for the finite volume element method (1.13), with initial data v in $\dot H^2$, $\dot H^1$ and L_2. For $v \in \dot H^2$, the error estimate is the same as that for the standard Galerkin finite element method, and this is also the case for $v \in \dot H^1$, provided the family of finite element spaces is quasi-uniform. In the case $v \in L_2$, with discrete initial data $v_h = P_h v$, in order to derive an optimal order estimate analogous to (1.6), we need to impose condition (1.17) for the quadrature error operator Q_h. In Section 4 we verify this condition for symmetric meshes. In the general case we are only able to show a nonoptimal order $O(h)$ error bound in L_2, whereas for the gradient of the error an optimal order $O(h)$ bound still holds.

The estimates and their proofs are analogous to those for the lumped mass method derived in [4], since the operators $\widetilde E_h$, $\widetilde\Delta_h$ and Q_h, defined in Section 2, have properties similar to those of the corresponding operators for the lumped mass method. References to [4] will therefore be given in some of the proofs below. We begin with smooth initial data, $v \in \dot H^2$.

Theorem 3.1. *Let u and $\tilde u_h$ be the solutions of* (1.1) *and* (2.4)*. Then*

$$\|\tilde u_h(t) - u(t)\| \leqslant Ch^2|v|_2, \quad \text{if } \|v_h - v\| \leqslant Ch^2|v|_2, \quad \text{for } t \geqslant 0.$$

Proof. Since, by (1.5), the corresponding error bound holds for the solution u_h of the standard Galerkin method, it suffices to consider the difference $\delta = \tilde u_h - u_h$. Also, by the stability estimates of Lemma 2.1, we may assume that $v_h = R_h v$. By the definition (1.15) of Q_h, δ satisfies (1.18), and hence

$$\delta_t - \widetilde\Delta_h \delta = \widetilde\Delta_h Q_h u_{h,t}, \quad \text{for } t \geqslant 0, \quad \text{with } \delta(0) = 0, \tag{3.1}$$

where u_h is the solution of (1.4). By Duhamel's principle this shows

$$\delta(t) = \int_0^t \widetilde E_h(t-s)\widetilde\Delta_h Q_h u_{h,t}(s)\, ds. \tag{3.2}$$

Using the fact that $\widetilde E_h(t)\widetilde\Delta_h = D_t \widetilde E_h(t)$, and Lemmas 2.1 and 2.3, we easily get

$$\|\widetilde E_h(t)\widetilde\Delta_h Q_h \chi\| \leqslant Ct^{-1/2}\|\nabla Q_h\chi\| \leqslant Ch^2 t^{-1/2}\|\nabla\chi\|, \quad \text{for } \chi \in S_h, \tag{3.3}$$

and hence

$$\|\delta(t)\| \leqslant Ch^2 \int_0^t (t-s)^{-1/2}\|\nabla u_{h,t}(s)\|\, ds.$$

Here, since $\Delta_h R_h = P_h \Delta$, we obtain, by first applying Lemma 2.1,

$$\|\nabla u_{h,t}(s)\| \leqslant Cs^{-1/2}\|u_{h,t}(0)\| = Cs^{-1/2}\|\Delta_h R_h v\| \leqslant Cs^{-1/2}\|\Delta v\| = Cs^{-1/2}|v|_2,$$

and hence

$$\|\delta(t)\| \leqslant Ch^2 \int_0^t (t-s)^{-1/2} s^{-1/2}\, ds\, |v|_2 = C\, h^2 |v|_2,$$

which completes the proof. □

We now consider mildly nonsmooth initial data, $v \in \dot{H}^1$. Here we shall need to assume the stability of P_h in $\dot{H}^1$, or $\|\nabla P_h w\| \leqslant C|w|_1$, which does not hold for arbitrary families of triangulations. However, a sufficient condition for such stability of P_h is the global quasi-uniformity of $\{\mathcal{T}_h\}$. Indeed, this assumption implies the inverse inequality $\|\nabla\chi\| \leqslant Ch^{-1}\|\chi\|$, which combined with the error bound $\|R_h w - w\| \leqslant Ch|w|_1$ shows the desired stability of P_h.

Theorem 3.2. *Let u and $\tilde{u}_h$ be the solutions of* (1.1) *and* (2.4)*. Then for $t > 0$*

$$\|\tilde{u}_h(t) - u(t)\| \leqslant Ch^2 t^{-1/2}|v|_1, \quad \text{if } v_h = P_h v \quad \text{and} \quad \|\nabla P_h v\| \leqslant C|v|_1.$$

Proof. Since by (1.7), the corresponding error estimate holds for the solution u_h of the standard Galerkin method (without the condition on ∇P_h), it suffices as above to bound $\delta = \tilde{u}_h - u_h$. We use (3.2) to write

$$\delta(t) = \Big\{\int_0^{t/2} + \int_{t/2}^t\Big\}\widetilde{E}_h(t-s)\widetilde{\Delta}_h Q_h u_{h,t}(s)\,ds = \delta_1(t) + \delta_2(t). \tag{3.4}$$

Using again (3.3), we have, since $\|\nabla u_{h,t}(s)\| \leqslant Cs^{-1}\|\nabla P_h v\| \leqslant Cs^{-1}\,|v|_1$, that

$$\|\delta_2(t)\| \leqslant Ch^2\int_{t/2}^t (t-s)^{-1/2}\,\|\nabla u_{h,t}(s)\|\,ds \leqslant Ch^2 t^{-1/2}\,|v|_1.$$

Integrating by parts, we obtain

$$\delta_1(t) = \big[\widetilde{E}_h(t-s)\widetilde{\Delta}_h Q_h u_h(s)\big]_0^{t/2} - \int_0^{t/2} D_s\widetilde{E}_h(t-s)\widetilde{\Delta}_h Q_h u_h(s)\,ds. \tag{3.5}$$

Employing (3.3), Lemmas 2.1 and 2.3 we now find, similarly to the above,

$$\begin{aligned}\|\delta_1(t)\| &\leqslant Ch^2 t^{-1/2}\big(\|\nabla u_h(t/2)\| + \|\nabla P_h v\|\big) + Ch^2\int_0^{t/2}(t-s)^{-3/2}\,\|\nabla u_h(s)\|\,ds\\ &\leqslant Ch^2 t^{-1/2}|v|_1.\end{aligned}$$

Together these estimates complete the proof. □

The analogous result and its proof also hold for the lumped mass method, which should replace the case $q = 1$ in [4, Theorem 3.1], since (1.7) does not hold for $v_h = R_h v$.

Next, we turn to the nonsmooth initial data error estimate.

Theorem 3.3. *Let u and $\tilde{u}_h$ be the solutions of* (1.1) *and* (2.4)*. If* (1.17) *holds and $v_h = P_h v$, then*

$$\|\tilde{u}_h(t) - u(t)\| \leqslant Ch^2 t^{-1}\|v\|, \quad \text{for } t > 0.$$

Proof. This follows easily from the fact that for Q_h satisfying (1.17) we have,

$$\|\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h P_h v\| \leqslant Ct^{-1}\|Q_h P_h v\| \leqslant Ch^2 t^{-1}\|v\|, \quad \text{for } t > 0. \tag{3.6}$$

This inequality is the necessary and sufficient condition for the desired bound to hold by the following lemma, which is proved in the same way as [4, Theorem 4.1]. □

Lemma 3.1. *Let u and $\tilde{u}_h$ be the solutions of* (1.1) *and* (2.4). *Then*

$$\|\tilde{u}_h(t) - u(t) + \widetilde{E}_h(t)\widetilde{\Delta}_h Q_h v_h\| \leqslant Ch^2 t^{-1}\|v\|, \quad \text{if } v_h = P_h v, \quad \text{for } t > 0.$$

Condition (1.17) will be discussed in more detail in Section 4 below. Note that, by Lemma 2.3, without additional assumptions on the mesh, we have

$$\|Q_h\chi\| \leqslant C\|\nabla Q_h\chi\| \leqslant Ch\|\chi\|, \quad \forall\, \chi \in S_h,$$

and that the lower order error estimate of the following theorem always holds. The proof is the same as that of [4, Theorem 4.3]. We shall show in Section 5 that a $O(h)$ bound is the best possible for general triangulation families $\{\mathcal{T}_h\}$.

Theorem 3.4. *Let u and $\tilde{u}_h$ be the solutions of* (1.1) *and* (2.4). *Then*

$$\|\tilde{u}_h(t) - u(t)\| \leqslant Cht^{-1/2}\|v\|, \quad \text{if } v_h = P_h v, \quad \text{for } t > 0.$$

We end this section by stating optimal order estimates for the gradient of the error. Note that no additional assumption on $\{\mathcal{T}_h\}$ is required.

Theorem 3.5. *Let u and $\tilde{u}_h$ be the solutions of* (1.1) *and* (2.4). *Then, for $t > 0$,*

$$\|\nabla(\tilde{u}_h(t) - u(t))\| \leqslant \begin{cases} Ch|v|_2, & \text{if } \|\nabla(v_h - v)\| \leqslant Ch|v|_2, \\ Cht^{-1/2}|v|_1, & \text{if } \|v_h - v\| \leqslant Ch|v|_1, \\ Cht^{-1}\|v\|, & \text{if } v_h = P_h v. \end{cases}$$

Proof. For the first two estimates it suffices, by the stability and smoothness estimates of Lemma 2.1, to consider $v_h = R_h v$. For this choice of the initial data the proofs are identical to those in [4, Theorem 3.1]. In the nonsmooth data case, the proof is the same as that of [4, Theorem 4.4]. □

4. Symmetric and Almost Symmetric Triangulations

In this section we first show that for families of triangulations $\{\mathcal{T}_h\}$ that are *symmetric*, in a sense to be defined below, assumption (1.17) is satisfied and therefore, by Theorem 3.3, the optimal order nonsmooth data error estimate holds. We shall then relax the symmetry requirements and consider *almost symmetric* families of triangulations, consisting of $O(h^2)$ perturbations of symmetric triangulations. In this case we show that (1.17) is satisfied with an additional logarithmic factor and, as a consequence, an almost optimal order nonsmooth data error estimate holds. Finally for the less restrictive class of *piecewise almost symmetric* families $\{\mathcal{T}_h\}$ we derive a $O(h^{3/2})$ order nonsmooth data error estimate.

In addition to the quadrature error operator Q_h defined in (1.16) we shall work with the symmetric operator $M_h : S_h \to S_h$, defined by

$$\varepsilon_h(\psi, \chi) = [\psi, M_h\chi], \quad \forall\, \psi, \chi \in S_h, \tag{4.1}$$

where we use the inner product

$$[\psi, \chi] = \sum_{z \in Z_h^0} \psi(z)\chi(z), \quad \forall\, \psi, \chi \in S_h. \tag{4.2}$$

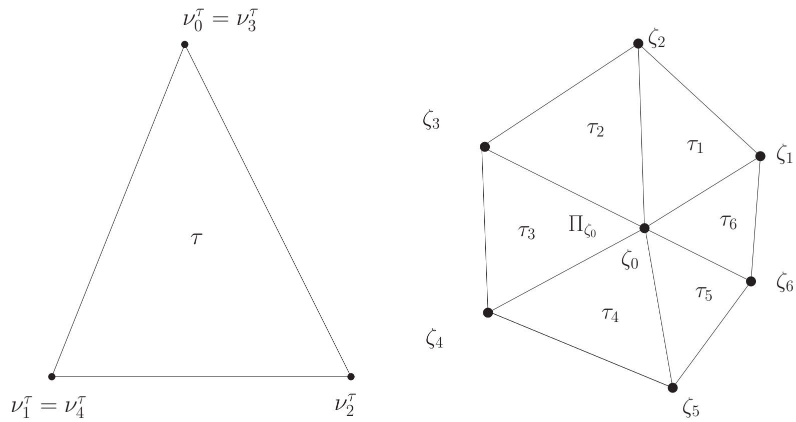

Figure 2. *Left:* A triangle τ. *Right:* A patch Π_{ζ_0} around a vertex ζ_0.

To determine the form of this operator, we introduce some notation. For $z \in Z_h^0$ an interior vertex of $\mathcal{T}_h$, we define the patch $\Pi_z = \{\bigcup \tau : \tau \in \mathcal{T}_h,\ z \in \partial\tau\}$, where for simplicity we have assumed that $\tau = \bar{\tau}$. Further, for z a vertex of $\tau \in \mathcal{T}_h$, we denote by z_+^τ and z_-^τ the other two vertices of τ. We then define

$$M_h^{\Pi_z}\chi := -\frac{1}{54}\sum_{\tau\subset\Pi_z}|\tau|\big(\chi(z_+^\tau) - 2\chi(z) + \chi(z_-^\tau)\big), \tag{4.3}$$

for which the following holds.

Lemma 4.1. *For the operator M_h defined by (4.1) we have, for $z \in Z_h^0$,*

$$M_h\chi(z) = M_h^{\Pi_z}\chi \quad \textit{with } M_h^{\Pi_z}\chi \textit{ given by (4.3).} \tag{4.4}$$

Proof. In view of (1.16), we may write

$$\varepsilon_h(\psi,\chi) = (\psi, J_h\chi) - (\psi,\chi) = \sum_{\tau\in\mathcal{T}_h}\int_\tau (\psi J_h\chi - \psi\chi)\,dx. \tag{4.5}$$

For $\tau \in \mathcal{T}_h$ we denote its vertices by $\nu_1^\tau, \nu_2^\tau, \nu_3^\tau$ and set $\nu_4^\tau = \nu_1^\tau$, $\nu_0^\tau = \nu_3^\tau$, see Figure 2. Writing $w_j = w(\nu_j^\tau)$ for a function w on τ, we obtain, after simple calculations,

$$\int_\tau \psi J_h\chi\,dx = \frac{|\tau|}{108}\sum_{j=1}^{3}\psi_j(22\chi_j + 7\chi_{j-1} + 7\chi_{j+1}), \tag{4.6}$$

and

$$\int_\tau \psi\chi\,dx = \frac{|\tau|}{12}\sum_{j=1}^{3}\psi_j(2\chi_j + \chi_{j-1} + \chi_{j+1}).$$

Thus

$$\int_\tau (\psi J_h\chi - \psi\,\chi)\,dx = -\frac{|\tau|}{54}\sum_{j=1}^{3}\psi_j(\chi_{j+1} - 2\chi_j + \chi_{j-1}).$$

Summation over $\tau \in \mathcal{T}_h$, (4.1) and (4.5) show

$$[\psi, M_h\chi] = \sum_{z\in Z_h^0}\psi(z)M_h^{\Pi_z}\chi, \quad \forall\,\psi,\chi \in S_h.$$

This implies (4.4) and thus completes the proof. □

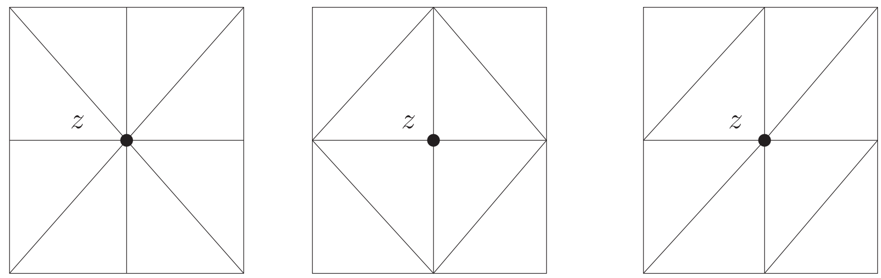

Figure 3. Patches which are symmetric with respect to the vertex z.

We say that $\mathcal{T}_h$ is *symmetric at* $z \in Z_h^0$, if the corresponding patch Π_z is symmetric around z, in the sense that if $x \in \Pi_z$, then $z - (x - z) = 2z - x \in \Pi_z$. We say that $\mathcal{T}_h$ is *symmetric* if it is symmetric at each $z \in Z_h^0$. The patch Π_{ζ_0} in Figure 2 is nonsymmetric with respect to ζ_0, whereas triangulations which are built up of either of the patches shown in Figure 3 are symmetric. Symmetric triangulations exist only for special domains, such as parallelograms, but not for general polygonal domains.

We now show the sufficiency of symmetry of $\{\mathcal{T}_h\}$ for condition (1.17) for the operator Q_h, and hence, by Theorem 3.3, for the nonsmooth data error estimate.

Theorem 4.1. *If the family* $\{\mathcal{T}_h\}$ *is symmetric, then* (1.17) *holds.*

Proof. The proof, by duality, follows that of [4, Theorem 5.1]. For given $\chi \in S_h$ we define $\varphi = \varphi_\chi \in \dot{H}^1$ as the solution of the Dirichlet problem $-\Delta\varphi = \chi$ in Ω, $\varphi = 0$ on $\partial\Omega$. Since Ω is convex, we have $\varphi \in \dot{H}^2$ and $|\varphi|_2 \leqslant C\|\chi\|$. With I_h the finite element interpolation operator into S_h, we have, for any $\psi \in S_h$,

$$\begin{aligned} \|Q_h\psi\| &= \sup_{\chi \in S_h} \frac{(Q_h\psi, \chi)}{\|\chi\|} = \sup_{\chi \in S_h} \frac{(\nabla Q_h\psi, \nabla\varphi)}{\|\chi\|} \\ &\leqslant \sup_{\chi \in S_h} \frac{|(\nabla Q_h\psi, \nabla(\varphi - I_h\varphi))|}{\|\chi\|} + \sup_{\chi \in S_h} \frac{|(\nabla Q_h\psi, \nabla I_h\varphi)|}{\|\chi\|} = I + II. \end{aligned} \tag{4.7}$$

By the obvious error estimate for I_h and Lemma 2.3, with $p = 0$, we find

$$|I| \leqslant Ch \sup_{\chi \in S_h} \frac{\|\nabla Q_h\psi\|\,|\varphi|_2}{\|\chi\|} \leqslant Ch^2\|\psi\|. \tag{4.8}$$

To estimate *II*, we employ (1.15) and (4.1) to rewrite the numerator in the form

$$(\nabla Q_h\psi, \nabla I_h\varphi) = \varepsilon_h(\psi, I_h\varphi) = [\psi, M_h I_h\varphi]. \tag{4.9}$$

To bound $M_h I_h\varphi$, we consider an arbitrary vertex $z = \zeta_0 \in Z_h^0$. Let Π_{ζ_0} be the corresponding patch of $\mathcal{T}_h$, with vertices $\{\zeta_j\}_{j=1}^K$, numbered counter-clockwise, with $\zeta_{j+K} = \zeta_j$ for all j. Also denote by $\{\tau_j\}_{j=1}^K$ the triangles of $\mathcal{T}_h$ in Π_{ζ_0}, with τ_j having vertices ζ_0, ζ_j, ζ_{j+1}, and set $\tau_0 = \tau_K$ (see Figure 2). Then Lemma 4.1 implies

$$M_h I_h\varphi(\zeta_0) = M_h^{\Pi_{\zeta_0}} I_h\varphi = -\frac{1}{54}\sum_{j=1}^{K} \omega_j(\varphi(\zeta_j) - \varphi(\zeta_0)), \tag{4.10}$$

with $\omega_j = |\tau_{j-1}| + |\tau_j|$. By assumption, the patch Π_{ζ_0} is symmetric and hence, by (4.10), we can express $M_h I_h\varphi(\zeta_0)$ as a linear combination of terms of the form $\varphi(\zeta_j) - 2\varphi(\zeta_0) + \varphi(\zeta_j')$,

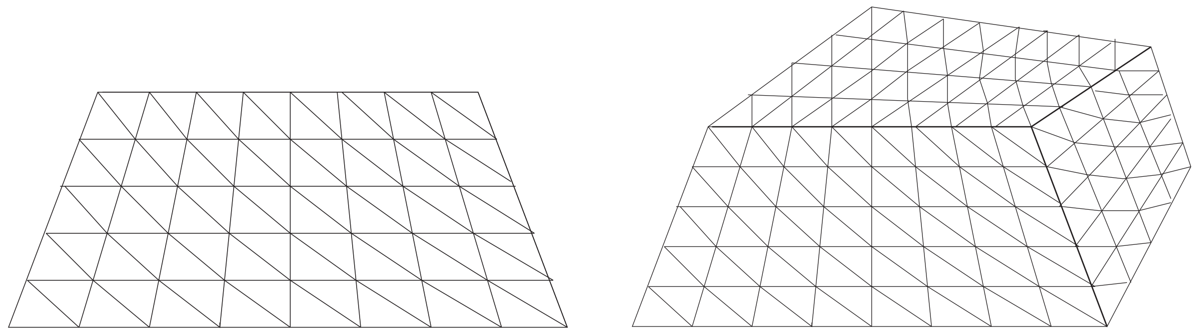

Figure 4. *Left:* An almost symmetric triangulation. *Right:* A piecewise almost symmetric triangulation.

where ζ_0 is the midpoint of the vertices ζ_j and ζ_j' of Π_{ζ_0}. Hence $M_h I_h \varphi(\zeta_0) = 0$ for φ linear in Π_{ζ_0} and, as in [4], we may apply the Bramble–Hilbert lemma to obtain

$$|M_h I_h \varphi(\zeta_0)| \leqslant Ch^2 |\Pi_{\zeta_0}|^{1/2} \|\varphi\|_{H^2(\Pi_{\zeta_0})} \leqslant Ch^3 \|\varphi\|_{H^2(\Pi_{\zeta_0})}. \tag{4.11}$$

Employing this estimate for all patches Π_z of $\mathcal{T}_h$, we obtain, for any $\psi \in S_h$,

$$\big|[\psi, M_h I_h \varphi]\big| \leqslant Ch^3 \sum_{z \in Z_h^0} |\psi(z)| \, \|\varphi\|_{H^2(\Pi_z)} \leqslant Ch^2 \|\psi\| \, |\varphi|_2 \leqslant Ch^2 \|\psi\| \, \|\chi\|. \tag{4.12}$$

Hence, in view of (4.7) and (4.9), we obtain $|II| \leqslant Ch^2 \|\psi\|$. Together with (4.8) this completes the proof. □

We now want to slightly weaken the assumption about symmetry. We say that a family of triangulations $\{\mathcal{T}_h\}$ is *almost symmetric* if each $\mathcal{T}_h$ is a perturbation by $O(h^2)$ of a symmetric triangulation, uniformly in h, in the sense that with each patch Π_z of $\mathcal{T}_h$ there is an associated symmetric patch from which Π_z is obtained by moving each of its vertices by $O(h^2)$. Such triangulations exist for any convex quadrilateral, cf. Figure 4. We note that various special triangulations have been used in the past for obtaining higher order accuracy for the gradient of the finite element solution (super-convergent rates of $O(h^2)$ or $O(h^2 \ell_h)$), see, e.g., [7, 10, 13]. For example, the strongly regular triangulations from [10], requiring that any two adjacent triangles form almost a parallelogram (a deviation of a parallelogram by $O(h^2)$), are almost symmetric meshes in our terminology. We shall show that, in this case, we have almost optimal order convergence for nonsmooth initial data.

Theorem 4.2. *If the family $\{\mathcal{T}_h\}$ is almost symmetric, then*

$$\|Q_h \psi\| \leqslant Ch^2 \ell_h^{1/2} \, \|\psi\|, \quad \forall\, \psi \in S_h, \quad \text{where } \ell_h = 1 + |\log h|. \tag{4.13}$$

Hence, for the solution of (1.13), *with $v_h = P_h v$, we have*

$$\|\tilde{u}_h(t) - u(t)\| \leqslant Ch^2 \ell_h^{1/2} \, t^{-1} \|v\|, \quad \text{for } t > 0. \tag{4.14}$$

In the proof we shall need the following Sobolev type inequality, where the $|\cdot|_{H^k}$ denote seminorms with only the derivatives of highest order k.

Lemma 4.2. *Let B be a fixed bounded domain, satisfying the cone property. Then we have, for $0 < \epsilon < 1$,*

$$\sup_{z, z' \in B,\, z' \neq z} \frac{|\varphi(z') - \varphi(z)|}{|z' - z|^{1-\epsilon}} \leqslant C\epsilon^{-1/2} \big(|\varphi|_{H^1(B)} + |\varphi|_{H^2(B)}\big), \quad \forall\, \varphi \in H^2(B).$$

Proof. We find from [1, pp. 109–110], for ϵ small, with C independent of ϵ,

$$\sup_{z,z'\in B,\, z'\neq z} \frac{|\varphi(z')-\varphi(z)|}{|z'-z|^{1-\epsilon}} \leqslant C\|\nabla\varphi\|_{L_p(B)}, \quad \text{with } p=2/\epsilon, \quad \forall\,\varphi\in W_p^1(B). \tag{4.15}$$

We shall also apply the Sobolev inequality, with explicit dependence on p,

$$\|\varphi\|_{L_p(B)} \leqslant C\,p^{1/2}\|\varphi\|_{H^1(B)}, \quad \text{for } p<\infty, \quad \forall\,\varphi\in H^1(B). \tag{4.16}$$

For $\varphi\in H_0^1(B)$ a proof was sketched in [12, Lemma 6.4]. For the general case of $\varphi\in H^1(B)$, we make a bounded extension of φ from $H^1(B)$ to $H_0^1(\widetilde{B})$, with $\widetilde{B}\subset\bar{B}$, cf. [1, Chapter IV] and apply (4.16) to $H^1(\widetilde{B})$ to complete the proof.

Employing (4.16) yields

$$\|\nabla\varphi\|_{L_p(B)} \leqslant C\,p^{1/2}\big(|\varphi|_{H^1(B)}+|\varphi|_{H^2(B)}\big), \quad \forall\,\varphi\in H^2(B).$$

Combining this with (4.15), using $p^{1/2}=(2/\epsilon)^{1/2}$, completes the proof. □

Proof of Theorem 4.2. The proof proceeds as that of Theorem 4.1, starting with (4.7) and noting that the bound (4.8) for I remains valid. In order to bound II, we follow the steps above, but now, instead of (4.11), we show

$$|M_hI_h\varphi(\zeta_0)| \leqslant Ch^3\ell_h^{1/2}\|\varphi\|_{H^2(\Pi_{\zeta_0})}. \tag{4.17}$$

Using (4.17) as (4.11) in (4.12), we find

$$\big|[\psi, M_hI_h\varphi]\big| \leqslant Ch^2\ell_h^{1/2}\|\psi\|\,\|\chi\|, \quad \forall\,\psi,\chi\in S_h, \tag{4.18}$$

and hence $|II|\leqslant Ch^2\ell_h^{1/2}\|\psi\|$. Together with (4.8), this completes the proof of (4.13). The error estimate (4.14) now follows from Lemma 3.1 and

$$\|\widetilde{E}_h(t)\widetilde{\Delta}_hQ_hP_hv\| \leqslant Ct^{-1}\|Q_hP_hv\| \leqslant Ch^2\ell_h^{1/2}t^{-1}\|v\|, \quad \text{for } t>0.$$

It remains to show (4.17). Let $\widetilde{\Pi}_{\zeta_0'}$ be the symmetric patch associated with Π_{ζ_0} by the definition of almost symmetric. After a preliminary translation of $\widetilde{\Pi}_{\zeta_0'}$ by $O(h^2)$, we may assume that $\zeta_0'=\zeta_0$. Further, without loss of generality, we may assume that $\widetilde{\Pi}_{\zeta_0}\subset\Pi_{\zeta_0}$. In fact, if this is not the case originally, it will be satisfied by shrinking $\widetilde{\Pi}_{\zeta_0}$ by a suitable factor $1-ch^2$ with $c\geqslant 0$. Starting with $\widetilde{\Pi}_{\zeta_0}$ we may now move the vertices one by one by $O(h^2)$ to obtain Π_{ζ_0} in a finite number of steps, through a sequence of intermediate patches $\widehat{\Pi}_{\zeta_0}\subset\Pi_{\zeta_0}$.

Applying (4.10) we will show that for each of these

$$\big|M_h^{\widehat{\Pi}_{\zeta_0}}I_h\varphi\big| \leqslant C_\epsilon h^{3-\epsilon}\|\varphi\|_{H^2(\Pi_{\zeta_0})}, \quad \text{where } C_\epsilon=C\epsilon^{-1/2},\ \epsilon>0, \tag{4.19}$$

which implies (4.17), by taking $\epsilon=\ell_h^{-1}$ and $\widehat{\Pi}_{\zeta_0}=\Pi_{\zeta_0}$.

Since (4.19) holds for the symmetric patch $\widetilde{\Pi}_{\zeta_0}$, by (4.11), it remains to show that if it holds for a given patch $\widehat{\Pi}_{\zeta_0}$ then it also holds for the next patch in the sequence. Assuming thus that (4.19) holds for $\widehat{\Pi}_{\zeta_0}$, we consider the effect of moving one of its vertices, ζ_2, say, to ζ_2', with $|\zeta_2'-\zeta_2|=O(h^2)$.

Applying Lemma 4.2 to the function $\varphi(h\cdot)$, with B suitable, we obtain

$$\sup_{z,z'\in\Pi_{\zeta_0},\,z'\neq z}\frac{|\varphi(z')-\varphi(z)|}{|z'-z|^{1-\epsilon}}\leqslant C_\epsilon h^{-1+\epsilon}\big(|\varphi|_{H^1(\Pi_{\zeta_0})}+h|\varphi|_{H^2(\Pi_{\zeta_0})}\big)$$
$$\leqslant C_\epsilon h^{-1+\epsilon}\|\varphi\|_{H^2(\Pi_{\zeta_0})}. \tag{4.20}$$

Moving only the vertex ζ_2 in $\widehat{\Pi}_{\zeta_0}$ changes only the triangles τ_1 and τ_2 and thus the terms corresponding to $j=1,2,3$ in (4.10).

Letting τ_1' and τ_2' be the new triangles, the change in the term with $j=1$ is then bounded, since $||\tau_1'|-|\tau_1||\leqslant Ch^3$, by

$$\big|(\omega_1'-\omega_1)\big(\varphi(\zeta_1)-\varphi(\zeta_0)\big)\big|\leqslant C\big||\tau_1'|-|\tau_1|\big|h^{1-\epsilon}\frac{|\varphi(\zeta_1)-\varphi(\zeta_0)|}{|\zeta_1-\zeta_0|^{1-\epsilon}}$$
$$\leqslant C_\epsilon h^3\|\varphi\|_{H^2(\Pi_{\zeta_0})},$$

and thus by the right-hand side of (4.19). The change in the term with $j=3$ is bounded in the same way. For $j=2$ the change is bounded by the modulus of

$$\omega_2'\big(\varphi(\zeta_2')-\varphi(\zeta_0)\big)-\omega_2\big(\varphi(\zeta_2)-\varphi(\zeta_0)\big)=(\omega_2'-\omega_2)\big(\varphi(\zeta_2)-\varphi(\zeta_0)\big)+\omega_2'\big(\varphi(\zeta_2')-\varphi(\zeta_2)\big).$$

The first term on the right is bounded as the terms with $j=1,3$, and the second is bounded, using (4.20), since $|\zeta_2'-\zeta_2|\leqslant Ch^2$, in the following way,

$$\big|\omega_2'\big(\varphi(\zeta_2')-\varphi(\zeta_2)\big)\big|\leqslant C_\epsilon h^2|\zeta_2'-\zeta_2|^{1-\epsilon}h^{-1+\epsilon}\|\varphi\|_{H^2(\Pi_{\zeta_0})}\leqslant C_\epsilon h^{3-\epsilon}\|\varphi\|_{H^2(\Pi_{\zeta_0})}.$$

This shows that (4.19) remains valid after moving ζ_2, which concludes the proof. □

More generally, we shall consider families of *piecewise almost symmetric* triangulations $\{\mathcal{T}_h\}$, in which Ω is partitioned into a fixed set of subdomains $\{\Omega_k\}_{k=1}^K$, and each of these is supplied with an almost symmetric family $\{\mathcal{T}_h(\Omega_k)\}$ so that $\mathcal{T}_h=\bigcup_{k=1}^K\mathcal{T}_h(\Omega_k)$. Such families may be constructed for any convex polygonal domain, cf. Figure 4, by successively refining an initial coarse mesh, a procedure routinely used in computational practice. For such meshes we show the following result.

Theorem 4.3. *If the family* $\{\mathcal{T}_h\}$ *is piecewise almost symmetric, then*

$$\|Q_h\psi\|\leqslant Ch^{3/2}\|\psi\|,\quad\forall\,\psi\in S_h. \tag{4.21}$$

Hence, for the solution of (2.4) *with* $v_h=P_hv$, *we have*

$$\|\tilde u_h(t)-u(t)\|\leqslant Ch^{3/2}t^{-1}\|v\|,\quad \textit{for } t>0. \tag{4.22}$$

Proof. Following again the steps in the proof of Theorem 4.1, we note that (4.8) still holds, and it remains to bound *II*. For each internal vertex ζ_0 of one of the $\mathcal{T}_h(\Omega_k)$, the corresponding patch Π_{ζ_0} is a $O(h^2)$ perturbation of a symmetric patch, and thus (4.17) holds. For $\zeta_0\in Z_h^0$ a vertex on the boundary of two of the $\mathcal{T}_h(\Omega_k)$ we see that by (4.10)

$$|M_h\chi(\zeta_0)|\leqslant Ch^3\max_{x\in\Pi_{\zeta_0}}|\nabla\chi(x)|\leqslant Ch^2\|\nabla\chi\|_{L_2(\Pi_{\zeta_0})},$$

and by the use of approximation properties of the interpolation operator I_h we get

$$|M_h I_h \varphi(\zeta_0)| \leqslant Ch^2 \|I_h\varphi\|_{H^1(\Pi_{\zeta_0})} \leqslant Ch^2 \big(\|\varphi\|_{H^1(\Pi_{\zeta_0})} + h|\varphi|_{H^2(\Pi_{\zeta_0})}\big). \tag{4.23}$$

Using (4.17) and (4.23) as earlier (4.11) in (4.12), we conclude

$$\big|[\psi, M_h I_h \varphi]\big| \leqslant Ch^2 \ell_h^{1/2} \|\psi\| \, |\varphi|_2 + Ch\|\psi\| \, \|\varphi\|_{H^1(\Omega_S)},$$

where Ω_S is a strip of width $O(h)$ around the interface between the subdomains Ω_k of Ω. Using now the inequality $\|\varphi\|_{H^1(\Omega_S)} \leqslant Ch^{1/2}\|\varphi\|_{H^2(\Omega)} \leqslant Ch^{1/2}\|\chi\|$, we get

$$\big|[\psi, M_h I_h \varphi]\big| \leqslant Ch^{3/2} \|\psi\| \, \|\chi\|, \quad \forall\, \psi, \chi \in S_h, \tag{4.24}$$

and hence $|II| \leqslant Ch^{3/2}\|\psi\|$. Together with (4.8), this completes the proof of (4.21). The error estimate (4.22) now follows by Lemma 3.1 and

$$\|\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h P_h v\| \leqslant Ct^{-1}\|Q_h P_h v\| \leqslant Ch^{3/2} t^{-1} \|v\|, \quad \text{for } t > 0. \quad \square$$

We remark that the operator M_h used here, modulo a constant factor, is the same as the operator Δ_h^* in [4]. The arguments in the proofs of Theorems 4.2 and 4.3 therefore show that the following result holds for the lumped mass method.

Corollary 4.1. *Assume that $\{\mathcal{T}_h\}$ is almost or piecewise almost symmetric. Then the nonsmooth data error estimates for the lumped mass method, corresponding to* (4.13) *and* (4.21)*, respectively, hold.*

We finish this section by remarking that, in one space dimension, the full $O(h^2)$ L_2-norm bound (1.17) for Q_h holds also for almost symmetric partitions, without a logarithmic factor. Let $\Omega = (0, 1)$ be partitioned by $0 = x_0 < x_1 < \cdots < x_{N_h+1} = 1$. Denote now $\mathcal{T}_h = \{\tau_i\}_{i=1}^{N_h+1}$, with $\tau_i = [x_{i-1}, x_i]$, and let S_h be the set of the continuous piecewise linear functions over $\mathcal{T}_h$, vanishing at $x = 0, 1$. We set $h_i = x_i - x_{i-1}$ and $h = \max_i h_i$. The control volumes are $V_i = (x_i - h_i/2, x_i + h_{i+1}/2)$ and $J_h\psi(x) = \psi(x_i)$ for $x \in V_i$. We say that $\mathcal{T}_h$ is almost symmetric if $|h_{i+1} - h_i| \leqslant Ch^2$ for all i.

Simple calculations show, with $(\chi, \psi) = \int_0^1 \chi\psi \, dx$ and $\langle \chi, \psi \rangle = (\chi, J_h\psi)$, for $\chi, \psi \in S_h$,

$$\varepsilon_h(\psi, \chi) = \langle \psi, \chi \rangle - (\psi, \chi) = -\frac{1}{24} \sum_{i=1}^{N_h} \psi_i \big(h_{i+1}(\chi_{i+1} - \chi_i) - h_i(\chi_i - \chi_{i-1})\big),$$

where $w_i = w(x_i)$ for a function w on Ω, and the one-dimensional version of (4.10) at x_i becomes

$$M_h I_h \varphi(x_i) = -\frac{1}{24}\big(h_{i+1}(\varphi_{i+1} - \varphi_i) + h_i(\varphi_{i-1} - \varphi_i)\big), \quad i = 1, \ldots, N_h.$$

The crucial step to prove (1.17) is then to show an analogue of (4.11), in this case

$$|M_h I_h \varphi(x_i)| \leqslant Ch^{5/2} \|\varphi\|_{H^2(\Pi_{x_i})}, \quad i = 1, \ldots, N_h, \quad \text{with } \Pi_{x_i} = \tau_i \cup \tau_{i+1}, \tag{4.25}$$

from which (1.17) follows as earlier. Using the Taylor formula

$$\varphi(x) = \varphi(x_i) + (x - x_i)\varphi'(x_i) + \int_{x_i}^{x} (x - y)\, \varphi''(y) \, dy,$$

we find easily

$$M_h I_h \varphi(x_i) = -\frac{1}{24}(h_{i+1}^2 - h_i^2)\varphi'(x_i) + O\big(h^{5/2}\,\|\varphi''\|_{L_2(\Pi_{x_i})}\big), \quad i = 1, \dots, N_h.$$

By the almost symmetry, $|h_{i+1}^2 - h_i^2| \leqslant Ch^3$ and by the Sobolev type inequality

$$|\varphi'(x_i)| \leqslant Ch^{-1/2}\big(\|\varphi'\|_{L_2(\Pi_{x_i})} + h\|\varphi''\|_{L_2(\Pi_{x_i})}\big) \leqslant Ch^{-1/2}\|\varphi\|_{H^2(\Pi_{x_i})},$$

for $i = 1, \dots, N_h$, we now conclude that (4.25) holds.

5. Examples of Nonoptimal Nonsmooth Initial Data Estimates

In this section we present two examples where the necessary and sufficiency condition (3.6) for an optimal $O(h^2)$ nonsmooth data error estimate for $t > 0$ is not satisfied. In the first example we construct a family of nonsymmetric meshes $\{\mathcal{T}_h\}$ for which the norm on the left-hand side of (3.6) is bounded below by ch, thus showing that the first order error bound of Theorem 3.5 is the best possible. In the second example we exhibit a piecewise symmetric mesh for which this norm is bounded below by $ch^{3/2}$, implying that the error estimate of Theorem 4.3 is best possible.

In our first example we choose $\Omega = (0,1) \times (0,1)$ and introduce a quasi-uniform family of triangulations $\{\mathcal{T}_h\}$ of Ω as follows. Let N be a positive integer divisible by 4, $h = 4/(3N)$, $x_0 = 0$, and set, for $j = 1, \dots, N$ and $m = 0, 1, \dots, M = \frac{3}{4}N$,

$$x_j = x_{j-1} + \begin{cases} \frac{1}{2}h, & \text{for } j \text{ odd}, \\ h, & \text{for } j \text{ even}, \end{cases} \quad \text{and} \quad y_m = mh. \tag{5.1}$$

We split the rectangle $(x_j, x_{j+1}) \times (y_m, y_{m+1})$ into two triangles by connecting the nodes (x_j, y_m) and (x_{j+1}, y_{m-1}), see Figure 5. This defines a triangulation $\mathcal{T}_h$ that is not symmetric at any vertex.

Let now $\zeta_0 = (x_{2j}, y_m)$, $\zeta_0 \in Z_h^0$, and let Π_{ζ_0} be the corresponding nonsymmetric patch shown in Figure 5, with vertices $\{\zeta_j\}_{j=1}^6$. Let τ_j be the triangle in Π_{ζ_0} with vertices ζ_0, ζ_j, ζ_{j+1}, where $\zeta_7 = \zeta_1$. We then have $|\tau_j| = \frac{1}{4}h^2$, for $j = 1, 2, 3$, and $|\tau_j| = \frac{1}{2}h^2$, for $j = 4, 5, 6$. Thus, using (4.10), for $\psi \in S_h$, we obtain with $\psi_j = \psi(\zeta_j)$,

$$\begin{aligned} M_h\psi(\zeta_0) &= -\frac{1}{54}\sum_{j=1}^{6}\omega_j(\psi_j - \psi_0) = -\frac{1}{54}\frac{h^2}{4}\Big(3(\psi_1 + \psi_4 - 2\psi_0) \\ &\quad + 2(\psi_2 - \psi_0) + 2(\psi_3 - \psi_0) + 4(\psi_5 - \psi_0) + 4(\psi_6 - \psi_0)\Big). \end{aligned} \tag{5.2}$$

Because $\nabla\psi$ is piecewise constant over Π_{ζ_0}, we easily see that (5.2) implies

$$|M_h\psi(\zeta_0)| \leqslant Ch^2\|\nabla\psi\|_{L_2(\Pi_{\zeta_0})}, \quad \forall\, \psi \in S_h. \tag{5.3}$$

For a smooth function φ we have, by Taylor expansion,

$$\varphi(\zeta_j) - \varphi(\zeta_0) = \nabla\varphi(\zeta_0) \cdot (\zeta_j - \zeta_0) + O(h^2),$$

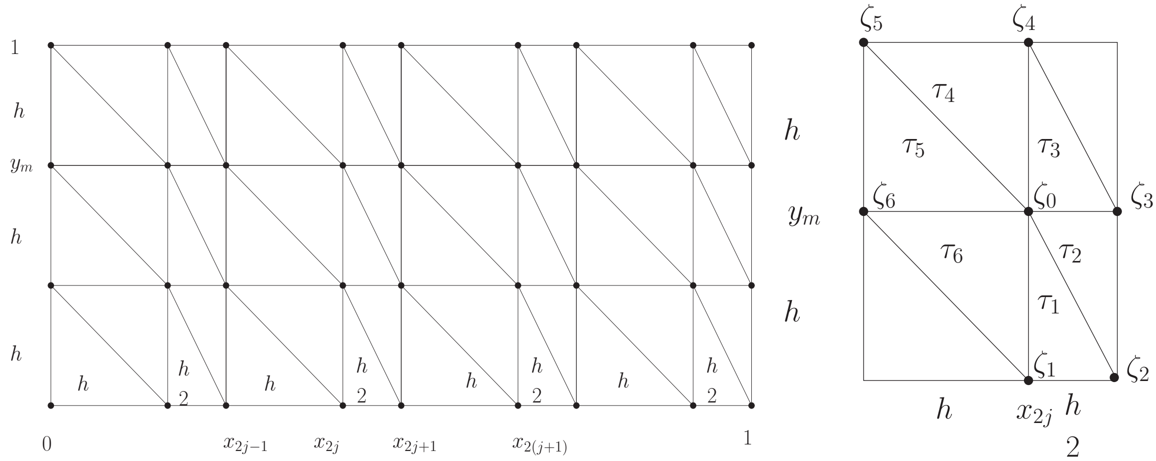

Figure 5. *Left:* A nonsymmetric mesh. *Right:* A nonsymmetric patch Π_{ζ_0}, around ζ_0.

where ζ_j is considered as a vector with components its Cartesian coordinates and the dot denotes the Euclidean inner product in $\mathbb{R}^2$. Employing this in (5.2), we find, after a simple calculation,

$$M_h I_h \varphi(\zeta_0) = \frac{h^3}{108} \nabla\varphi(\zeta_0) \cdot (3, -1) + O(h^4). \tag{5.4}$$

Let $\phi_1(x, y) = 2\sin(\pi x)\sin(\pi y)$ be the eigenfunction of $-\Delta$, corresponding to the smallest eigenvalue $\lambda_1 = 2\pi^2$. We then easily find that $\nabla\phi_1(1/4, 1/4) \cdot (3, -1) = 2\pi$. Hence, there exists a square $\mathcal{P} = [1/4 - d, 1/4 + d]^2$, with $0 < d < 1/4$, such that

$$\nabla\phi_1(z) \cdot (3, -1) \geqslant 1, \quad \forall\, z \in \mathcal{P}. \tag{5.5}$$

Letting now $z \in Z_h^0 \cap \mathcal{P}$ we then have that $M_h I_h \phi_1(z) \geqslant ch^3$, $c > 0$, for h small. We shall prove the following proposition.

Proposition 5.1. *Let* $\mathcal{T}_h$ *be defined by* (5.1), $\mathcal{P}_h = \{z = (x_{2j}, y_m) \in \mathcal{P}\}$ *and consider the initial value problem* (2.4) *with* $v_h = \sum_{z\in\mathcal{P}_h} \Phi_z$, *where* $\Phi_z \in S_h$ *is the nodal basis function of* S_h *at* z. *Then we have, for* h *small,*

$$\|\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h v_h\| \geqslant c(t)h\|v_h\|, \quad \text{with } c(t) > 0, \quad \text{for } t > 0.$$

Proof. Letting $\tilde{\lambda}_j^h$ and $\tilde{\phi}_j^h$ be the eigenvalues and eigenfunctions of $-\widetilde{\Delta}_h$, and using Parseval's relation in S_h, equipped with $\langle\cdot,\cdot\rangle$, we have

$$|||\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h v_h|||^2 = \sum_{j=1}^{N_h} e^{-2t\tilde{\lambda}_j^h} \langle\widetilde{\Delta}_h Q_h v_h, \tilde{\phi}_j^h\rangle^2 \geqslant e^{-2t\tilde{\lambda}_1^h} \langle\widetilde{\Delta}_h Q_h v_h, \tilde{\phi}_1^h\rangle^2. \tag{5.6}$$

Combining (2.3), (1.15) and (4.1), we find

$$-\langle\widetilde{\Delta}_h Q_h v_h, \psi\rangle = (\nabla Q_h v_h, \nabla\psi) = \varepsilon_h(v_h, \psi) = [v_h, M_h\psi], \quad \forall\, \psi \in S_h. \tag{5.7}$$

Note now that for $z \in \mathcal{P}_h$, the corresponding patch Π_z has the same form as the patch Π_{ζ_0} considered above. Thus employing (5.3) for $\zeta_0 = z$, we get, for $\psi \in S_h$,

$$\big|[v_h, M_h\psi]\big| \leqslant \sum_{z\in\mathcal{P}_h} \big|[\Phi_z, M_h\psi]\big| = \sum_{z\in\mathcal{P}_h} |M_h\psi(z)| \leqslant Ch\|\nabla\psi\|, \tag{5.8}$$

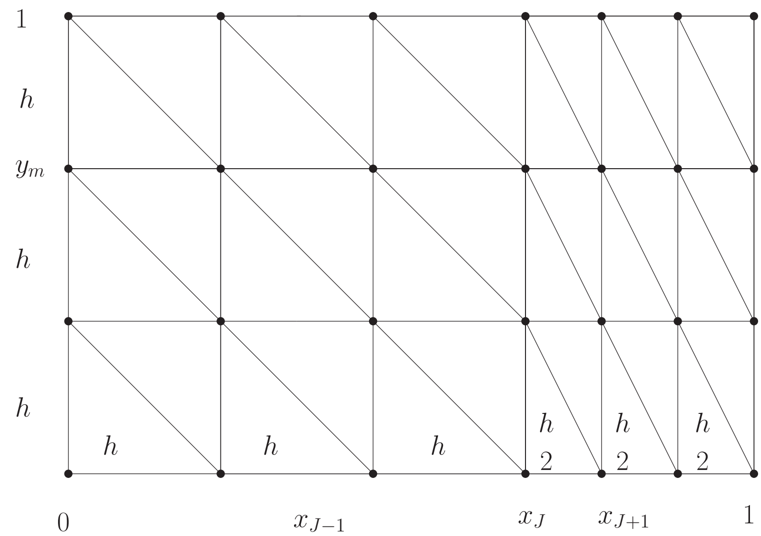

Figure 6. A piecewise symmetric mesh.

where in the last inequality we have used the fact that the number of points in $\mathcal{P}_h$ is $O(N^2) = O(h^{-2})$. We recall from [8] that

$$\|\tilde{\phi}_1^h - \phi_1\|_{H^1} = O(h) \quad \text{and} \quad \tilde{\lambda}_1^h \to \lambda_1, \text{ as } h \to 0,$$

and, since obviously $\|\phi_1 - I_h\phi_1\|_{H^1} = O(h)$, (5.8) with $\psi = \tilde{\phi}_1^h - I_h\phi_1$ gives

$$\big|[v_h, M_h(\tilde{\phi}_1^h - I_h\phi_1)]\big| \leqslant Ch\|\nabla(\tilde{\phi}_1^h - I_h\phi_1)\| \leqslant Ch^2. \tag{5.9}$$

For every $z \in \mathcal{P}_h$, (5.5) holds, and thus, using (5.4) with $\varphi = \phi_1$ and $\zeta_0 = z$, we obtain, for h small, since the number of vertices in $\mathcal{P}_h$ is bounded below by cN^2,

$$[v_h, M_h I_h\phi_1] = \sum_{z\in\mathcal{P}_h} M_h I_h \phi_1(z) \geqslant ch^3N^2 = ch, \quad \text{with } c > 0.$$

Combining this with (5.9), we obtain, for h small,

$$[v_h, M_h\tilde{\phi}_1^h] \geqslant [v_h, M_h I_h\phi_1] - \big|[v_h, M_h(\tilde{\phi}_1^h - I_h\phi_1)]\big| \geqslant ch - Ch^2 \geqslant ch, \quad \text{with } c > 0.$$

Since $|||v_h||| = O(1)$, (5.6) and (5.7) now show

$$|||\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h v_h||| \geqslant e^{-t\tilde{\lambda}_1^h}[v_h, M_h\tilde{\phi}_1^h] \geqslant c(t)h\,|||v_h|||, \quad \text{for } t > 0.$$

Since $|||\cdot|||$ and $\|\cdot\|$ are equivalent norms, the proof is complete. □

It follows from Proposition 5.1 and Lemma 3.1 that the highest order of convergence that can hold, uniformly for all $v \in L_2$, and for any family of triangulations $\{\mathcal{T}_h\}$, is $O(h)$, i.e., Theorem 3.4 is best possible, in this case.

We now turn to our second example, in which $\{\mathcal{T}_h\}$ is a piecewise symmetric family. Let again $\Omega = (0,1)\times(0,1)$ and consider a triangulation $\mathcal{T}_h$ of Ω, where the nodes (x_j, y_m) are given as follows. With J a positive integer, let $N = 7J$, $M = 4J$ and $h = 1/(4J)$, and set for $j = 0,\dots,N$ and $m = 0,\dots,M$,

$$x_j = \begin{cases} jh, & \text{for } 0 \leqslant j \leqslant J, \\ 1/4 + (j-J)h/2, & \text{for } J < j \leqslant N, \end{cases} \quad \text{and} \quad y_m = mh, \tag{5.10}$$

see Figure 6. This time we consider the set of vertices in $\mathcal{P}$ with $x = 1/4$ and prove the following proposition.

Proposition 5.2. *Let $\mathcal{T}_h$ be defined by (5.10) and $\mathcal{P}'_h = \{z = (x_J, y_m) \in \mathcal{P}\}$. For the initial value problem (2.4), with $v_h = \sum_{z\in\mathcal{P}'_h} \Phi_z$, where $\Phi_z \in S_h$ is the nodal basis function of S_h at z, we have, for h small,*

$$\|\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h v_h\| \geqslant c(t)h^{3/2}\|v_h\|, \quad \text{with } c(t) > 0, \quad \text{for } t > 0.$$

Proof. Again, using (5.6) and (5.7), we have,

$$|||\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h v_h|||^2 \geqslant e^{-2t\tilde{\lambda}_1^h}[v_h, M_h\tilde{\phi}_1^h]^2. \tag{5.11}$$

For $z \in \mathcal{P}'_h$, the corresponding patch Π_z has the same form as the patch Π_{ζ_0} considered above, see Figure 5 (right). Thus employing (5.3) for $\zeta_0 = z$ and taking into account that the number of vertices in $\mathcal{P}'_h$ is $O(N)$ we now obtain, for $\psi \in S_h$,

$$\big|[v_h, M_h\psi]\big| \leqslant \sum_{z\in\mathcal{P}'_h}\big|[\Phi_z, M_h\psi]\big| = \sum_{z\in\mathcal{P}'_h}|M_h\psi(z)| \leqslant Ch^{3/2}\|\nabla\psi\|.$$

Similarly to (5.9) this now shows

$$\big|[v_h, M_h(\tilde{\phi}_1^h - I_h\phi_1)]\big| \leqslant Ch^{5/2}, \tag{5.12}$$

and, again using (5.4), for h small,

$$[v_h, M_h I_h\phi_1] = \sum_{z\in\mathcal{P}'_h} M_h I_h\phi_1(z) \geqslant ch^3 J = ch^2, \quad \text{with } c > 0.$$

Combined with (5.12) this gives, for h small,

$$[v_h, M_h\tilde{\phi}_1^h] \geqslant ch^2 - Ch^{5/2} \geqslant ch^2, \quad \text{with } c > 0. \tag{5.13}$$

Since $|||v_h||| = O(h^{1/2})$ we obtain from (5.11) and (5.13)

$$|||\widetilde{E}_h(t)\widetilde{\Delta}_h Q_h v_h||| \geqslant c(t)h^2 \geqslant c(t)h^{3/2}\,|||v_h|||, \quad \text{for } t > 0. \quad \square$$

It follows from Proposition 5.2 and Lemma 3.1 that the highest order of convergence that can hold, uniformly for all $v \in L_2$, and for all piecewise symmetric families $\{\mathcal{T}_h\}$, is $O(h^{3/2})$, i.e., Theorem 4.3 is best possible in this regard.

Remark 5.1. Since M_h is proportional to the operator Δ_h^* used in [4], the arguments in this section also apply to the lumped mass method. In particular, the analogue of Proposition 5.1 then shows that the first order nonsmooth data estimate for $t > 0$ of [4, Theorem 4.3] is best possible for general triangulations $\{T_h\}$. Further, the $O(h^{3/2})$ estimate stated in Corollary 4.1 is best possible for piecewise almost symmetric triangulations. Our examples here may be thought of as generalizations to two space dimensions of the one-dimensional counter-examples in [4, Section 7].

6. Some Fully Discrete Schemes

In this section we discuss briefly the generalization of our above results for the spatially semidiscrete finite volume method to some basic fully discrete schemes, namely the backward Euler and Crank–Nicolson methods.

With $k > 0$, $t_n = n\,k$, $n = 0, 1, \ldots$, the backward Euler finite volume method approximates $u(t_n)$ by $\widetilde{U}^n \in S_h$ for $n \geqslant 0$ such that, with $\bar\partial \widetilde{U}^n = (\widetilde{U}^n - \widetilde{U}^{n-1})/k$,

$$\langle \bar\partial \widetilde{U}^n, \chi\rangle + (\nabla \widetilde{U}^n, \nabla\chi) = 0, \quad \forall\, \chi \in S_h, \quad \text{for } n \geqslant 1, \quad \text{with } \widetilde{U}^0 = v_h,$$

or,

$$\bar\partial \widetilde{U}^n - \widetilde{\Delta}_h \widetilde{U}^n = 0, \quad \text{for } n \geqslant 1, \quad \text{with } \widetilde{U}^0 = v_h. \tag{6.1}$$

Introducing the discrete solution operator $\widetilde{E}_{kh} = (I - k\widetilde{\Delta}_h)^{-1}$ we may write $\widetilde{U}^n = \widetilde{E}_{kh}\widetilde{U}^{n-1} = \widetilde{E}_{kh}^n \widetilde{U}^0$, $n \geqslant 1$. Using eigenfunction expansion and Parseval's relation, we obtain, analogously to [12, Chapter 7], the stability property

$$\|\nabla^p \widetilde{E}_{kh}^n \chi\| \leqslant C\|\nabla^p \chi\|, \quad \forall\, \chi \in S_h, \quad \text{for } p = 0, 1. \tag{6.2}$$

The estimates that follow and their proofs are analogous to those for the lumped mass method derived in [4], since the operators $\widetilde{E}_h(t)$, $\widetilde{\Delta}_h$ and Q_h, defined in Section 2, have properties analogous to those of the corresponding operators for the lumped mass method. For simplicity we will only sketch the proof of Theorem 6.1.

We shall use the following abstract lemma shown in [4], in the case $\mathcal{H} = S_h$, normed by $|||\cdot|||$, and with $\mathbf{A} = -\widetilde{\Delta}_h$.

Lemma 6.1. *Let* $\mathbf{A}$ *be a linear, selfadjoint, positive definite operator in a Hilbert space* $\mathcal{H}$, *with compact inverse, let* $\mathbf{u} = \mathbf{u}(t)$ *be the solution of*

$$\mathbf{u}' + \mathbf{A}\mathbf{u} = 0, \quad \textit{for } t > 0, \quad \textit{with } \mathbf{u}(0) = \mathbf{v},$$

and let $\mathbf{U} = \{\mathbf{U}^n\}_{n=0}^{\infty}$ *be defined by*

$$\bar\partial \mathbf{U}^n + \mathbf{A}\mathbf{U}^n = 0, \quad \textit{for } n \geqslant 1, \quad \textit{with } \mathbf{U}^0 = \mathbf{v}.$$

Then, for $p = 0, 1$, $-1 \leqslant q \leqslant 3$, *with* $p + q \geqslant 0$, *we have*

$$\|\mathbf{A}^{p/2}(\mathbf{U}^n - \mathbf{u}(t_n))\| \leqslant Ckt_n^{-(1-q/2)}\|\mathbf{A}^{(p+q)/2}\mathbf{v}\|, \quad \textit{for } n \geqslant 1.$$

The error estimates of the following theorem for (6.1) are of optimal order under the same assumptions as in Section 3.

Theorem 6.1. *Let* u *and* $\widetilde{U}$ *be the solutions of* (1.1) *and* (6.1). *Then, for* $n \geqslant 1$,

$$\|\widetilde{U}^n - u(t_n)\| \leqslant \begin{cases} C(h^2 + k)|v|_2, & \textit{if } \|v_h - v\| \leqslant Ch^2|v|_2, \\ C(h^2 + k)t_n^{-1/2}|v|_1, & \textit{if } v_h = P_h v \textit{ and } \|\nabla P_h v\| \leqslant C|v|_1, \\ C(h^2 + k)t_n^{-1}\|v\|, & \textit{if } v_h = P_h v \textit{ and } (1.17) \textit{ holds.} \end{cases}$$

Proof. Analogously to the proof of [4, Theorem 8.1], we split the error as

$$\widetilde{U}^n - u(t_n) = (\widetilde{U}^n - \tilde{u}_h(t_n)) + (\tilde{u}_h(t_n) - u(t_n)) = \beta_n + \eta_n.$$

By Theorems 3.1–3.3, η_n is bounded as required. In order to bound $\beta_n = (\widetilde{E}^n_{kh} - \widetilde{E}_h(t_n))v_h$ in the smooth data case, it suffices, using the stability estimates (6.2) and Lemma 2.1, to consider $v_h = R_h v$. We obtain by Lemma 6.1, with $\mathbf{A} = A_h = -\widetilde{\Delta}_h$, and $q = 2, 1, 0$,

$$|||\beta_n||| = |||\widetilde{U}^n - \tilde{u}_h(t_n)||| \leqslant Ckt_n^{-(1-q/2)}|||A_h^{q/2}v_h||| \leqslant Ckt_n^{-(1-q/2)}|v|_q,$$

where for $q = 2$, the last inequality follows from

$$|||A_h R_h v|||^2 = (\nabla R_h v, \nabla A_h R_h v) = (\nabla v, \nabla A_h R_h v) = -(\Delta v, A_h R_h v),$$

for $q = 1$ from $|||A_h^{1/2} P_h v||| = \|\nabla P_h v\| \leqslant C|v|_1$ and for $q = 0$ from $|||P_h v||| \leqslant C\|v\|$. □

Also for the lumped mass method the analogous result in the mildly nonsmooth data case $v \in \dot{H}^1$ holds, and should replace the result for $q = 1$ in [4, Theorem 8.1], cf. the remark after Theorem 3.2.

Recall that Q_h satisfies (1.17) if $\{\mathcal{T}_h\}$ is symmetric. For almost symmetric or piecewise almost symmetric $\{\mathcal{T}_h\}$ we obtain correspondingly the following nonsmooth initial data error estimates employing (4.14) and (4.22).

Theorem 6.2. *Let u and $\widetilde{U}$ be the solutions of* (1.1) *and* (6.1)*, with $v_h = P_h v$. Then, for $n \geqslant 1$,*

$$\|\widetilde{U}^n - u(t_n)\| \leqslant \begin{cases} C(h^2\ell_h^{1/2} + k)t_n^{-1}\|v\|, & \text{if } \{\mathcal{T}_h\} \text{ is almost symmetric,} \\ C(h^{3/2} + k)t_n^{-1}\|v\|, & \text{if } \{\mathcal{T}_h\} \text{ is piecewise almost symmetric.} \end{cases}$$

For the gradient of the error we may prove as in [4, Theorem 8.2], the following smooth and nonsmooth data error estimates, without additional assumptions on $\mathcal{T}_h$. For smooth initial data we assumed in [4] that $v_h = R_h v$, but the more general choices of v_h are permitted by the stability estimates (6.2) and Lemma 2.1.

Theorem 6.3. *Let u and $\widetilde{U}$ be the solutions of* (1.1) *and* (6.1)*. Then, for $n \geqslant 1$,*

$$\|\nabla(\widetilde{U}^n - u(t_n))\| \leqslant \begin{cases} C(h + k)|v|_3, & \text{if } \|\nabla(v_h - v)\| \leqslant Ch|v|_2, \\ C(h\,t_n^{-1} + k\,t_n^{-3/2})\|v\|, & \text{if } v_h = P_h v. \end{cases}$$

We now turn to the Crank–Nicolson method, defined by

$$\bar{\partial}\widetilde{U}^n - \widetilde{\Delta}_h\widetilde{U}^{n-\frac{1}{2}} = 0, \quad \text{for } n \geqslant 1, \quad \text{with } U^0 = v_h, \quad \widetilde{U}^{n-\frac{1}{2}} = \frac{1}{2}(\widetilde{U}^n + \widetilde{U}^{n-1}). \tag{6.3}$$

Denoting again the discrete solution operator by $\widetilde{E}_{kh} = (I + \frac{1}{2}k\widetilde{\Delta}_h)(I - \frac{1}{2}k\widetilde{\Delta}_h)^{-1}$ we may write $\widetilde{U}^n = \widetilde{E}_{kh}\widetilde{U}^{n-1} = \widetilde{E}^n_{kh}\widetilde{U}^0$, $n \geqslant 1$. Using eigenfunction expansion and Parseval's relation, we find that (6.2) also holds for this method.

The Crank–Nicolson method does not have as advantageous smoothing properties as the backward Euler method, which is reflected in the fact that the following analogue of Lemma 6.1, shown in [4, Lemma 8.2], does not allow $q = 0$.

Lemma 6.2. *Let* $\mathbf{A}$ *and* $\mathbf{u}(t)$ *be as in Lemma 6.1 and let* $\mathbf{U}^n$ *satisfy*

$$\bar\partial \mathbf{U}^n + \mathbf{A}\mathbf{U}^{n-\frac12} = 0, \quad \text{for } n \geqslant 1, \quad \text{with } \mathbf{U}^0 = \mathbf{v}.$$

Then

$$\|\mathbf{A}^{p/2}(\mathbf{U}^n - \mathbf{u}(t_n))\| \leqslant Ck^2 t_n^{-(2-q)} \|\mathbf{A}^{p/2+q}\mathbf{v}\|, \quad \text{for } n \geqslant 1,\ p = 0,1,\ q = 1,2.$$

This time optimal order estimates for the error in L_2 and in H^1 hold uniformly down to $t = 0$, if $v \in \dot H^4$ and $v \in \dot H^5$, respectively. The proofs are analogous to those of [4, Theorems 8.3 and 8.4], where we assumed $v_h = R_h v$. Again the stability estimates (6.2) and Lemma 2.1 permit the more general choices for v_h.

Theorem 6.4. *Let* u *and* $\widetilde U$ *be the solutions of* (1.1) *and* (6.3)*. Then, with* $q = 1, 2$*, we have, for* $n \geqslant 1$*,*

$$\|\widetilde U^n - u(t_n)\| \leqslant C(h^2 + k^2 t_n^{-(2-q)})|v|_{2q}, \qquad \text{if } \|v_h - v\| \leqslant Ch^2|v|_2,$$
$$\|\nabla(\widetilde U^n - u(t_n))\| \leqslant C(h + k^2 t_n^{-(2-q)})|v|_{2q+1}, \quad \text{if } \|\nabla(v_h - v)\| \leqslant Ch|v|_2.$$

For optimal order convergence for initial data only in L_2, one may modify the Crank–Nicolson scheme by taking the first two steps by the backward Euler method, which has a smoothing effect. We may show then the following result, analogously to that of [4, Theorem 8.5], with the obvious modifications for almost symmetric and piecewise almost symmetric families $\{\mathcal{T}_h\}$.

Theorem 6.5. *Let* u *be the solution of* (1.1) *and* $\widetilde U^n$ *that of* (6.1)*, for* $n = 1, 2$*, and of* (6.3)*, for* $n \geqslant 3$*, with* $v_h = P_h v$ *and assume* (1.17) *holds. Then we have*

$$\|\widetilde U^n - u(t_n)\| \leqslant C(h^2 t_n^{-1} + k^2 t_n^{-2})\|v\|, \quad \text{for } n \geqslant 1.$$

7. Problems with More General Elliptic Operators

This final section is devoted to the extension of our earlier results to the more general problem (1.19), and we recall that we shall consider the finite volume method (1.26) where the bilinear form $\tilde a_h(\cdot,\cdot)$ is defined by (1.25). Our error analysis is again based on estimates for the standard Galerkin finite element method, in this case defined by (1.20) and (1.21). It is well known that for this method the stability and smoothing estimates (2.2) hold as do the error estimates (1.5)–(1.7), where the norms $|\cdot|_q$ are defined analogously to the norms (1.2), using the eigenvalues and eigenfunctions of A.

We introduce the discrete elliptic operator $\widetilde A_h : S_h \to S_h$ by

$$\langle \widetilde A_h \psi, \chi\rangle = \tilde a_h(\psi, J_h\chi), \quad \forall\, \chi, \psi \in S_h, \tag{7.1}$$

which is symmetric and positive definite with respect to the inner product $\langle\cdot,\cdot\rangle$ by (1.24), since $(\tilde\beta\psi, J_h\chi)$ is symmetric, positive semidefinite on S_h. This follows from the fact that $\int_\tau \chi J_h\psi\,dx$ is symmetric by (4.6) and $\tilde\beta$ is constant and non-negative in each τ of $\mathcal{T}_h$. We may then rewrite (1.26) as

$$\tilde u_{h,t} + \widetilde A_h \tilde u_h = 0, \quad \text{for } t \geqslant 0, \quad \text{with } \tilde u_h(0) = v_h, \tag{7.2}$$

and the solution is given by $\tilde{u}_h(t) = \widetilde{E}_h(t)v_h$, where $\widetilde{E}_h(t) = e^{-\widetilde{A}_h t}$ is defined as in (2.5), with $\{\tilde{\lambda}_j^h\}$ and $\{\tilde{\phi}_j^h\}$ the eigenvalues and eigenfunctions of $\widetilde{A}_h$, orthonormal with respect to $\langle \cdot, \cdot \rangle$.

Note that a slightly different finite volume element method for (1.19) has been considered in [9]. This method differs in the discretization of the lower order term, using the bilinear $\bar{a}_h(\cdot, \cdot)$ defined by

$$\bar{a}_h(\psi, J_h\chi) = (\tilde{\alpha}\nabla\psi, \nabla\chi) + (\beta J_h\psi, J_h\chi), \quad \forall\, \psi, \chi \in S_h.$$

For this method analogous results to Theorems 7.1–7.3 hold.

Following our error analysis in the previous sections we introduce $\delta = \tilde{u}_h - u_h$ and split the error into $\tilde{u}_h - u = \delta + (u_h - u)$, where $u_h - u$ and $\nabla(u_h - u)$ are estimated by the analogues of (1.5)–(1.7). It therefore suffices to derive estimates for δ, which satisfies, for $t \geqslant 0$,

$$\langle \delta_{h,t}, \chi \rangle + \tilde{a}_h(\delta, J_h\chi) = -\varepsilon_h(u_{h,t}, \chi) - \tilde{\varepsilon}_h(u_h, \chi), \quad \forall\, \chi \in S_h, \quad \text{with } \delta(0) = 0, \tag{7.3}$$

where $\varepsilon_h(\cdot, \cdot)$ is given by (1.16) and $\tilde{\varepsilon}_h(\cdot, \cdot)$ is defined by

$$\tilde{\varepsilon}_h(\psi, \chi) = \tilde{a}_h(\psi, J_h\chi) - a(\psi, \chi), \quad \forall\, \psi, \chi \in S_h. \tag{7.4}$$

Now let $Q_h : S_h \to S_h$ and $\widetilde{Q}_h : S_h \to S_h$ be the quadrature error operators given by

$$\tilde{a}_h(Q_h\psi, J_h\chi) = \varepsilon_h(\psi, \chi) \quad \text{and} \quad \tilde{a}_h(\widetilde{Q}_h\psi, J_h\chi) = \tilde{\varepsilon}_h(\psi, \chi), \quad \forall\, \psi, \chi \in S_h. \tag{7.5}$$

Using (7.1), the equation (7.3) for δ can then be written in operator form as

$$\delta_t + \widetilde{A}_h\delta = -\widetilde{A}_h Q_h u_{h,t} - \widetilde{A}_h \widetilde{Q}_h u_h, \quad \text{for } t \geqslant 0, \quad \text{with } \delta(0) = 0.$$

This problem is similar to (3.1), except that the operator $-\widetilde{\Delta}_h$ is replaced by $\widetilde{A}_h$ and that on the right-hand side we have an additional term resulting from the approximation of the bilinear form $a(\cdot, \cdot)$. By Duhamel's principle we have

$$\begin{aligned} \delta(t) &= -\int_0^t \widetilde{E}_h(t-s)\widetilde{A}_h Q_h u_{h,t}(s)\, ds - \int_0^t \widetilde{E}_h(t-s)\widetilde{A}_h \widetilde{Q}_h u_h(s)\, ds \\ &=: \tilde{\delta}(t) + \widehat{\delta}(t), \quad \text{for } t \geqslant 0. \end{aligned} \tag{7.6}$$

To estimate δ it therefore suffices to bound $\tilde{\delta}$ and $\widehat{\delta}$. For this we need some auxiliary results, which are discussed below.

Lemma 7.1. *Let $\alpha, \beta \in \mathcal{C}^2$. For the error functional $\tilde{\varepsilon}_h$, defined by (7.4), we have*

$$|\tilde{\varepsilon}_h(\psi, \chi)| \leqslant Ch^{p+q}\|\nabla^q\psi\|\,\|\nabla^p\chi\|, \quad \forall\, \psi, \chi \in S_h, \quad \textit{with } p, q = 0, 1.$$

Proof. In view of (7.4), we may write

$$\tilde{\varepsilon}_h(\psi, \chi) = ((\tilde{\alpha} - \alpha)\nabla\psi, \nabla\chi) + (\tilde{\beta}\psi, J_h\chi) - (\beta\psi, \chi).$$

We then split $\tilde{\varepsilon}_h(\psi, \chi)$ as a sum of integrals over $\tau \in \mathcal{T}_h$. Since $\tilde{\alpha} = \alpha(z_\tau)$, we see that $\int_\tau (f - f(z_\tau))dx = 0$ for linear functions f, and hence

$$\Big|\int_\tau (f - f(z_\tau))dx\Big| \leqslant Ch_\tau^2|\tau|\|f\|_{\mathcal{C}^2}, \quad \text{for } f \in \mathcal{C}^2, \tag{7.7}$$

with h_τ the maximal side length of τ. Therefore, using this and the fact that $\nabla\psi\cdot\nabla\chi$ is constant in τ, we get

$$\Big|\int_\tau(\tilde\alpha-\alpha)\nabla\psi\cdot\nabla\chi\,dx\Big|\leqslant Ch_\tau^2\|\alpha\|_{\mathcal{C}^2}\int_\tau|\nabla\psi\cdot\nabla\chi|\,dx\leqslant Ch_\tau^2\|\nabla\psi\|_{L_2(\tau)}\|\nabla\chi\|_{L_2(\tau)}.$$

Employing an inverse inequality locally and summing over $\tau\in\mathcal{T}_h$, we obtain

$$|((\tilde\alpha-\alpha)\nabla\psi,\nabla\chi)|\leqslant Ch^{p+q}\|\nabla^q\psi\|\,\|\nabla^p\chi\|. \tag{7.8}$$

In a similar manner we estimate the zero order term. Obviously,

$$(\tilde\beta\psi,J_h\chi)-(\beta\psi,\chi)=\varepsilon_h(\tilde\beta\psi,\chi)+((\tilde\beta-\beta)\psi,\chi). \tag{7.9}$$

Using Lemma 2.2 we can bound the first term on the right-hand side of (7.9), as desired. We then split the second term, in the following way,

$$\begin{aligned}\int_\tau(\tilde\beta-\beta)\psi\,\chi\,dx&=\int_\tau(\tilde\beta-\beta)(\psi\chi)(z_\tau)dx+\int_\tau(\tilde\beta-\beta)(\psi\,\chi-(\psi\chi)(z_\tau))dx\\&=:I+II.\end{aligned} \tag{7.10}$$

Employing (7.7) we easily get

$$|I|\leqslant Ch_\tau^2\|\beta\|_{\mathcal{C}^2}|\tau||(\psi\chi)(z_\tau)|=Ch_\tau^2|\tau|^{-1}\Big|\int_\tau\psi\,dx\Big|\,\Big|\int_\tau\chi\,dx\Big|\leqslant Ch^2\|\psi\|_{L_2(\tau)}\|\chi\|_{L_2(\tau)},$$

and since $|\beta-\tilde\beta|\leqslant Ch_\tau\|\tilde\beta\|_{\mathcal{C}^1}$ in τ,

$$\begin{aligned}|II|&\leqslant Ch_\tau^2\int_\tau\big(|\nabla\psi\,\chi|+|\psi\,\nabla\chi|\big)\,dx\\&\leqslant Ch^2\big(\|\nabla\psi\|_{L_2(\tau)}\|\chi\|_{L_2(\tau)}+\|\psi\|_{L_2(\tau)}\|\nabla\chi\|_{L_2(\tau)}\big).\end{aligned}$$

Combining the bounds for I and II with (7.10), using an inverse inequality locally, summing over $\tau\in\mathcal{T}_h$ and using (7.8), we conclude the proof. □

For the solution operator $\widetilde{E}_h(t)=e^{-\widetilde{A}_h t}$ of (7.2), one shows, as in Lemma 2.1, the following smoothing property.

Lemma 7.2. *For $\widetilde{E}_h$, the solution operator of (7.2), we have, for $v_h\in S_h$ and $t>0$,*

$$\|\nabla^p D_t^\ell\widetilde{E}_h(t)v_h\|\leqslant Ct^{-\ell-(p-q)/2}\|\nabla^q v_h\|,\quad \ell\geqslant 0,\ p,q=0,1,\ 2\ell+p\geqslant q.$$

Further, following the steps in the proof of Lemma 2.3 we can get easily the following estimate.

Lemma 7.3. *Let $\widetilde{A}_h$, Q_h and $\widetilde{Q}_h$ be the operators defined by (7.1) and (7.5). Then*

$$\|\nabla Q_h\chi\|+h\|\widetilde{A}_hQ_h\chi\|\leqslant Ch^{p+1}\|\nabla^p\chi\|,\quad \forall\,\chi\in S_h,\quad \textit{for } p=0,1,$$

and the same bounds hold if we replace Q_h by $\widetilde{Q}_h$.

Proof. Using the fact that $\tilde{a}_h(\chi, J_h\chi) \geqslant c\|\nabla\chi\|^2$, for $\chi \in S_h$, (7.5) and Lemma 2.2, with $\psi = Q_h\chi$, we obtain for $p = 0, 1$,

$$c\|\nabla Q_h\chi\|^2 \leqslant \tilde{a}_h(Q_h\chi, J_hQ_h\chi) = \varepsilon_h(\chi, Q_h\chi) \leqslant Ch^{p+1}\|\nabla^p\chi\|\,\|\nabla Q_h\chi\|,$$

which bounds $Q_h\chi$ as desired. By the definition of $\widetilde{A}_h$ and Lemma 2.2 with $q = 0$, we also get for $p = 0, 1$,

$$|||\widetilde{A}_hQ_h\chi|||^2 = \varepsilon_h(\chi, \widetilde{A}_hQ_h\chi) \leqslant Ch^p\|\nabla^p\chi\|\,\|\widetilde{A}_hQ_h\chi\|.$$

Since the norms $|||\cdot|||$ and $\|\cdot\|$ are equivalent on S_h, this shows the bound stated.

To prove the corresponding bounds for $\widetilde{Q}_h$, analogously we use Lemma 7.1 instead of Lemma 2.2. □

We now show an estimate for $\widehat{\delta}$ defined in (7.6), including exceptionally the exponential decay of the bound.

Lemma 7.4. *For the error $\widehat{\delta}$ defined by* (7.6), *we have*

$$\|\widehat{\delta}(t)\| + h\|\nabla\widehat{\delta}(t)\| \leqslant Ch^2e^{-ct}\|v_h\|, \quad \text{for } t \geqslant 0, \quad v_h \in S_h, \quad \text{with } c > 0.$$

Proof. Using the fact that $\widetilde{E}_h(t)\widetilde{A}_h = -D_t\widetilde{E}_h(t)$, Lemmas 7.2 and 7.3, and the smoothing property (2.2), we find this time taking into account the exponential decay of $\widetilde{E}_h(t)$ and $u_h(t)$ for large t,

$$\begin{aligned}
\|\widehat{\delta}(t)\| + h\|\nabla\widehat{\delta}(t)\| &\leqslant \int_0^t \Big(\|\widetilde{E}_h'(t-s)\widetilde{Q}_hu_h(s)\| + h\|\nabla\widetilde{E}_h(t-s)\widetilde{A}_h\widetilde{Q}_hu_h(s)\|\Big)\,ds \\
&\leqslant C\int_0^t (t-s)^{-1/2}e^{-c(t-s)}\Big(\|\nabla\widetilde{Q}_hu_h(s)\| + h\|\widetilde{A}_h\widetilde{Q}_hu_h(s)\|\Big)\,ds \\
&\leqslant Ch^2\int_0^t (t-s)^{-1/2}e^{-c(t-s)}\|\nabla u_h(s)\|\,ds \\
&\leqslant Ch^2\int_0^t (t-s)^{-1/2}e^{-c(t-s)}s^{-1/2}e^{-cs}\,ds\,\|v_h\| = Ch^2e^{-ct}\|v_h\|,
\end{aligned}$$

which is the desired result. □

We are now ready for the error estimates for the solution of (7.2).

Theorem 7.1. *Let u and $\tilde{u}_h$ be the solutions of* (1.19) *and* (7.2). *Then for $t > 0$,*

$$\|\tilde{u}_h(t) - u(t)\| \leqslant \begin{cases} Ch^2|v|_2, & \text{if } \|v_h - v\| \leqslant Ch^2|v|_2, \\ Ch^2t^{-1/2}|v|_1, & \text{if } v_h = P_hv \text{ and } \|\nabla P_hv\| \leqslant C|v|_1. \end{cases}$$

Further, the estimates for the gradient of the error of Theorem 3.5 remain valid.

Proof. As in Section 3, it suffices to estimate $\delta = \tilde{u}_h - u_h$. Using the splitting (7.6), $\delta = \tilde{\delta} + \widehat{\delta}$, the term $\widehat{\delta}$ is easily bounded by Lemma 7.4, and $\tilde{\delta}$ is bounded as in Theorems 3.1 and 3.2, now applying Lemmas 7.2 and 7.3. □

Turning to nonsmooth initial data, we begin with the following lemma.

Lemma 7.5. *Let u and $\tilde{u}_h$ be the solutions of* (1.19) *and* (7.2)*. Then, for $t > 0$,*

$$\|\tilde{u}_h(t) - u(t) - \widetilde{E}_h(t)\widetilde{A}_h Q_h v_h\| \leqslant Ch^2 t^{-1}\|v\|, \quad \text{if } v_h = P_h v.$$

Proof. Using Lemma 7.4 for $\widehat{\delta}$, it remains to bound $\tilde{\delta}(t) - \widetilde{E}_h(t)\widetilde{A}_h Q_h v_h$, which as for Lemma 3.1, is done as in [4, Theorem 4.1]. □

The following is now our nonsmooth data error estimate. Its proof is an obvious modification of that of Theorem 3.3, using Lemmas 7.2, 7.3 and 7.5.

Theorem 7.2. *Let u and $\tilde{u}_h$ be the solutions of* (1.19) *and* (7.2)*, and let Q_h be defined by* (7.5)*. Then, if* (1.17) *holds, we have*

$$\|\tilde{u}_h(t) - u(t)\| \leqslant Ch^2 t^{-1}\|v\|, \quad \text{if } v_h = P_h v, \quad \text{for } t > 0.$$

Condition (1.17) on Q_h is again satisfied for symmetric meshes:

Theorem 7.3. *For $\{\mathcal{T}_h\}$ symmetric,* (1.17) *holds for Q_h defined by* (7.5)*.*

Proof. We follow the steps in the proof of Theorem 4.1. For given $\chi \in S_h$ we define $\varphi = \varphi_\chi \in \dot{H}^1$ as the solution of the Dirichlet problem $A\varphi = \chi$ in Ω, $\varphi = 0$ on $\partial\Omega$. Since Ω is convex, we have $\varphi \in \dot{H}^2$ and $|\varphi|_2 \leqslant C\|\chi\|$. For $\psi \in S_h$, we have

$$\begin{aligned}\|Q_h\psi\| &= \sup_{\chi\in S_h} \frac{(Q_h\psi, \chi)}{\|\chi\|} = \sup_{\chi\in S_h} \frac{a(Q_h\psi, \varphi)}{\|\chi\|} \\ &\leqslant \sup_{\chi\in S_h} \frac{|a(Q_h\psi, \varphi - I_h\varphi)|}{\|\chi\|} + \sup_{\chi\in S_h} \frac{|a(Q_h\psi, I_h\varphi)|}{\|\chi\|} = I + II.\end{aligned}$$

By the obvious error estimate for I_h and Lemma 7.3, with $p = 0$, we get

$$|I| \leqslant Ch \sup_{\chi\in S_h} \frac{\|\nabla Q_h\psi\|\, |\varphi|_2}{\|\chi\|} \leqslant Ch^2\|\psi\|.$$

To estimate II, we rewrite the numerator in the form

$$a(Q_h\psi, I_h\varphi) = -\tilde{\varepsilon}_h(Q_h\psi, I_h\varphi) + \tilde{a}_h(Q_h\psi, J_h I_h\varphi) = ii_1 + ii_2.$$

In order to complete the proof it suffices to show that

$$|ii_1 + ii_2| \leqslant Ch^2\|\chi\|\,\|\psi\|.$$

Using Lemmas 7.1 and 7.3 we obtain

$$|ii_1| \leqslant Ch^2\|\nabla Q_h\psi\|\,\|\nabla I_h\varphi\| \leqslant Ch^2\|\nabla Q_h\psi\|\,\|\varphi\|_{H^2} \leqslant Ch^2\|\psi\|\,\|\chi\|.$$

Also, employing (7.5) and (4.1) we get

$$ii_2 = \varepsilon_h(\psi, I_h\varphi) = [\psi, M_h I_h\varphi].$$

Since the family $\{\mathcal{T}_h\}$ is symmetric, (4.12) shows the required bound for ii_2. □

The results of Theorems 4.2 and 4.3 for our less restrictive assumptions on the family $\{\mathcal{T}_h\}$ also remain valid, with the obvious modified proofs.

The above results for the spatially semidiscrete finite volume method (1.26) extend in the obvious way to the fully discrete backward Euler method (6.1) and the Crank–Nicolson method (6.3), with $-\widetilde{\Delta}_h$ replaced by $\widetilde{A}_h$, so that Theorems 6.1–6.5 remain literally valid in the general case.

Acknowledgments

The research of R. D. Lazarov was supported in parts by US NSF Grant DMS-1016525 and by award KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). The research of P. Chatzipantelidis was partly supported by the FP7-REGPOT-2009-1 project "Archimedes Center for Modeling Analysis and Computation", funded by the European Commission.

References

[1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] R. E. Bank and D. J. Rose, *Some error estimates for the box method*, SIAM J. Numer. Anal., **24** (1987), 777–787.

[3] P. Chatzipantelidis, R. D. Lazarov, and V. Thomée, *Error estimates for a finite volume element method for parabolic equations in convex polygonal domains*, Numer. Methods Partial Differential Equations, **20** (2004), 650–674.

[4] P. Chatzipantelidis, R. D. Lazarov, and V. Thomée, *Some error estimates for the lumped mass finite element method for a parabolic problem*, Math. Comp., **81** (2012), 1–20.

[5] S. H. Chou and Q. Li, *Error estimates in L^2, H^1 and L^∞ in co–volume methods for elliptic and parabolic problems: A unified approach*, Math. Comp., **69** (2000), 103–120.

[6] J. Huang and S. Xi, *On the finite volume element method for general self-adjoint elliptic problems*, SIAM J. Numer. Anal., **35** (1998), 1762–1774.

[7] M. Křižek and P. Neittaanmäki, *On superconvergence techniques*, Acta Appl. Math., **9** (1987), 175–198.

[8] S. Liang, X. Ma, and A. Zhou, *Finite volume methods for eigenvalue problems*, BIT, **41** (2001), 345–363.

[9] X. Ma, S. Shu, and A. Zhou, *Symmetric finite volume discretizations for parabolic problems*, Comput. Methods Appl. Mech. Engrg., **192** (2003), 4467–4485.

[10] L. Qun and J. Xu, *Linear finite elements with high accuracy*, J. Comput. Math., **3** (1985), 115–133.

[11] R. Sinha, R. Ewing, and R. D. Lazarov, *Some new error estimates of a semidiscrete finite volume element method for a parabolic integro-differential equation with nonsmooth initial data*, SIAM J. Numer. Anal., **43** (2006), 2320–2344.

[12] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd edn., Springer, Berlin, 2006.

[13] L. Wahlbin, *Superconvergence in Galerkin Finite Element Methods*, Lecture Note in Math., 1605, Springer, Berlin, 1995.