Finite Volume Convergence for a Heat Equation with a Measure-Valued Source

Matthias K. Gobbert[1], Martin Kružík[2], Andreas Meister[3], Michael Muscedere[1]

[1]Department of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hilltop Circle, MD 21250 Baltimore, USA.
[2]Czech Academy of Sciences, Institute of Information Theory and Automation, Pod vodárenskou věží 4, 182 00 Praha 8, Czechia & Czech Technical University, Faculty of Civil Engineering Thákurova 7, 166 29 Praha 6, Czechia.
[3]Institute of Mathematics, University of Kassel, Untere Königsstr. 86, 34117 Kassel, Germany.

Contributing authors: gobbert@umbc.edu; kruzik@utia.cas.cz; meister@mathematik.uni-kassel.de; mmusce1@umbc.edu;

Abstract

Past work with Tom Seidman analyzed the spatial convergence order of the numerical finite element method applied to time-dependent diffusion-reaction equations with a measure-valued source term. For smooth source terms, the numerical method has second-order accuracy, but the non-smooth source terms introduce a limitation of the order that depends on the spatial dimension. This work in honor of Tom extends the analysis to the finite volume method in its classical two-point flux form and shows that the regularity of the source term reduces the convergence order to 1 in two and 0.5 in three spatial dimensions which is reduced from the nominal second-order and depends on the space dimension. Computational tests with a point source modeled by a Dirac delta distribution exhibit the behavior predicted by the analytical results.

Keywords: Diffusion-reaction equation; Dirac delta distribution; Finite element method; Finite volume method; Spatial convergence order

This paper is dedicated to the memory of Tom Seidman.

1 Introduction

For co-author MG, Thomas I. "Tom" Seidman was a dear departmental colleague, and many memories connect to him, from the job interview, many conversations and visits to each other's offices which led to joint papers on various topics, e.g., [1, 2], co-teaching a special topics course that leveraged complementary skills on the subject, to serving as the lead-organizer for his 70th birthday conference in 2007 for his 50 years of research [3]. Tom also served on thesis committees, where he became aware of the question of convergence order of the finite element method in the face of thousands of point sources modeling calcium injection sites in a heart cell as Dirac delta distributions [4, 5]. Tom was never afraid to tackle anything, and in this spirit he also engaged with numerical topics without hesitation, thus starting the work on this project, based on the purely numerical evidence in [5–7]. Co-author MM in the same way was connected to Tom by many conversations, colloquium talks, the conference, and a joint research project with MG and Tom. The other co-authors MK and AM enjoyed the wide-ranging conversations with Tom during their visits to the department at UMBC, where the same question was discussed, leading with co-author MK to the joint analysis paper of the finite element method for this problem [8] and with co-author AM to the implementation of the finite volume method for this application problem [9]. The occasion of this special issue in honor of Tom Seidman has now brought the co-authors together for this paper that is an extension of [8] from the finite element method (FEM) to the finite volume method (FVM).

Like in [8] for finite elements, we are interested in the convergence behavior of finite volume schemes applied to diffusion-reaction equations of the form

$$u_t(\mathbf{x},t) - \nabla\cdot(D\,\nabla u(\mathbf{x},t)) = \varphi(\mathbf{x},t) \qquad \text{in } \mathcal{Q} = \Omega\times\mathbb{R}_0^+, \tag{1}$$

$$u(\mathbf{x},t) = 0 \qquad \text{on } \partial\Omega\times\mathbb{R}_0^+ \tag{2}$$

$$u(\mathbf{x},0) = 0 \qquad \text{on } \overline{\Omega}, \tag{3}$$

where $\Omega\subset\mathbb{R}^d$ is a bounded Lipschitz domain in d spatial dimensions, φ includes a smooth or measure-valued source term, combined with Dirichlet boundary conditions, and an initial condition.

In [8, Theorem 4.2], the finite element solution u_h on a mesh with spacing h was shown to have a dimension-dependent convergence order in the L^2-norm with order $\|u(t)-u_h(t)\|_{L^2(\Omega)} \le Ch^\lambda$ with $\lambda\approx 2-d/2$ in $d=2$ and 3 dimensions, that is, order h in 2-D and $h^{1/2}$ in 3-D. We are now able to extend the result to the finite volume method, in the form of the classical two-point flux approximation described in Section 2, with result

$$\|u(t)-u_h(t)\|_{L^2(\Omega)} \le \begin{cases} Ch|\log h|^{1/2}, & d=2,\\ Ch^{1/2}, & d=3, \end{cases} \tag{4}$$

in (60) in Section 5 below, that shows essentially the same convergence behavior. The question of extending the finite element results to the finite volume method had come

up among the co-authors ever since the numerical results in [9], of which Tom was well-aware. Unfortunately, we could not find the time to extend the results together with Tom, but it is very fitting that we have been able to come together on the occasion of this paper.

For the model heat equation (1) with $D = 1$ and $\varphi = \mu$, where μ is a bounded Radon measure, the problem is well understood at the level of the continuous PDE. In particular, the theory of renormalized solutions for nonlinear parabolic equations with general measure data includes the heat equation as the special case $p = 2$ of

$$u_t - \Delta_p u = \mu.$$

A relevant reference is Petitta [10].

For finite volume methods, the situation appears to be less complete. There are finite volume convergence results for parabolic convection–diffusion equations with L^1 right-hand sides, for instance Gallouët, Larcher and Latché [11]. There are also finite volume results for nonlinear parabolic problems with L^1 data in the framework of renormalized solutions, see Guibé and Leclavier [12]. However, a direct finite volume analogue of the elliptic result of Droniou, Gallouët and Herbin [13], with a genuinely Radon-measure right-hand side, does not seem to be as standard for the heat equation.

Thus, one may summarize the state of the literature as follows. The heat equation with measure-valued right-hand side is solved at the PDE level. For finite volume schemes, the available convergence theory is well developed for L^1 parabolic data, while the fully Radon-measure finite volume theory is clearly established in the elliptic case and seems less explicit in the standard parabolic finite volume literature. From the point of view of applications, heat equations with low-regularity/measure right-hand side appear in many models of the continuum mechanics of solids, as e.g. in modeling of shape memory alloys [14].

Since the analysis concerns a specific finite volume scheme, Section 2 first describes this scheme, namely the classical two-point flux approximation, which is the simplest finite volume method for diffusion problems, and introduces the notation used in the analysis. Section 2 also highlights key features of the finite volume method and explains why it is particularly suitable for problems with thousands of measure-valued Dirac source terms that motivated this work originally [4, 5]. The following analysis uses the same approach developed with Tom in [8] by first considering the stationary problem corresponding to (1) before returning to the time-dependent problem. More precisely, Section 3 considers the Poisson equation with a Dirac source, before Sections 4 and 5 consider the heat equation with a space-time Dirac and a persistent Dirac source, respectively. Each section analyzes the convergence in $d = 2$ and $d = 3$ space dimensions and demonstrates the effect of dimensionality in the final results (31) and (32) in Section 3, (53) and (55) in Section 4, and (60) and (63) in Section 5. For the Poisson equation, the scaling arguments of Subsections 3.1 and 3.2 are complemented in Subsection 3.3 by a proof that also accounts for the global propagation of local truncation errors by the discrete solution operator.

Finally, Section 6 provides numerical demonstrations that the behavior predicted by the analysis is exhibited in actual computations; it is comprised of two subsections

for the stationary and time-dependent problems, respectively. Specifically, the convergence orders in Table 1 in Subsection 6.1 exhibit the behavior predicted in (31) and (32) in Section 3 and the convergence orders in Table 3 in Subsection 6.2 the behavior predicted in (63) in Section 5. Both tables' computational results point to the hypothesis of the convergence order being $2 - d/2 = 1$ and $1/2$ in $d = 2$ and 3 space dimensions, respectively.

In summary, this paper shows the analytical proof for the extension of the old result for the finite element method, obtained with Tom Seidman in [8], to the finite volume method, and potentially to other methods like the discontinuous Galerkin method. The accompanying computational evidence confirms the analytical behavior and is obtained by sample code in Matlab for test problems in 2 and 3 spatial dimensions that aligns with previous results from general purpose C code.

2 Finite Volume Method Discretization

2.1 General formulation

The essence of the finite volume method, see e.g. [15] lies in integrating the PDE (1) of the problem over so-called control volumes $\sigma_k \subset \overline{\Omega}$ to derive an integral form of the PDE. To this end, the open domain Ω is decomposed into a finite number of disjoint open control volumes $\sigma_1, \ldots, \sigma_m \subset \Omega$ such that $\cup_{k=1}^{m} \overline{\sigma_k} = \overline{\Omega}$. In our context of a regular quadrilateral domain $\Omega = (-50, 50)^d$ in $d = 1, 2, 3$ dimensions in Section 6, we use a regular mesh with spacing h and N control volumes in each dimension, for a total of $m = N^d$ volumes. The integration of the PDE (1) over each control volume σ_k yields then

$$\frac{d}{dt} \int_{\sigma_k} u(\mathbf{x}, t)\, d\mathbf{x} - \int_{\sigma_k} \nabla \cdot (D\, \nabla u(\mathbf{x}, t))\; d\mathbf{x} = \int_{\sigma_k} \varphi(\mathbf{x}, t)\, d\mathbf{x}, \tag{5}$$

where time differentiation and spatial integration were already interchanged. Introducing the cell averages

$$u_k(t) = \frac{1}{|\sigma_k|} \int_{\sigma_k} u(\mathbf{x}, t)\, d\mathbf{x} \tag{6}$$

makes the following integral form of (5)

$$\frac{d}{dt} u_k(t) - \frac{1}{|\sigma_k|} \int_{\partial\sigma_k} \mathbf{n} \cdot (D\, \nabla u(\mathbf{x}, t))\; ds = \frac{1}{|\sigma_k|} \int_{\sigma_k} \varphi(\mathbf{x}, t)\, d\mathbf{x} \tag{7}$$

eventually a system of equations for the cell averages as unknowns, where the integral over the diffusion term is converted to a surface integral using the Gaussian integral theorem with outer unit normal vector $\mathbf{n}$.

It remains to approximate the normal fluxes in (7). Since (1) contains no convective term, the only fluxes are the diffusive fluxes $\mathbf{n} \cdot (D\, \nabla u)$, which depend linearly on u and are continuous across the faces of the control volumes. Consequently, no reconstruction of left and right states at the faces and no (approximate) Riemann solver are needed, in contrast to finite volume schemes for hyperbolic or convection-dominated problems.

We use the classical two-point flux approximation [16], which is the simplest finite volume method for diffusion problems: each control volume σ_k carries one unknown u_k, which approximates the cell average (6) and is associated with a point $\mathbf{x}_k \in \overline{\sigma_k}$, and the flux through an interior face $e = \partial\sigma_k \cap \partial\sigma_l$ is approximated by the difference quotient

$$\int_e \mathbf{n}_{k,e} \cdot (D\,\nabla u)\,ds \approx F_{k,e}(\mathbf{u}) := D\,|e|\,\frac{u_l - u_k}{d_e}, \tag{8}$$

where $\mathbf{n}_{k,e}$ is the unit normal pointing out of σ_k, $d_e = |\mathbf{x}_k - \mathbf{x}_l|$, and the segment from $\mathbf{x}_k$ to $\mathbf{x}_l$ is orthogonal to e (admissibility of the mesh). For a face $e \subset \partial\Omega$, u_l is replaced by the boundary value $u = 0$ and d_e by the distance from $\mathbf{x}_k$ to e. The fluxes are conservative, $F_{l,e} = -F_{k,e}$. Each unknown thus interacts only with its face neighbors, and no higher-order reconstruction is used. This is what we mean by a *low-order*, more precisely, lowest-order, finite volume method throughout this paper; its accuracy is summarized in Subsection 2.2 below. On the uniform Cartesian meshes of Section 6, (8) reduces in the interior of the domain to central difference quotients, i.e., to the classical five-point ($d = 2$) and seven-point ($d = 3$) stencils, which leads to the Kronecker product form of the matrices given below.

The introduction of the surface integral of the diffusion term in the first term on the right-hand side of (7) allows the finite volume method to conserve the mass in the system at the discrete level, which is one key advantage in application problems such as the calcium flow in a heart cell [4, 5] that motivated this work originally. The other key advantage lies in the ability of the integral in the second term on the right-hand side of (7) to rigorously handle a point source modeled by a Dirac delta $\varphi(\mathbf{x}, t) = \kappa\,\delta_0(\mathbf{x})$ with injection rate κ as

$$\frac{1}{|\sigma_k|}\int_{\sigma_k} \varphi(\mathbf{x}, t)\,d\mathbf{x} = \frac{1}{|\sigma_k|}\int_{\sigma_k} \kappa\,\delta_0(\mathbf{x})\,d\mathbf{x} = \frac{\kappa}{|\sigma_k|} \tag{9}$$

for that control volume σ_k that contains the location of the injection site and 0 otherwise. An analogue of this is possible for the finite element method as in [8], but it is not as convenient for instance for a finite difference method; for such a method, a discrete delta/grid function is often used. In particular, the finite volume method uses only the integrals of the source over the control volumes. It therefore does not distinguish between the Dirac source $\kappa\,\delta_0$ and the regularized source $\kappa\,|\sigma_k|^{-1}\mathbf{1}_{\sigma_k}$ that is uniformly distributed over the control volume containing the injection site; this observation is used in the analysis in Section 5.

For the vector of unknown cell averages as function of time and with the vector of integrals of the source term

$$\mathbf{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}, \quad \mathbf{\Phi}(t) \begin{pmatrix} \frac{1}{|\sigma_1|}\int_{\sigma_1} \varphi(\mathbf{x}, t)\,d\mathbf{x} \\ \vdots \\ \frac{1}{|\sigma_m|}\int_{\sigma_m} \varphi(\mathbf{x}, t)\,d\mathbf{x} \end{pmatrix},$$

the semi-discrete form of the finite volume method can then be formulated in vector notation as

$$\hat{M}\frac{d\mathbf{u}(t)}{dt} + A\,\mathbf{u}(t) = \hat{M}\mathbf{\Phi}(t) \tag{10}$$

with lumped mass matrix $\hat{M} \in \mathbb{R}^{m\times m}$ and stiffness matrix $A \in \mathbb{R}^{m\times m}$ with $m = N^d$ in $d = 1, 2, 3$ spatial dimensions for N the number of control volumes in each spatial dimension. These matrices are block diagonal or tridiagonal matrices with diagonal or tridiagonal blocks, respectively, that can be very effectively pre-computed as sums of Kronecker products $\otimes$ in software such as Matlab. For instance, the matrices are $\hat{M} = S \in \mathbb{R}^{N\times N}$ and $A = \frac{D}{h^2}\,T \in \mathbb{R}^{N\times N}$ in the one-dimensional case, $\hat{M} = S\otimes S \in \mathbb{R}^{N^2\times N^2}$ and $A = \frac{D}{h^2}(S\otimes T + T\otimes S) \in \mathbb{R}^{N^2\times N^2}$ in two dimensions, and $\hat{M} = S\otimes S\otimes S \in \mathbb{R}^{N^3\times N^3}$ and $A = \frac{D}{h^2}(S\otimes S\otimes T + S\otimes T\otimes S + T\otimes S\otimes S) \in \mathbb{R}^{N^3\times N^3}$ in three dimensions, constructed from the diagonal matrix $S \in \mathbb{R}^{N\times N}$ and the tridiagonal matrix $T \in \mathbb{R}^{N\times N}$. Matlab's sparse storage scheme is the key to the efficiency of operations such as matrix-vector products with these matrices.

The exact coefficients in the matrices S, T, A, and $\hat{M}$ depend on the discretization of the diffusion term and the boundary conditions and are designed to facilitate the Kronecker products for efficient implementation, resulting in $S \in \mathbb{R}^{N\times N}$ diagonal with $\frac{1}{2}, 1, \ldots, 1, \frac{1}{2}$ on the diagonal and $T \in \mathbb{R}^{N\times N}$ tridiagonal with $1, 2, \ldots, 2, 1$ on the main diagonal and -1 on the adjacent lower and upper bands. In the current implementation, we use central differences in the interior of the domain and one-sided approximations along the boundary; in the interior, this is exactly the two-point flux approximation (8) on the uniform mesh. The original method uses flux boundary conditions yielding the specific matrices above. The Dirichlet boundary conditions are implemented as correction of certain values in A, namely all diagonal values, including the ones corresponding to control volumes at the boundary, being the same value as the diagonal value for interior control volumes. With this sign convention, A has positive diagonal and nonpositive off-diagonal entries and represents the operator $-\nabla\cdot(D\,\nabla\,\cdot)$, i.e., $\hat{M}^{-1}A$ corresponds to $-\Delta_h$ in the notation of Subsection 2.2 (for $D = 1$); this is why $A\,\mathbf{u}(t)$ appears on the left-hand side of (10).

The implementation just described is the *vertex-centered* variant of the two-point flux approximation: the unknowns are associated with the nodes $\mathbf{x}_k$ of a uniform grid with spacing h that includes the nodes on $\partial\Omega$, and the control volumes are the squares or cubes of side length h centered at the nodes, intersected with Ω. The control volumes at the boundary are therefore halves (and, at the corners, quarters or eighths) of the interior ones, which is the origin of the weights $\frac{1}{2}$ in S. In the interior, the fluxes are given by (8) with $d_e = h$, and the injection site $\mathbf{0}$ is a node of the grid, i.e., the center of its control volume. At the boundary, the Dirichlet condition is imposed by the diagonal correction of A described above rather than by the boundary fluxes of (8). All numerical results in Section 6, in particular those in Tables 1 and 3, were computed with this vertex-centered variant. The analysis in Subsection 3.3 is carried out for the cell-centered variant, in which the points $\mathbf{x}_k$ are the centers of the cells of a uniform Cartesian mesh; Remark 4 explains why the results also hold for the vertex-centered variant.

The time discretization of the semi-discrete problem (10) results necessarily in a stiff system of ordinary differential equations, which necessitates the use of an appropriate ODE solver. The implicit Euler time-stepping method is both appropriate for this situation, simple, and sufficiently accurate if using small time steps. Solving the implicit system for the solution at every new time step by the iterative conjugate gradient method is highly accurate if a tight tolerance is used and still very efficient, since the solution at the old time step can be used as an excellent initial guess.

For a stationary test problem, the time-derivative term in (10) is not present and no variable depends on time, thus we have one linear system, $A\,\mathbf{u} = \hat{M}\boldsymbol{\Phi}$, for the unknowns to solve, which we solve by sparse direct Gaussian elimination, since it is robust and requires no tunable parameters such as stopping tolerances. For the symmetric, diagonally dominant system matrix A, Gaussian elimination is backward stable, so that the relative error of the computed solution is bounded by approximately $\kappa(A)\,\varepsilon_{\text{mach}}$, where $\kappa(A)$ denotes the condition number of A and $\varepsilon_{\text{mach}} \approx 1.1 \cdot 10^{-16}$ the unit roundoff of double-precision arithmetic [17]. The computed solution is therefore not accurate to full double precision: here, $\kappa(A) = O(h^{-2})$ grows under mesh refinement, which is also the origin of the stiffness of the time-dependent problem noted above. However, even on the finest meshes used in Section 6, the resulting algebraic error is many orders of magnitude smaller than the discretization errors reported in Table 1 and therefore does not affect the convergence study.

The size of the system in (10) is given by the number m of control volumes. This becomes a significant scale in higher dimensions as $m = N^d$ in $d = 1, 2, 3$ spatial dimensions for N the number of control volumes in each spatial dimension. The method is implemented in Matlab, fully taking advantage of its efficient vectorization capabilities.

2.2 Notation and accuracy of the scheme used in the analysis

In the analysis of Sections 3–5, we set $D = 1$, denote the mesh by $\mathcal{T}_h = \{\sigma_1, \ldots, \sigma_m\}$, write $K, L \in \mathcal{T}_h$ for generic control volumes, $\mathcal{E}$ for the set of faces, and h for the mesh size. The mesh is assumed to be quasi-uniform, i.e., $\operatorname{diam}(K) \sim h$ and $|K| \sim h^d$, and K_0 denotes the control volume containing the source point $\mathbf{x}_0$. A vector of unknowns $v = (v_K)_{K\in\mathcal{T}_h}$ is identified with the piecewise constant function that equals v_K on K, and P_h denotes the projection onto cell averages, $(P_h u)_K = |K|^{-1} \int_K u \, d\mathbf{x}$. We use the discrete inner product, norm, and energy seminorm

$$(v, w)_h = \sum_{K\in\mathcal{T}_h} |K|\, v_K w_K, \qquad \|v\|_h = (v, v)_h^{1/2}, \qquad |v|_{1,h}^2 = \sum_{e\in\mathcal{E}} \frac{|e|}{d_e} (D_e v)^2, \tag{11}$$

where $D_e v = v_K - v_L$ for an interior face e between K and L (with a fixed orientation), $D_e v = v_K$ for a face $e \subset \partial K \cap \partial\Omega$, and d_e is the corresponding distance in (8). Note that $\|v\|_h$ is the $L^2(\Omega)$ norm of the piecewise constant function v, and that $|\cdot|_{1,h}$ is the discrete energy norm (discrete H_0^1 seminorm), which is denoted by $\|\cdot\|_{\text{energy}}$ in Section 4. The discrete Laplacian is defined by $(\Delta_h v)_K = |K|^{-1} \sum_{e\subset\partial K} F_{K,e}(v)$ with

the fluxes (8) for $D = 1$, so that

$$(-\Delta_h v, w)_h = \sum_{e\in\mathcal{E}} \frac{|e|}{d_e} D_e v\, D_e w =: a_h(v,w), \qquad |v|_{1,h}^2 = a_h(v,v).$$

The finite volume approximations of $-\Delta u = f$ and of $u_t - \Delta u = f$ then read

$$-\Delta_h u_h = f_h \qquad \text{and} \qquad u_h' - \Delta_h u_h = f_h, \qquad \text{with } (f_h)_K = |K|^{-1} \int_K f\, d\mathbf{x}.$$

In particular, $f = \delta_{\mathbf{x}_0}$ leads to $f_h = \mu_h := |K_0|^{-1} \mathbf{1}_{K_0}$, which is also the source vector of the regularized source μ_h; this is used in Section 5. The matrix of a_h is a symmetric M-matrix, so that the scheme satisfies a discrete maximum principle.

Throughout this paper, *low-order* refers to this lowest-order finite volume method with one unknown per control volume, piecewise constant approximations, and two-point fluxes without reconstruction. For smooth solutions on the uniform meshes used in Section 6, it is second-order accurate in the discrete L^2 norm, $\|P_h u - u_h\|_h = O(h^2)$, and (at least) first-order accurate in the discrete energy norm, $|P_h u - u_h|_{1,h} = O(h)$. These are the standard estimates used in Section 4, and the second-order accuracy is confirmed by the smooth test problems mentioned in Section 6. On general admissible meshes, the available error estimates for the two-point flux approximation are in general of first order only [16]. Note that the $L^2(\Omega)$ norm of $u - u_h$ for the piecewise constant function u_h additionally contains the projection error $\|u - P_h u\|_{L^2(\Omega)} = O(h)$; therefore, second-order statements for smooth solutions refer to the discrete norm $\|P_h u - u_h\|_h$, equivalently, to the errors in the cell centers. For the singular solutions in Sections 3 and 5, the error bounds hold both for $\|u - u_h\|_{L^2(\Omega)}$ and for $\|P_h u - u_h\|_h \le \|u - u_h\|_{L^2(\Omega)}$. The implementation uses a discretization at the boundaries that is, in formal terms, first-order accurate; this is appropriate given the behavior of the solution in combination with the homogeneous Dirichlet boundary conditions, and the numerical tests confirm this expectation in Section 6.1.

3 Finite Volume Error for a Poisson Equation with a Dirac Source

3.1 The two-dimensional case

We study the Poisson equation with a Dirac measure right-hand side

$$-\Delta u = \delta_{\mathbf{x}_0} \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{12}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain and $\mathbf{x}_0 \in \Omega$.

Because the right-hand side is a measure, the solution is singular and classical finite element or finite volume convergence rates are not valid.

Let $G(\cdot, \mathbf{x}_0)$ denote the Green function of the Laplacian in Ω with homogeneous Dirichlet boundary conditions. In two dimensions the Green function has the form

$$G(\mathbf{x}, \mathbf{x}_0) = -\frac{1}{2\pi} \log |\mathbf{x} - \mathbf{x}_0| + w(\mathbf{x}), \tag{13}$$

where w is smooth in a neighborhood of $\mathbf{x}_0$.

Therefore the solution of (12) can be written as

$$u(\mathbf{x}) = -\frac{1}{2\pi} \log |\mathbf{x} - \mathbf{x}_0| + w(\mathbf{x}). \tag{14}$$

Let $\mathcal{T}_h$ be a quasi-uniform mesh of Ω with mesh size h, and let u_h denote the finite volume approximation of (12) by the two-point flux scheme of Section 2, identified with a piecewise constant function on $\mathcal{T}_h$ (see Subsection 2.2). We first present a heuristic scaling argument that identifies the sources of the error. It treats u_h as a local approximation of u on each cell; a rigorous proof, which also accounts for the propagation of local errors through Ω by the globally coupled discrete problem, is given in Subsection 3.3.

Let K_0 be the mesh cell containing $\mathbf{x}_0$. Since the mesh is quasi-uniform,

$$\operatorname{diam}(K_0) \sim h, \qquad |K_0| \sim h^2.$$

We split the error as

$$\|u - u_h\|^2_{L^2(\Omega)} = \|u - u_h\|^2_{L^2(K_0)} + \|u - u_h\|^2_{L^2(\Omega \setminus K_0)}. \tag{15}$$

We first estimate the L^2 norm of the singular part on a ball

$$B_h(\mathbf{x}_0) = \{\mathbf{x} : |\mathbf{x} - \mathbf{x}_0| < h\}.$$

Using polar coordinates,

$$\int_{B_h(\mathbf{x}_0)} |\log |\mathbf{x} - \mathbf{x}_0||^2 \, d\mathbf{x} = 2\pi \int_0^h (\log r)^2 r \, dr. \tag{16}$$

A primitive of $r(\log r)^2$ is

$$\frac{r^2}{2}(\log r)^2 - \frac{r^2}{2} \log r + \frac{r^2}{4}.$$

Hence

$$\int_0^h (\log r)^2 r \, dr = \frac{h^2}{2}(\log h)^2 - \frac{h^2}{2} \log h + \frac{h^2}{4}. \tag{17}$$

As $h \to 0$ the dominant term is

$$\frac{h^2}{2}(\log h)^2.$$

Therefore

$$\|u\|_{L^2(B_h(\mathbf{x}_0))} \sim h|\log h|. \tag{18}$$

The size $h|\log h|$ of u on $B_h(\mathbf{x}_0)$ does not by itself determine the error on K_0, because the constant value u_{K_0} of the discrete solution on K_0 need not be small. Denoting by $\overline{u}_{K_0} = |K_0|^{-1}\int_{K_0} u\,d\mathbf{x}$ the mean value of u on K_0, we split

$$\|u-u_h\|_{L^2(K_0)} \le \|u-\overline{u}_{K_0}\|_{L^2(K_0)} + |K_0|^{1/2}\,|\overline{u}_{K_0}-u_{K_0}|. \tag{19}$$

The first term is of order h. Indeed, since $\log|\mathbf{x}-\mathbf{x}_0| = \log h + \log(|\mathbf{x}-\mathbf{x}_0|/h)$, the logarithmic singularity minus the constant $\log h$ is invariant under the scaling $\mathbf{x}-\mathbf{x}_0 = h\,\hat{\mathbf{x}}$, and hence

$$\|u-\overline{u}_{K_0}\|_{L^2(K_0)} \le \frac{1}{2\pi}\left\|\log(|\mathbf{x}-\mathbf{x}_0|/h)\right\|_{L^2(K_0)} + Ch^2 = \frac{h}{2\pi}\left\|\log|\hat{\mathbf{x}}|\right\|_{L^2(\hat{K}_0)} + Ch^2 \le Ch,$$

where $\hat{K}_0 = h^{-1}(K_0-\mathbf{x}_0)$ has a diameter of order one. In particular, the large constant $\frac{1}{2\pi}|\log h|$, which is responsible for the size $h|\log h|$ of u on $B_h(\mathbf{x}_0)$, does not contribute to this term. The second term in (19) measures how well the discrete solution reproduces the mean value of u on the singular cell. It is a genuinely discrete quantity that cannot be inferred from properties of u alone, since u_{K_0} depends on the solution of the globally coupled discrete problem. It is estimated rigorously in Subsection 3.3 with the result $|K_0|^{1/2}|\overline{u}_{K_0}-u_{K_0}| \le Ch|\log h|^{1/2}$. Hence

$$\|u-u_h\|_{L^2(K_0)} \lesssim h|\log h|^{1/2}, \tag{20}$$

which is smaller than the size $h|\log h|$ of u itself on $B_h(\mathbf{x}_0)$: the discrete solution captures the logarithmic growth of u in the singular cell to leading order.

Away from $\mathbf{x}_0$, the solution is smooth.

For a cell K located at distance r from $\mathbf{x}_0$ we have

$$|\nabla^2 u(\mathbf{x})| \sim \frac{1}{r^2}.$$

Standard interpolation estimates give, if u_h is heuristically regarded as a local second-order approximation of u on K,

$$\|u-u_h\|_{L^2(K)} \lesssim h^2|\nabla^2 u||K|^{1/2}.$$

Since $|K|^{1/2} \sim h$ this yields

$$\|u-u_h\|_{L^2(K)} \sim \frac{h^3}{r^2}.$$

Summing contributions over annuli around $\mathbf{x}_0$ gives a total error

$$\|u-u_h\|_{L^2(\Omega\setminus K_0)} \lesssim h. \tag{21}$$

This heuristic far-field bound ignores that u_h is not a local interpolant of u: the local truncation errors, in particular the flux errors of order one at the faces of K_0, are propagated over $\Omega \setminus K_0$ by the discrete Green's function. Moreover, for the piecewise constant function u_h, the local error on a cell K is at best of order $h|\nabla u|\,|K|^{1/2}$, which after summation over the annuli yields $h|\log h|^{1/2}$ instead of h. Both effects are taken into account in Theorem 1 in Subsection 3.3, which shows that the overall error is of order $h|\log h|^{1/2}$, see (31).

3.2 The three-dimensional case

We now consider the analogous problem in three dimensions:

$$-\Delta u = \delta_{\mathbf{x}_0} \quad \text{in } \Omega \subset \mathbb{R}^3, \qquad u = 0 \quad \text{on } \partial\Omega. \tag{22}$$

In three dimensions the Green function has the form

$$u(\mathbf{x}) = \frac{1}{4\pi|\mathbf{x}-\mathbf{x}_0|} + w(\mathbf{x}), \tag{23}$$

where w is smooth near $\mathbf{x}_0$. Thus the singular part is

$$g(\mathbf{x}) = \frac{1}{4\pi r}, \qquad r = |\mathbf{x}-\mathbf{x}_0|.$$

Let K_0 be the mesh cell containing $\mathbf{x}_0$, with diameter h. Comparing K_0 with a ball $B_h(\mathbf{x}_0)$, we obtain

$$\|g\|^2_{L^2(B_h(\mathbf{x}_0))} \sim \int_{B_h(\mathbf{x}_0)} \frac{1}{r^2}\, d\mathbf{x}.$$

Using spherical coordinates,

$$\int_{B_h(\mathbf{x}_0)} \frac{1}{r^2}\, d\mathbf{x} = 4\pi \int_0^h \frac{1}{r^2} r^2\, dr = 4\pi \int_0^h dr \sim h.$$

Therefore

$$\|g\|_{L^2(B_h(\mathbf{x}_0))} \sim h^{1/2}. \tag{24}$$

We use again the splitting (19). In contrast to the two-dimensional case, the first term is now of the same order as $\|g\|_{L^2(B_h(\mathbf{x}_0))}$: under the scaling $\mathbf{x}-\mathbf{x}_0 = h\hat{\mathbf{x}}$, the singular part becomes $g = h^{-1}\hat{g}(\hat{\mathbf{x}})$ with $\hat{g}(\hat{\mathbf{x}}) = 1/(4\pi|\hat{\mathbf{x}}|)$, and since the non-constant function $\hat{g}$ cannot be approximated by constants on $\hat{K}_0$, $\|u - \overline{u}_{K_0}\|_{L^2(K_0)} \sim h^{-1}h^{3/2} = h^{1/2}$. The second term is again estimated in Subsection 3.3 (Remark 1) and is at most of order $h^{1/2}$. Since u_h is constant on K_0, the first term is also a lower bound for the error on K_0, and thus

$$\|u - u_h\|_{L^2(K_0)} \sim h^{1/2}. \tag{25}$$

Away from $\mathbf{x}_0$, the solution is smooth and

$$|\nabla^2 u(x)| \sim \frac{1}{r^3}.$$

Let K be a mesh cell at distance r from x_0. Then a standard local estimate gives, again regarding u_h heuristically as a local second-order approximation of u,

$$\|u - u_h\|_{L^2(K)} \lesssim h^2 |\nabla^2 u|\, |K|^{1/2}.$$

Since in three dimensions $|K|^{1/2} \sim h^{3/2}$, it follows that

$$\|u - u_h\|_{L^2(K)} \lesssim h^2 \frac{1}{r^3} h^{3/2} = \frac{h^{7/2}}{r^3}.$$

Squaring,

$$\|u - u_h\|^2_{L^2(K)} \lesssim \frac{h^7}{r^6}.$$

Now consider a spherical shell

$$A(r, r + dr) = \{x : r < |x - x_0| < r + dr\}.$$

Its volume is approximately

$$4\pi r^2\, dr.$$

Since each cell has volume $\sim h^3$, the number of cells in the shell is

$$N(r) \sim \frac{4\pi r^2\, dr}{h^3}.$$

Therefore the shell contributes

$$dE^2_{\text{far}}(r) \sim N(r) \frac{h^7}{r^6} = \frac{4\pi r^2\, dr}{h^3} \frac{h^7}{r^6} \sim h^4 \frac{dr}{r^4}.$$

Integrating from $r \sim h$ to $r \sim 1$ gives

$$E^2_{\text{far}} \sim h^4 \int_h^1 r^{-4}\, dr.$$

Since

$$\int r^{-4}\, dr = -\frac{1}{3r^3},$$

we obtain

$$E^2_{\text{far}} \sim h^4 \left(\frac{1}{3h^3} - \frac{1}{3} \right) \sim h.$$

Hence

$$\|u - u_h\|_{L^2(\Omega \setminus K_0)} \lesssim h^{1/2}. \tag{26}$$

As in two dimensions, this heuristic regards u_h as a local second-order approximation of u. For the piecewise constant function u_h, the local error on a cell K is at best of order $h|\nabla u|\, |K|^{1/2}$, i.e., of order $h^{5/2} r^{-2}$, since $|\nabla u(\mathbf{x})| \sim r^{-2}$. Summing the squares over the spherical shells as above gives $h^2 \int_h^1 r^{-2}\, dr \sim h$, so this projection

error is again of order $h^{1/2}$ and does not change the rate. The global propagation of the local errors is taken into account in Theorem 1 below.

Both the near-field and far-field contributions are of the same order:

$$\|u - u_h\|_{L^2(K_0)} \sim h^{1/2}, \qquad \|u - u_h\|_{L^2(\Omega\setminus K_0)} \lesssim h^{1/2}.$$

These heuristic scaling arguments therefore suggest

$$\|u - u_h\|_{L^2(\Omega)} \lesssim h^{1/2},$$

which is proved rigorously in Theorem 1 below; as in two dimensions, the rigorous proof also accounts for the global propagation of the local errors.

3.3 Rigorous error bound including the global propagation of local errors

The scaling arguments of Subsections 3.1 and 3.2 treat u_h as a local approximation of u on each cell. Since u_h is the solution of a globally coupled linear system, local truncation errors—in particular the flux errors of order one at the faces of the singular cell K_0—are propagated through Ω by the discrete solution operator, i.e., by the discrete Green's function. We now give a proof that takes this propagation into account. We work in the setting of the numerical experiments of Section 6 and use the notation of Subsection 2.2.

Assumption (A). $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, is a rectangle or a rectangular box, $\mathcal{T}_h$ is a uniform Cartesian mesh of Ω consisting of squares or cubes of side length h, the scheme is the two-point flux approximation (8) with the cell centers as points $\mathbf{x}_K$ (so that $d_e = h$ for interior faces and $d_e = h/2$ for boundary faces; the vertex-centered variant with the boundary treatment of Subsection 2.1, which was used for all computations in Section 6, is covered by Remark 4 below), and there is a constant $c_0 > 0$ independent of h such that $\operatorname{dist}(\mathbf{x}_0, \partial K_0) \geq c_0 h$ (for instance, $\mathbf{x}_0$ is the center of K_0).

For a face $e \in \mathcal{E}$ and a control volume K with $e \subset \partial K$, let $\mathbf{n}_{K,e}$ be the unit normal on e pointing out of K. We compare the exact fluxes $\Phi_{K,e}(u) := -\int_e \nabla u \cdot \mathbf{n}_{K,e}\, ds$ with the discrete fluxes $\Phi^h_{K,e}(v) := \frac{|e|}{d_e}(v_K - v_L) = -F_{K,e}(v)$, where L is the neighbor of K across e and $v_L := 0$ for $e \subset \partial\Omega$, and define the flux consistency errors

$$\tau_{K,e} := \Phi^h_{K,e}(P_h u) - \Phi_{K,e}(u). \tag{27}$$

Both fluxes are conservative, so that $\tau_{L,e} = -\tau_{K,e}$, and we write $\tau_e := \tau_{K,e}$ with the orientation used in the definition of D_e.

Lemma 1 (Error representation)**.** *Let Assumption (A) hold, let u be the solution of (12), and let u_h be its finite volume approximation, $-\Delta_h u_h = \mu_h$. Then $e_h := P_h u - u_h$ satisfies*

$$a_h(e_h, w) = \sum_{e\in\mathcal{E}} \tau_e\, D_e w \qquad \textit{for all grid functions } w. \tag{28}$$

Proof. Since $\mathbf{x}_0 \notin \partial K$ for all $K \in \mathcal{T}_h$, the solution u is smooth in a neighborhood of each face, and the divergence theorem together with $-\Delta u = \delta_{\mathbf{x}_0}$ gives the exact flux balance $\sum_{e\subset\partial K} \Phi_{K,e}(u) = 1$ if $K = K_0$ and $= 0$ otherwise. The scheme reads $\sum_{e\subset\partial K} \Phi^h_{K,e}(u_h) = |K|\,(\mu_h)_K$, which is the same right-hand side. Subtracting and using the linearity of $\Phi^h_{K,e}$ yields $\sum_{e\subset\partial K} \Phi^h_{K,e}(e_h) = \sum_{e\subset\partial K} \tau_{K,e}$ for every K. Multiplying by w_K, summing over K, and collecting the two contributions of each interior face gives (28). □

Lemma 2 (Flux consistency)**.** *Let Assumption (A) hold. Then there are constants C and $h_0 > 0$ independent of h such that for $h \le h_0$*

$$|\tau_e| \le \begin{cases} C, & \operatorname{dist}(e,\mathbf{x}_0) < 2h, \\ C\,h^{d+1}\operatorname{dist}(e,\mathbf{x}_0)^{-d-1}, & \operatorname{dist}(e,\mathbf{x}_0) \ge 2h,\ e \not\subset \partial\Omega, \\ C\,h^{d}, & e \subset \partial\Omega, \end{cases} \quad \text{and hence} \quad \sum_{e\in\mathcal{E}} |\tau_e| \le C. \tag{29}$$

Proof. By odd reflection across the sides of Ω, the solution u is smooth on $\overline{\Omega} \setminus \{\mathbf{x}_0\}$, and $|\nabla^k u(\mathbf{x})| \le C|\mathbf{x}-\mathbf{x}_0|^{2-d-k}$ for $k = 1,2,3$. Let e be an interior face between K and L with $\mathbf{n}_{K,e} = \mathbf{e}_i$. A direct computation shows that the discrete flux density is a weighted average of $\partial_i u$ over $Z_e := K \cup L$,

$$\frac{(P_h u)_L - (P_h u)_K}{h} = \frac{1}{|e|}\int_e \int_{-h}^{h} \phi_h(\sigma)\,\partial_i u(\mathbf{y} + \sigma\mathbf{e}_i)\,d\sigma\,ds(\mathbf{y}), \qquad \phi_h(\sigma) = \frac{h - |\sigma|}{h^2},$$

with the symmetric weight $\phi_h \ge 0$, $\int_{-h}^{h} \phi_h\,d\sigma = 1$, whereas the exact flux density is $|e|^{-1}\int_e \partial_i u\,ds$. If $\operatorname{dist}(e,\mathbf{x}_0) \ge 2h$, then $|\mathbf{x}-\mathbf{x}_0| \ge \operatorname{dist}(e,\mathbf{x}_0)/2$ on Z_e; a Taylor expansion in σ, in which the linear term vanishes by the symmetry of ϕ_h, gives $|\tau_e| \le C|e|\,h^2 \sup_{Z_e}|\nabla^3 u| \le Ch^{d+1}\operatorname{dist}(e,\mathbf{x}_0)^{-d-1}$. If $\operatorname{dist}(e,\mathbf{x}_0) < 2h$, then e does not intersect the open cell K_0, so that $\operatorname{dist}(e,\mathbf{x}_0) \ge \operatorname{dist}(\mathbf{x}_0,\partial K_0) \ge c_0 h$ and $|\Phi_{K,e}(u)| \le |e| \sup_e |\nabla u| \le Ch^{d-1}(c_0 h)^{1-d} \le C$; moreover, $Z_e \subset B_{5h}(\mathbf{x}_0)$ and $\phi_h \le h^{-1}$ give $|\Phi^h_{K,e}(P_h u)| \le h^{-1}\int_{Z_e}|\nabla u|\,d\mathbf{x} \le Ch^{-1}\int_{B_{5h}(\mathbf{x}_0)} |\mathbf{x}-\mathbf{x}_0|^{1-d}\,d\mathbf{x} \le C$. For a boundary face $e \subset \partial K \cap \partial\Omega$, a Taylor expansion in the normal direction, using $u = 0$ on e and $d_e = h/2$, gives $|\tau_e| \le C|e|\,h \sup_K |\nabla^2 u| \le Ch^d$, since K stays away from $\mathbf{x}_0$ for $h \le h_0$. Finally, there are at most C faces with $\operatorname{dist}(e,\mathbf{x}_0) < 2h$, at most Ck^{d-1} interior faces with $kh \le \operatorname{dist}(e,\mathbf{x}_0) < (k+1)h$, and at most Ch^{1-d} boundary faces. Hence $\sum_e |\tau_e| \le C\big(1 + \sum_{k\ge 2} k^{d-1}k^{-d-1} + h\big) \le C$. □

Lemma 3 (Stability of the discrete dual problem)**.** *Let Assumption (A) hold, let ψ be a grid function, and let z_h be the solution of $a_h(w,z_h) = (\psi,w)_h$ for all grid functions w, i.e., $-\Delta_h z_h = \psi$. Then*

$$\max_{e\in\mathcal{E}} |D_e z_h| \le C\,h\,\ell_h\,\|\psi\|_h, \qquad \ell_h := \begin{cases} (1 + |\log h|)^{1/2}, & d = 2, \\ h^{-1/2}, & d = 3, \end{cases} \tag{30}$$

with C independent of h and ψ.

Proof. Under Assumption (A), the matrix of $-\Delta_h$ is the Kronecker sum $\sum_{i=1}^{d} I \otimes \cdots \otimes T_i \otimes \cdots \otimes I$ of symmetric positive definite tridiagonal matrices T_i, one for each coordinate direction. Since these summands commute, expanding $\|\Delta_h z_h\|_h^2$ yields $\|(I \otimes \cdots \otimes T_i \otimes \cdots \otimes I) z_h\|_h \le \|\Delta_h z_h\|_h = \|\psi\|_h$, and similarly for the mixed terms involving $T_i^{1/2} \otimes T_j^{1/2}$, $i \neq j$. In terms of difference quotients, this is a discrete H^2-regularity estimate: all second difference quotients of z_h, including the mixed ones, are bounded by $C\|\psi\|_h$ in the discrete L^2 norm. Consequently, each first difference quotient $\partial_{h,i} z_h$, regarded as a grid function on the faces normal to the i-th coordinate direction, has a discrete H^1 norm bounded by $C\|\psi\|_h$. The discrete Sobolev inequality $\|v\|_\infty \le C(1 + |\log h|)^{1/2} \|v\|_{1,h}$ in two dimensions [18], and the discrete Sobolev inequality $\|v\|_{L_h^6} \le C\|v\|_{1,h}$ combined with the inverse estimate $\|v\|_\infty \le h^{-1/2} \|v\|_{L_h^6}$ in three dimensions, give $\max |\partial_{h,i} z_h| \le C\ell_h \|\psi\|_h$. Since $|D_e z_h| = h\,|\partial_{h,i} z_h|$ on the faces normal to the i-th direction ($h/2$ instead of h on boundary faces), (30) follows. □

Theorem 1. *Let Assumption (A) hold, let u be the solution of (12), and let u_h be its finite volume approximation, identified with a piecewise constant function. Then there are constants C and $h_0 > 0$ independent of h such that for all $h \le h_0$*

$$\|u - u_h\|_{L^2(\Omega)} \le C\, h\, |\log h|^{1/2}, \qquad d = 2, \tag{31}$$

$$\|u - u_h\|_{L^2(\Omega)} \le C\, h^{1/2}, \qquad d = 3. \tag{32}$$

The same bounds hold for the discrete error $\|P_h u - u_h\|_h$.

Proof. We split $u - u_h = (u - P_h u) + e_h$ with $e_h = P_h u - u_h$, and note that $h \le h\ell_h$, $h\ell_h = h^{1/2}$ for $d = 3$, and $h\ell_h \le Ch|\log h|^{1/2}$ for $d = 2$ and $h \le h_0 < 1$.

Projection error. For cells K with $\operatorname{dist}(K, \mathbf{x}_0) \ge h$, the Poincaré inequality on K gives $\|u - (P_h u)_K\|_{L^2(K)} \le Ch\|\nabla u\|_{L^2(K)}$, and summation yields

$$\sum_{\operatorname{dist}(K,\mathbf{x}_0)\ge h} \|u - (P_h u)_K\|_{L^2(K)}^2 \le Ch^2 \int_{\Omega \cap \{|\mathbf{x}-\mathbf{x}_0| \ge h\}} |\mathbf{x} - \mathbf{x}_0|^{2-2d}\, d\mathbf{x} \le Ch^2 \ell_h^2.$$

For each of the boundedly many remaining cells, which lie in $B_{3h}(\mathbf{x}_0)$, the scaling argument used for the first term in (19) gives $\|u - (P_h u)_K\|_{L^2(K)} \le Ch$ for $d = 2$, and $\|u - (P_h u)_K\|_{L^2(K)} \le \|u\|_{L^2(K)} \le C\big\||\mathbf{x} - \mathbf{x}_0|^{-1}\big\|_{L^2(B_{3h}(\mathbf{x}_0))} + Ch^{3/2} \le Ch^{1/2}$ for $d = 3$. Hence $\|u - P_h u\|_{L^2(\Omega)} \le Ch\ell_h$.

Discrete error. Let z_h be the solution of the dual problem of Lemma 3 with $\psi = e_h$. By Lemmas 1, 2 and 3,

$$\|e_h\|_h^2 = a_h(e_h, z_h) = \sum_{e\in\mathcal{E}} \tau_e\, D_e z_h \le \Big(\sum_{e\in\mathcal{E}} |\tau_e|\Big) \max_{e\in\mathcal{E}} |D_e z_h| \le C\, h\, \ell_h\, \|e_h\|_h.$$

Hence $\|P_h u - u_h\|_h = \|e_h\|_h \le Ch\ell_h$, and the triangle inequality completes the proof. □

Remark 1 (Error on the singular cell)**.** *Since $|K_0|^{1/2} |\overline{u}_{K_0} - u_{K_0}| = \|e_h\|_{L^2(K_0)} \le \|e_h\|_h$, the proof yields $|K_0|^{1/2} |\overline{u}_{K_0} - u_{K_0}| \le Ch\ell_h$, which is the bound for the second*

term in (19) used in Subsections 3.1 and 3.2. For $d = 2$, this means $|\overline{u}_{K_0} - u_{K_0}| \leq C|\log h|^{1/2}$, whereas $\overline{u}_{K_0} = \frac{1}{2\pi}|\log h| + O(1)$: the discrete solution reproduces the logarithmic growth of u in the singular cell to leading order. For $d = 3$, the error on K_0 is bounded from below by $\|u - \overline{u}_{K_0}\|_{L^2(K_0)} \geq c\, h^{1/2}$, so that the rate in (32) cannot be improved.

Remark 2 (Mechanism of the order reduction)**.** *The proof makes the propagation of local errors explicit. By Lemma 2, the far-field flux errors are summable, while the finitely many faces near $\mathbf{x}_0$ carry flux errors of order one, which do not decrease under mesh refinement. Through (28), these errors are propagated by the discrete dual solution, i.e., by the discrete Green's function, and they enter the global error multiplied by the discrete gradients $D_e z_h$ near $\mathbf{x}_0$. The size $h\ell_h$ of these gradients reflects that the dual solution is only in H^2, which does not embed into $W^{1,\infty}$ for $d = 2, 3$. This is the mechanism that reduces the convergence order from the nominal value 2 to $2 - d/2$ (up to a logarithmic factor for $d = 2$), in analogy with the finite element result of Scott [19], who proved the L^2 error bound $O(h)$ for piecewise linear finite elements in two dimensions.*

Remark 3 (Extensions)**.** *Assumption (A) enters only through the summability of the flux errors and the dual stability (30). On general quasi-uniform admissible meshes, the two-point flux approximation is in general only first-order consistent [16], so that the far-field flux errors are of order $h^d \operatorname{dist}(e, \mathbf{x}_0)^{-d}$ and their sum grows like $|\log h|$; provided a dual stability estimate like (30) is available, the argument then yields the same rates up to an additional logarithmic factor. Moreover, if the dual problem with data $\psi \in L^\infty(\Omega)$ satisfies the discrete $W^{1,\infty}$ estimate $\max_e |D_e z_h| \leq Ch\|\psi\|_{L^\infty(\Omega)}$, which for the five- and seven-point schemes is a consequence of discrete $W^{2,p}$-regularity with $p > d$, the same duality argument in L^1 gives $\|P_h u - u_h\|_{L^1(\Omega)} \leq Ch$. Since $\|u - P_h u\|_{L^1(\Omega)} \leq Ch\|\nabla u\|_{L^1(\Omega)}$ with $\nabla u \in L^1(\Omega)$, this implies $\|u - u_h\|_{L^1(\Omega)} \leq Ch$ for $d = 2, 3$.*

Remark 4 (Vertex-centered variant used in the computations)**.** *Theorem 1 also holds for the vertex-centered variant of Subsection 2.1, with which all results in Section 6 were computed. In this variant, the mesh is uniform and Cartesian away from $\partial\Omega$, the interior fluxes are (8) with $d_e = h$, and $\mathbf{x}_0$ is the center of K_0, so that $\operatorname{dist}(\mathbf{x}_0, \partial K_0) = h/2$. Instead of the boundary fluxes of (8), the diagonal correction of Subsection 2.1 adds to a_h the terms $\beta_K v_K w_K$ for the control volumes K at $\partial\Omega$, with $\beta_K > 0$ of order h^{d-2}. The matrix of a_h is thus again a symmetric M-matrix, with positive row sums in the boundary rows, so that the discrete maximum principle holds, and the discrete Poincaré inequality $\|v\|_h \leq C\, a_h(v, v)^{1/2}$ follows as for the two-point scheme. However, because of the diagonal correction, the matrix is in general not a Kronecker sum, so that Lemma 3 cannot be applied directly, and the boundary closure is only of first order. We therefore modify the proof of Theorem 1 in the treatment of the boundary only. Let $\chi \in C_c^\infty(\Omega)$ with $\chi = 1$ in a neighborhood of $\mathbf{x}_0$, and split $u = u^{\mathrm{s}} + u^{\mathrm{r}}$ with $u^{\mathrm{s}} := \chi u$ and $u^{\mathrm{r}} := (1 - \chi)u$. Then $-\Delta u^{\mathrm{s}} = \delta_{\mathbf{x}_0} + g$ and $-\Delta u^{\mathrm{r}} = -g$ with a smooth function g supported in $\operatorname{supp} \nabla\chi$, the function u^{r} is smooth on $\overline{\Omega}$ (cf. the proof of Lemma 2), and, by linearity, $u_h = u_h^{\mathrm{s}} + u_h^{\mathrm{r}}$ with the finite volume approximations u_h^{s} and u_h^{r} of these two problems.*

Singular part. *For h small enough, $P_h u^{\mathrm{s}}$ vanishes on all control volumes outside a fixed compact set $\Omega' \Subset \Omega$. The proof of Lemma 1 thus shows that $P_h u^{\mathrm{s}} - u_h^{\mathrm{s}}$ satisfies* (28) *with flux errors τ_e of u^{s} that vanish for all faces $e \not\subset \Omega'$, in particular for all faces at $\partial\Omega$, and that satisfy the interior bounds of Lemma 2, since the mesh is uniform and Cartesian in Ω' and $\mathbf{x}_0$ is the center of K_0; hence $\sum_e |\tau_e| \le C$. The duality argument in the proof of Theorem 1 therefore requires* (30) *only for faces $e \subset \Omega'$. To show this interior version of Lemma 3, let $\theta \in C_c^\infty(\Omega)$ with $\theta = 1$ in a neighborhood of Ω'. The grid function θz_h vanishes near $\partial\Omega$ and hence solves the five- or seven-point problem with homogeneous Dirichlet conditions on a rectangular box $B \Subset \Omega$, whose matrix is a Kronecker sum, so that the proof of Lemma 3 applies on B. The right-hand side of this problem is $\theta\psi$ plus terms that involve z_h and its first difference quotients multiplied by derivatives of θ; by the energy estimate and the discrete Poincaré inequality, their discrete L^2 norm is bounded by $C(\|z_h\|_h + a_h(z_h, z_h)^{1/2}) \le C\|\psi\|_h$. Hence $|D_e z_h| = |D_e(\theta z_h)| \le Ch\ell_h\|\psi\|_h$ for $e \subset \Omega'$, and the proof of Theorem 1 gives $\|P_h u^{\mathrm{s}} - u_h^{\mathrm{s}}\|_h \le Ch\ell_h$.*

Regular part. *We compare u_h^{r} with the nodal interpolant $I_h u^{\mathrm{r}}$, which vanishes at the boundary nodes. The residual of $I_h u^{\mathrm{r}}$ in the scheme $-\Delta_h v = f_h$ is of order h^2 in the interior rows, as usual for the five- and seven-point stencils, and of order h^{-1} in the boundary rows, which reflects the first-order boundary closure. Since the row sums of $-\Delta_h$ are bounded from below by ch^{-2} in the boundary rows, the discrete maximum principle with the barrier function $C\big(h + h^2(R^2 - |\mathbf{x}|^2)\big)$, with R sufficiently large, gives $\max_K |(I_h u^{\mathrm{r}})_K - (u_h^{\mathrm{r}})_K| \le Ch$, and hence $\|P_h u^{\mathrm{r}} - u_h^{\mathrm{r}}\|_h \le Ch$.*

Together with the estimate of the projection error $u - P_h u$, which is not affected by the boundary treatment, this proves (31), (32), *and the bound for $\|P_h u - u_h\|_h$ also for the vertex-centered variant. In particular, the first-order boundary closure contributes only $O(h)$, which does not exceed the order $h\ell_h$ caused by the singularity. The same modifications apply to Step 5 of the proof of Proposition 2 and to Proposition 3: the boundary rows are treated by the semidiscrete or fully discrete maximum principle with the same time-independent barrier function, which is possible since $\hat{M}^{-1}A$ has nonpositive off-diagonal entries, so that the semigroup $E_h(s) = e^{s\Delta_h}$ and the backward Euler step preserve positivity.*

4 Finite Volume Error for a Heat Equation with a Space-Time Dirac Source

We consider the heat equation

$$u_t - \Delta u = \delta_{(x_0,t_0)} \qquad \text{in } \Omega \times (0,T) \tag{33}$$

with homogeneous Dirichlet boundary conditions

$$u = 0 \qquad \text{on } \partial\Omega \times (0,T),$$

and vanishing initial condition

$$u(x,0) = 0.$$

The right-hand side is a Dirac mass concentrated at the space-time point (x_0, t_0).

4.1 Exact solution in the whole space

In $\mathbb{R}^d$, the fundamental solution of the heat equation is the heat kernel

$$G(\mathbf{x}, t) = (4\pi t)^{-d/2} \exp\left(-\frac{|\mathbf{x}|^2}{4t}\right), \qquad t > 0. \tag{34}$$

Therefore, in the whole space, the solution of (33) is

$$u(\mathbf{x}, t) = \begin{cases} 0, & t < t_0, \\ G(\mathbf{x} - \mathbf{x}_0, t - t_0), & t > t_0. \end{cases} \tag{35}$$

Thus, if we set

$$s := t - t_0 > 0,$$

then

$$u(\mathbf{x}, t) = (4\pi s)^{-d/2} \exp\left(-\frac{|\mathbf{x} - \mathbf{x}_0|^2}{4s}\right). \tag{36}$$

This shows that the singularity is instantaneous in time: at $t = t_0$ the solution is measure-valued, but for every fixed $t > t_0$ the solution is smooth in space.

4.2 Scaling of the solution

The heat kernel has the self-similar form

$$u(\mathbf{x}, t) = s^{-d/2}\, \Phi\left(\frac{\mathbf{x} - \mathbf{x}_0}{\sqrt{s}}\right), \qquad \Phi(\mathbf{y}) := (4\pi)^{-d/2} e^{-|\mathbf{y}|^2/4}. \tag{37}$$

Differentiating with respect to $\mathbf{x}$ gives

$$\nabla^m u(\mathbf{x}, t) = s^{-d/2} s^{-m/2} \Psi_m\left(\frac{\mathbf{x} - \mathbf{x}_0}{\sqrt{s}}\right), \tag{38}$$

where Ψ_m is a polynomial times a Gaussian distribution function. Therefore

$$|\nabla^m u(\mathbf{x}, t)| \sim s^{-d/2 - m/2} \times \text{Gaussian factor}. \tag{39}$$

Taking the L^2 norm and using the change of variables

$$\mathbf{y} = \frac{\mathbf{x} - \mathbf{x}_0}{\sqrt{s}}, \qquad d\mathbf{x} = s^{d/2}\, d\mathbf{y},$$

we obtain

$$\|\nabla^m u(\cdot, t)\|^2_{L^2(\mathbb{R}^d)} \sim \int_{\mathbb{R}^d} s^{-d-m} \left|\Psi_m\left(\frac{\mathbf{x} - \mathbf{x}_0}{\sqrt{s}}\right)\right|^2 d\mathbf{x} \tag{40}$$

$$= s^{-d-m} s^{d/2} \int_{\mathbb{R}^d} |\Psi_m(\mathbf{y})|^2 \, d\mathbf{y} \tag{41}$$

$$\sim s^{-m-d/2}. \tag{42}$$

Hence

$$\|\nabla^m u(\cdot,t)\|_{L^2(\mathbb{R}^d)} \sim s^{-m/2-d/4}. \tag{43}$$

In particular, for $m = 2$,

$$|u(\cdot,t)|_{H^2} \sim s^{-1-d/4}. \tag{44}$$

4.3 Finite volume approximation

Let $u_h(\cdot,t)$ be the semidiscrete approximation by the two-point flux finite volume method of Section 2, i.e., $u_h' - \Delta_h u_h = 0$ for $t > t_0$ with $u_h(t_0) = \mu_h$, on a quasi-uniform mesh of size h. As explained in Subsection 2.2, *low-order* refers to this lowest-order scheme with one unknown per control volume, piecewise constant approximations, and two-point fluxes without reconstruction. For smooth solutions on the uniform meshes considered here, it is second-order accurate in the discrete L^2 norm and first-order accurate in the discrete energy norm $\|\cdot\|_{\text{energy}} = |\cdot|_{1,h}$, where both norms are applied to the grid function $P_h u - u_h$ that compares the cell averages of u with the discrete solution, see Subsection 2.2. In this section, all error estimates are therefore stated for $P_h u - u_h$ in these discrete norms.

Accordingly, for smooth functions, standard spatial error estimates for this scheme have the form

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_h \lesssim h^2 \, |u(\cdot,t)|_{H^2(\Omega)}, \tag{45}$$

and, in a discrete energy norm,

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_{\text{energy}} \lesssim h \, |u(\cdot,t)|_{H^2(\Omega)}. \tag{46}$$

The point is that for every fixed $t > t_0$, the heat kernel is smooth, so these estimates apply. The only special feature is that the H^2 norm depends singularly on

$$s = t - t_0.$$

Substituting (44) into (45) and (46), we obtain

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_h \lesssim h^2 \, s^{-1-d/4}, \qquad s = t - t_0 > 0, \tag{47}$$

and

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_{\text{energy}} \lesssim h \, s^{-1-d/4}. \tag{48}$$

4.4 Explicit rates in two dimensions

If $d = 2$, then (44) becomes

$$|u(\cdot,t)|_{H^2} \sim s^{-3/2}. \tag{49}$$

Therefore

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_h \lesssim h^2\, s^{-3/2}, \qquad s = t - t_0 > 0, \tag{50}$$

and

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_{\text{energy}} \lesssim h\, s^{-3/2}. \tag{51}$$

Thus, in two dimensions, for every fixed positive time after the impulse, the spatial finite volume method recovers the standard rates:

$$O(h^2) \quad \text{in the discrete } L^2 \text{ norm}, \qquad O(h) \quad \text{in the discrete energy norm}.$$

4.5 Explicit rates in three dimensions

If $d = 3$, then (44) becomes

$$|u(\cdot,t)|_{H^2} \sim s^{-7/4}. \tag{52}$$

Therefore

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_h \lesssim h^2\, s^{-7/4}, \qquad s = t - t_0 > 0, \tag{53}$$

and

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_{\text{energy}} \lesssim h\, s^{-7/4}. \tag{54}$$

For the heat equation

$$u_t - \Delta u = \delta_{(\mathbf{x}_0,t_0)},$$

the low-order (two-point flux) finite volume method satisfies, for every fixed $t > t_0$,

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_h \lesssim \begin{cases} h^2 (t-t_0)^{-3/2}, & d = 2, \\ h^2 (t-t_0)^{-7/4}, & d = 3, \end{cases} \tag{55}$$

and

$$\|P_h u(\cdot,t) - u_h(\cdot,t)\|_{\text{energy}} \lesssim \begin{cases} h (t-t_0)^{-3/2}, & d = 2, \\ h (t-t_0)^{-7/4}, & d = 3. \end{cases} \tag{56}$$

For the piecewise constant function $u_h(\cdot,t)$ itself, the $L^2(\Omega)$ norm of $u(\cdot,t) - u_h(\cdot,t)$ additionally contains the projection error, which is only of first order: by the Poincaré inequality on each control volume and (43) with $m = 1$, $\|u(\cdot,t) - P_h u(\cdot,t)\|_{L^2(\Omega)} \le Ch\, |u(\cdot,t)|_{H^1(\Omega)} \lesssim h\, (t-t_0)^{-1/2-d/4}$, and hence

$$\|u(\cdot,t) - u_h(\cdot,t)\|_{L^2(\Omega)} \lesssim h\, (t-t_0)^{-1/2-d/4} + h^2\, (t-t_0)^{-1-d/4}.$$

5 Finite Volume Error for a Heat Equation with a Persistent Dirac Source

Proposition 2. *Let $\Omega \subset \mathbb{R}^d$ $(d \ge 2)$ be a bounded domain and $\mathbf{x}_0 \in \Omega$. Consider*

$$\partial_t u - \Delta u = \delta_{\mathbf{x}_0}\, \mathbf{1}_{t \ge t_0} \quad \text{in } \Omega \times (0,T) \qquad u = 0 \text{ on } \partial\Omega, \qquad u(0) = 0. \tag{57}$$

Let $\mathcal{T}_h$ be a quasi-uniform admissible finite volume mesh and let $K_0 \in \mathcal{T}_h$ be the control volume containing $\mathbf{x}_0$; for the estimate of the discretization error (Step 5 of the proof), we assume that Ω, $\mathcal{T}_h$, and $\mathbf{x}_0$ satisfy Assumption (A) of Subsection 3.3, or that the vertex-centered variant of Remark 4 is used. Define the regularized source

$$\mu_h := \frac{1}{|K_0|}\mathbf{1}_{K_0}.$$

Let $u^{(h)}$ solve

$$\partial_t u^{(h)} - \Delta u^{(h)} = \mu_h \, \mathbf{1}_{t\geq t_0}, \qquad u^{(h)}|_{\partial\Omega} = 0, \qquad u^{(h)}(0) = 0, \tag{58}$$

and let $u_h(t)$ be the semidiscrete finite volume approximation, $u_h' - \Delta_h u_h = \mu_h \mathbf{1}_{t\geq t_0}$, $u_h(0) = 0$, see Subsection 2.2.

Then for every $\tau > 0$ there exists $C > 0$ such that for all $t \in [t_0 + \tau, T]$,

$$\|u(t) - u_h(t)\|_{L^1(\Omega)} \leq Ch, \tag{59}$$

$$\|u(t) - u_h(t)\|_{L^2(\Omega)} \leq \begin{cases} Ch|\log h|^{1/2}, & d = 2, \\ Ch^{1/2}, & d = 3. \end{cases} \tag{60}$$

Moreover, for any compact $K \Subset \Omega \setminus \{\mathbf{x}_0\}$,

$$\|P_h u(t) - u_h(t)\|_{L^p(K)} \leq C(h^r),$$

where r is the order of the underlying FV scheme for smooth solutions; for the piecewise constant function $u_h(t)$ itself, the projection error $\|u(t) - P_h u(t)\|_{L^p(K)} = O(h)$ has to be added, cf. Subsection 2.2.

Proof. We split the error as

$$u - u_h = (u - u^{(h)}) + (u^{(h)} - u_h) =: \eta_h + \xi_h.$$

Step 1: representation formula. Let $G_D(t, \mathbf{x}, \mathbf{y})$ be the Dirichlet heat kernel. Then

$$u(\mathbf{x}, t) = \int_{t_0}^{t} G_D(t - s, \mathbf{x}, \mathbf{x}_0)\, ds, \quad u^{(h)}(\mathbf{x}, t) = \int_{t_0}^{t} \frac{1}{|K_0|} \int_{K_0} G_D(t - s, \mathbf{x}, \mathbf{y})\, d\mathbf{y}\, ds.$$

Hence

$$\eta_h(\mathbf{x}, t) = \int_{t_0}^{t} \left(G_D(t - s, \mathbf{x}, \mathbf{x}_0) - \frac{1}{|K_0|} \int_{K_0} G_D(t - s, \mathbf{x}, \mathbf{y})\, d\mathbf{y} \right) ds.$$

Step 2: pointwise estimate. For $|\mathbf{x}-\mathbf{x}_0| \geq 2h$, by the mean value theorem in $\mathbf{y}$,

$$|\eta_h(\mathbf{x},t)| \leq Ch \int_{t_0}^{t} \sup_{\mathbf{y}\in K_0} |\nabla_{\mathbf{y}} G_D(t-s,\mathbf{x},\mathbf{y})|\, ds.$$

Using the classical estimate

$$\int_0^\infty |\nabla_{\mathbf{y}} G_D(\tau,\mathbf{x},\mathbf{y})|\, d\tau \lesssim |\mathbf{x}-\mathbf{x}_0|^{1-d},$$

we obtain

$$|\eta_h(\mathbf{x},t)| \lesssim h|\mathbf{x}-\mathbf{x}_0|^{1-d}, \qquad |\mathbf{x}-\mathbf{x}_0| \geq 2h.$$

Step 3: L^1 estimate. Integrating in polar coordinates,

$$\|\eta_h(t)\|_{L^1(\Omega)} \lesssim h \int_h^1 r^{1-d} r^{d-1} dr \lesssim h.$$

Step 4: L^2 estimate. Similarly,

$$\|\eta_h(t)\|_{L^2(\Omega)}^2 \lesssim h^2 \int_h^1 r^{1-d} dr,$$

which gives

$$\|\eta_h(t)\|_{L^2} \lesssim \begin{cases} h|\log h|^{1/2}, & d=2, \\ h^{1/2}, & d=3. \end{cases}$$

Here the pointwise bound of Step 2 has been used for $|\mathbf{x}-\mathbf{x}_0| \geq 2h$ only. In the remaining ball $B_{2h}(\mathbf{x}_0)$, we use that $\int_0^\infty G_D(\sigma,\mathbf{x},\mathbf{y})\, d\sigma = G(\mathbf{x},\mathbf{y})$ is the Dirichlet Green's function of the Laplacian, so that Step 1 gives $\eta_h(t) = b_h - E(t-t_0)b_h$ with $b_h(\mathbf{x}) := G(\mathbf{x},\mathbf{x}_0) - |K_0|^{-1}\int_{K_0} G(\mathbf{x},\mathbf{y})\, d\mathbf{y}$, where $E(s)$ denotes the Dirichlet heat semigroup. For $s = t-t_0 \geq \tau$, we have $\|E(s)b_h\|_{L^\infty(\Omega)} \leq C(\tau)\|b_h\|_{L^1(\Omega)} \leq C$, and on $B_{2h}(\mathbf{x}_0)$ we have $|b_h(\mathbf{x})| \leq C(1+|\log(|\mathbf{x}-\mathbf{x}_0|/h)|)$ for $d=2$ (the constant $\frac{1}{2\pi}|\log h|$ cancels in b_h, as in (19)) and $|b_h(\mathbf{x})| \leq C(|\mathbf{x}-\mathbf{x}_0|^{-1}+h^{-1})$ for $d=3$. Hence the near field contributes at most Ch for $d=2$ and $Ch^{1/2}$ for $d=3$ to the L^2 norm, and $O(h^2)$ to the L^1 norm, so that the bounds of Steps 3 and 4 hold on all of Ω.

Step 5: discretization error. The function $u^{(h)}$ solves a heat equation with a bounded source, but this bound degenerates: $\|\mu_h\|_{L^\infty(\Omega)} = |K_0|^{-1} \sim h^{-d}$, and the higher derivatives of $u^{(h)}$ blow up near K_0 as $h \to 0$; for instance, $\|u^{(h)}(t)\|_{H^2(\Omega)} \sim \|\mu_h\|_{L^2(\Omega)} \sim h^{-d/2}$. Standard consistency estimates of the form $\|\xi_h(t)\|_{L^2(\Omega)} \leq Ch \sup_{s\leq t}\|u^{(h)}(s)\|_{H^2(\Omega)}$ would therefore only give the useless bound $O(h^{1-d/2})$. We avoid derivatives of $u^{(h)}$ near K_0 altogether. Since $\xi_h = (u-u_h)-\eta_h$ and η_h has been estimated in Steps 3 and 4, it suffices to estimate $u-u_h$.

Let $u_\infty := G(\cdot,\mathbf{x}_0)$ be the solution of the Poisson problem (12), and let $u_{\infty,h}$ be its finite volume approximation, $-\Delta_h u_{\infty,h} = \mu_h$; recall from Subsection 2.2 that the

finite volume source vectors of $\delta_{\mathbf{x}_0}$ and of μ_h coincide. With the semigroup $E(s)$ of Step 4 and its semidiscrete counterpart $E_h(s) := e^{s\Delta_h}$, Duhamel's formula gives for $s := t - t_0 \geq \tau$

$$u(t) = u_\infty - E(s)u_\infty, \qquad u_h(t) = u_{\infty,h} - E_h(s)u_{\infty,h},$$

and hence, with $v(t) := E(s)u_\infty$ and $\varepsilon_h(t) := P_h v(t) - E_h(s)u_{\infty,h}$,

$$u(t) - u_h(t) = (u_\infty - u_{\infty,h}) - \big(v(t) - P_h v(t)\big) - \varepsilon_h(t).$$

(a) The first term is the error of the stationary problem, $\|u_\infty - u_{\infty,h}\|_{L^2(\Omega)} \leq Ch\ell_h$ by Theorem 1. This is the only place where the singularity at $\mathbf{x}_0$ enters, and Theorem 1 controls it through the flux errors of order one at the finitely many faces near $\mathbf{x}_0$ (Lemma 2) and the stability of the dual problem (Lemma 3), not through derivatives near $\mathbf{x}_0$.

(b) For $s \geq \tau$, $v(t)$ is smooth, and $\|v(t) - P_h v(t)\|_{L^2(\Omega)} \leq Ch\|\nabla v(t)\|_{L^2(\Omega)} \leq C(\tau)h$.

(c) The function ε_h starts from $\varepsilon_h(t_0) = P_h u_\infty - u_{\infty,h}$, with $\|\varepsilon_h(t_0)\|_h \leq Ch\ell_h$ by the proof of Theorem 1, and satisfies $\varepsilon_h' - \Delta_h \varepsilon_h = \rho_h$, where, as in Lemma 1, $(\rho_h(t), w)_h = \sum_e \tau_e(t)\, D_e w$ with the flux consistency errors $\tau_e(t)$ of $v(t)$. The standard first-order flux consistency of the two-point scheme [16] gives $\|\rho_h(t)\|_{-1,h} \leq Ch\,|v(t)|_{H^2(\Omega)}$, where $\|\cdot\|_{-1,h}$ is the norm dual to $|\cdot|_{1,h}$. Since $u_\infty \in H^{1/4}(\Omega)$ (in fact, $u_\infty \in H^\beta(\Omega)$ for all $\beta < 2 - d/2$), parabolic smoothing and H^2-regularity on the convex domain Ω yield $|v(t)|_{H^2(\Omega)} \leq C\|\Delta v(t)\|_{L^2(\Omega)} \leq C(t - t_0)^{-7/8}\|u_\infty\|_{H^{1/4}(\Omega)}$. Using that $E_h(s)$ is a contraction in $\|\cdot\|_h$ and the discrete smoothing property $\|E_h(s)\rho\|_h \leq Cs^{-1/2}\|\rho\|_{-1,h}$, which follows from the spectral representation of the symmetric operator Δ_h, Duhamel's formula gives

$$\|\varepsilon_h(t)\|_h \leq \|\varepsilon_h(t_0)\|_h + Ch \int_0^s (s - \sigma)^{-1/2}\sigma^{-7/8}\, d\sigma \leq Ch\ell_h + C(\tau)h.$$

Combining (a)–(c) shows $\|u(t) - u_h(t)\|_{L^2(\Omega)} \leq Ch\ell_h$ for $t \in [t_0 + \tau, T]$, and thus $\|\xi_h(t)\|_{L^2(\Omega)} \leq \|u(t) - u_h(t)\|_{L^2(\Omega)} + \|\eta_h(t)\|_{L^2(\Omega)} \leq Ch\ell_h$, which is the asserted order. For the L^1 bound, the stationary errors in (a) and in $\varepsilon_h(t_0)$ are estimated in $L^1(\Omega)$ as in Remark 3, the propagation of $\varepsilon_h(t_0)$ is controlled since $E_h(s)$ is a contraction in the discrete L^1 norm (the matrix of $-\Delta_h$ is an M-matrix, and the scheme is conservative), and the remaining terms are bounded by Ch in $L^2(\Omega)$ and hence in $L^1(\Omega)$. Standard finite volume consistency yields optimal rates on compact subsets $K \Subset \Omega \setminus \{\mathbf{x}_0\}$, which gives the local estimate.

Combining both terms gives the result. □

Proposition 3 (Fully discrete FV error with backward Euler)**.** *Under the assumptions of the previous proposition, consider time steps $\Delta t > 0$ and $t_n = n\Delta t$.*

Let u_h^n be defined by the backward Euler finite volume scheme

$$\frac{u_h^n - u_h^{n-1}}{\Delta t} - \Delta_h u_h^n = f_h^n, \tag{61}$$

with

$$f_h^n = \begin{cases} 0, & t_n < t_0, \\ \mu_h, & t_n \geq t_0. \end{cases}$$

Then for every $\tau > 0$ there exists $C > 0$ such that for all $t_n \in [t_0 + \tau, T]$,

$$\|u(t_n) - u_h^n\|_{L^1(\Omega)} \leq C(h + \Delta t), \tag{62}$$

$$\|u(t_n) - u_h^n\|_{L^2(\Omega)} \leq \begin{cases} C(h|\log h|^{1/2} + \Delta t), & d = 2, \\ C(h^{1/2} + \Delta t), & d = 3. \end{cases} \tag{63}$$

Moreover, for any $K \Subset \Omega \setminus \{\mathbf{x}_0\}$,

$$\|P_h u(t_n) - u_h^n\|_{L^p(K)} \leq C(h^r + \Delta t),$$

where r is the spatial order of the FV scheme for smooth solutions; as in Proposition 2, the projection error $O(h)$ has to be added for u_h^n itself.

Proof. Let $U_h(t)$ denote the semidiscrete solution. Then

$$u(t_n) - u_h^n = (u - u^{(h)})(t_n) + (u^{(h)}(t_n) - U_h(t_n)) + (U_h(t_n) - u_h^n).$$

Step 1: source regularization. From the semidiscrete analysis,

$$\|u(t_n) - u^{(h)}(t_n)\|_{L^1} \lesssim h, \quad \|u(t_n) - u^{(h)}(t_n)\|_{L^2} \lesssim \begin{cases} h|\log h|^{1/2}, & d = 2, \\ h^{1/2}, & d = 3. \end{cases}$$

Step 2: spatial discretization. The term $u^{(h)} - U_h$ is estimated as in the semidiscrete case and does not exceed the same order.

Step 3: time discretization. Since U_h solves a linear parabolic problem with a source that is constant for $t \geq t_0$, standard backward Euler estimates give

$$\|U_h(t_n) - u_h^n\| \leq C\Delta t,$$

uniformly for $t_n \geq t_0 + \tau$, with C independent of h. Indeed, as in Step 5 of the previous proof, $U_h(t) = u_{\infty,h} - E_h(t - t_0)u_{\infty,h}$, and $u_h^n = u_{\infty,h} - r(-\Delta t\, \Delta_h)^k u_{\infty,h}$ with $r(\lambda) = (1 + \lambda)^{-1}$, where k is the number of time steps since the source has been switched on. The nonsmooth-data estimate $\sup_{\lambda \geq 0} |r(\lambda)^k - e^{-k\lambda}| \leq C/k$ for the backward Euler method [20] gives $\|U_h(t_n) - u_h^n\|_h \leq C\, \Delta t\, (t_n - t_0)^{-1} \|u_{\infty,h}\|_h$, up to a shift of the switching-on time by at most one step, which causes an additional error of the same order. Since $\|u_{\infty,h}\|_h \leq \|P_h u_\infty\|_h + \|P_h u_\infty - u_{\infty,h}\|_h \leq C$ uniformly in h by Theorem 1, the factor h^{-d} in μ_h does not enter the constant. □

6 Stationary and Time-Dependent Numerical Tests

Analogous to the analysis, we present numerical demonstrations for a stationary test case in Section 6.1 and a time-dependent test case in Section 6.2. All results in this section, i.e., Tables 1 and 3 and Figures 1 and 2, were computed with the vertex-centered variant of the two-point flux approximation described in Subsection 2.1, with the injection site at a grid node. By Remark 4, Theorem 1 and Propositions 2 and 3 apply to this variant.

6.1 Stationary diffusion-reaction equation with homogeneous Dirichlet boundary conditions

The stationary measure-valued test problem is given by the diffusion-reaction portion of the time-dependent problem (1)

$$-\nabla \cdot (D\,\nabla u) = \kappa\,\delta_{\mathbf{0}}(\mathbf{x}) \qquad \text{for } \mathbf{x} \in \Omega, \tag{64}$$

$$u(\mathbf{x}) = 0 \qquad \text{for } \mathbf{x} \in \partial\Omega, \tag{65}$$

on the domain $\Omega = (-50, 50)^d \subset \mathbb{R}^d$ in $d = 2, 3$ spatial dimensions. This problem models a pollutant that diffuses with diffusivity $D = 40$ and is injected into the domain at a single location. The injection is modeled as a point source by the measure-valued Dirac delta distribution $\delta_{\mathbf{0}}(\mathbf{x})$ located at the center $\mathbf{0}$ of the domain Ω with constant injection rate $\kappa = 10$. Analogous to the analysis, the boundary condition is a homogeneous Dirichlet condition.

Here we present convergence studies for this test case. For smooth forcing functions on the right-hand side of the stationary problem with an analytical solution u, the expected second-order convergence *uniformly for all spatial dimensions* of the numerical solution u_h on a discretization with mesh size h is verified by computing the true error $u - u_h$; the results are not included here since they are not novel. They do serve to confirm that the formally first-order accurate boundary discretization does not pollute the second-order convergence overall. For problems, whose true solution u is not available, it is customary to use the numerical solution on the finest available mesh as a reference solution u_{ref} in place of u in the norm of the errors. This procedure is necessary for this non-smooth test problem. The error norm is computed in Matlab by interpolating the numerical solution on the coarser mesh to the fine mesh of the reference solution before computing their difference.

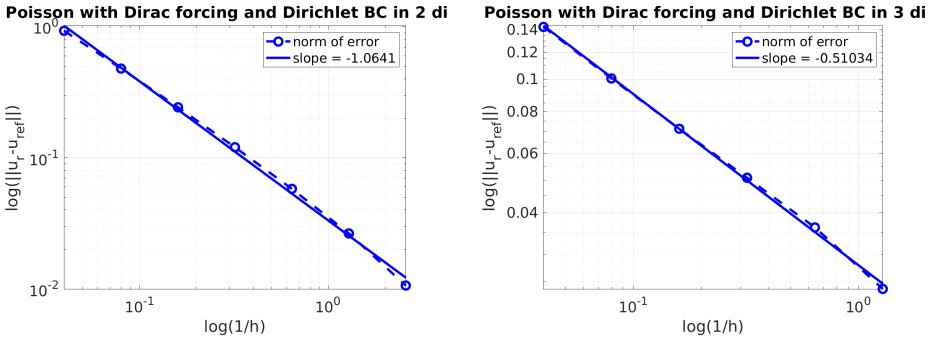
Table 1 shows, for 2-D and 3-D, the errors against the reference solution for several uniform refinement levels $r = 0, 1, 2, \ldots$. In parentheses are listed the convergence order p_{est} estimated as logarithm of the ratio between the error norms of consecutive refinement levels

$$p_{est} = \log_2 \left(\frac{\|u_{r-1} - u_{ref}\|}{\|u_r - u_{ref}\|} \right), \quad r = 1, 2, \ldots. \tag{66}$$

The result (31) in Section 3 leads us to expect $O(h|\log h|^{1/2}) \to O(h)$ as $h \to 0$ in 2-D and (32) to expect $O(h^{1/2}) = O(h^{0.5})$ in 3-D.

Table 1: Stationary diffusion-reaction equation with homogeneous Dirichlet boundary conditions and Dirac source term: observed L^2-error against reference solution (estimated convergence order) vs. spatial mesh refinement in 2-D and 3-D.

	2-D		3-D	
r	$\|u_r - u_{ref}\|$	(p_{est})	$\|u_r - u_{ref}\|$	(p_{est})
0	9.322030e-01		1.430697e-01	
1	4.799837e-01	(0.9577)	1.004787e-01	(0.5098)
2	2.430508e-01	(0.9817)	7.117283e-02	(0.4975)
3	1.205341e-01	(1.0118)	5.086844e-02	(0.4846)
4	5.801228e-02	(1.0550)	3.618990e-02	(0.4912)
5	2.649574e-02	(1.1306)	2.373690e-02	(0.6085)
6	1.061522e-02	(1.3196)	—	—



2-D 3-D

Fig. 1: Stationary diffusion-reaction equation with homogeneous Dirichlet boundary conditions and Dirac source term: observed $\log(\|u_r - u_{ref}\|)$-error against $\log(1/h)$ where the slope estimates convergence order in 2-D and 3-D.

We also show the convergence order estimate in the log-log-plot for 2-D and 3-D in Figure 1 as a linear fit of $\log(\|u_r - u_{ref}\|)$ against $\log(1/h)$, whose slope estimates the convergence order.

We observe in both Table 1 and Figure 1 that the numerical convergence order is approaching $p = 1$ for 2-D while close to $p = 0.5$ in 3-D, thus the values listed in the table and the trends shown in the figure are consistent with the results presented in Section 3 for 2-D and 3-D. Since the reference solution u_{ref} is itself a numerical solution with an error of the same structure, as r approaches the reference level, u_r and u_{ref} become increasingly correlated, so $\|u_r - u_{ref}\|$ underestimates the true error and the estimated order appears too large. This is particularly noticeable for $r = 5$ and 6 in 2-D ($p_{est} = 1.1306$ and 1.3196, respectively) and $r = 5$ in 3-D ($p_{est} = 0.6085$). The orders should therefore be judged from the coarser levels.

6.2 Time-dependent diffusion-reaction equation with homogeneous Dirichlet boundary conditions

The time-dependent measure-valued test problem is given by the diffusion-reaction equation (1) for the concentration $u(\mathbf{x}, t)$

$$u_t - \nabla \cdot (D \nabla u) = \kappa \, \delta_{\mathbf{0}}(\mathbf{x}) \qquad \text{for } \mathbf{x} \in \Omega \text{ and } 0 < t \leq t_{\text{final}}, \tag{67}$$

$$u(\mathbf{x}, t) = 0 \qquad \text{for } \mathbf{x} \in \partial\Omega \text{ and } 0 < t \leq t_{\text{final}}, \tag{68}$$

$$u(\mathbf{x}, 0) = 0 \qquad \text{for } \mathbf{x} \in \overline{\Omega} \text{ at } t = 0, \tag{69}$$

on the domain $\Omega = (-50, 50)^d \subset \mathbb{R}^d$ in $d = 2, 3$ spatial dimensions. The time-dependent test problem is inspired by the model of pollution entering at an injection rate κ as a concentrated point source for 24 hours at the center $\mathbf{0}$ of the domain Ω and diffusing with diffusivity D. The variables have the meaning and values collected in Table 2. Analogous to the analysis, the boundary condition is a homogeneous Dirichlet condition.

Table 2: Variables and their units for a pollution example.

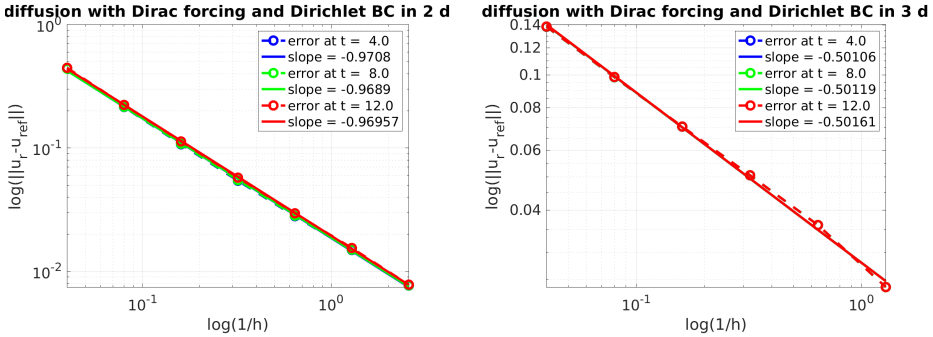
Variable	Definition	Value and/or Units
d	spatial dimension	2, 3
$\mathbf{x} \in \mathbb{R}^d$	spatial position	km
t	time variable	h
$u(\mathbf{x}, t)$	concentration of pollutant	kg/km^d
Ω	spatial domain	$(-50, 50)^d$
L	domain length	100 km
t_{final}	time domain	24 h
D	diffusivity coefficient	40 km^2/h
κ	injection rate	10 kg/h
$\delta_{\mathbf{0}}(\mathbf{x})$	Dirac delta distribution	1/km^d

Now, we present convergence results of the finite volume method for the time-dependent diffusion-reaction equation with homogeneous Dirichlet boundary conditions and Dirac source term still conforming to the semi-discretization developed in Section 2. The implicit Euler with a small time step $\Delta t = 1/128$ and the conjugate gradient method with a tight tolerance 10^{-12} are used to ensure that their errors do not obscure the spatial convergence behavior in (63) in Section 5.

Table 3 shows, for 2-D and 3-D, the errors against the reference solution for several uniform refinement levels $r = 0, 1, 2, \ldots$ at three representative points in time $t = 4, 8, 12$ during the transient simulation. In parentheses are listed the convergence order p_{est} estimated as logarithm of the ratio between the error norms of consecutive refinement levels, estimated in the same way as described in the previous Subsection 6.1. As in the stationary results, the results (60) and (63) in Section 5 lead us to expect convergence orders of $O(h|\log h|^{1/2}) \to O(h)$ as $h \to 0$ in 2-D and $O(h^{1/2}) = O(h^{0.5})$ in 3-D.

Table 3: Time-dependent diffusion-reaction equation with Dirichlet boundary conditions and one point source: observed L^2-error against reference solution (observed convergence order) vs. spatial mesh refinement at three points in time.

	$t = 4$		$t = 8$		$t = 12$	
r	$\|u_r - u_{ref}\|$	(p_{est})	$\|u_r - u_{ref}\|$	(p_{est})	$\|u_r - u_{ref}\|$	(p_{est})
	diffusion with Dirac forcing and Dirichlet BC in 2 dim					
0	4.429665e-01		4.394271e-01		4.467623e-01	
1	2.150050e-01	(1.0428)	2.160408e-01	(1.0243)	2.235983e-01	(0.9986)
2	1.072135e-01	(1.0039)	1.082159e-01	(0.9974)	1.131661e-01	(0.9825)
3	5.422757e-02	(0.9834)	5.479832e-02	(0.9817)	5.756817e-02	(0.9751)
4	2.801808e-02	(0.9527)	2.828863e-02	(0.9539)	2.959463e-02	(0.9599)
5	1.491316e-02	(0.9098)	1.501291e-02	(0.9140)	1.549514e-02	(0.9335)
6	7.685259e-03	(0.9564)	7.707547e-03	(0.9619)	7.816026e-03	(0.9873)
	diffusion with Dirac forcing and Dirichlet BC in 3 dim					
0	1.384470e-01		1.385138e-01		1.386870e-01	
1	9.838274e-02	(0.4929)	9.840875e-02	(0.4932)	9.854894e-02	(0.4929)
2	7.037154e-02	(0.4834)	7.038765e-02	(0.4835)	7.044464e-02	(0.4844)
3	5.060385e-02	(0.4757)	5.060914e-02	(0.4759)	5.062910e-02	(0.4765)
4	3.611849e-02	(0.4865)	3.611993e-02	(0.4866)	3.612553e-02	(0.4869)
5	2.372452e-02	(0.6064)	2.372477e-02	(0.6064)	2.372576e-02	(0.6066)



2-D 3-D

Fig. 2: Time-dependent diffusion-reaction equation with Dirichlet boundary conditions and one point source: observed $\log(\|u_r - u_{ref}\|)$-error against $\log(1/h)$ where the slope estimates convergence order in 2-D and 3-D.

Figure 2 shows the log-log-plot of the convergence order estimate and the linear fit of $\log(\|u_r - u_{ref}\|)$ against $\log(1/h)$, whose slope estimates the convergence order.

Both Table 3 and Figure 2 show the numerical convergence order tending to $p = 1$ for 2-D, while close to $p = 0.5$ in 3-D across all three times $t = 4, 8, 12$. The values agree with the expected values in the results of Section 5 for 2-D and 3-D. As in the stationary case, the increase of p_{est} at the finest level $r = 5$ in 3-D is due to the self-comparison bias discussed in Subsection 6.1.

Acknowledgments

We acknowledge the computational resources in the UMBC High Performance Computing Facility (hpcf.umbc.edu) and the financial contributions from NIH, NSF, CIRC, and UMBC for this work. MK is indebted to Petr Knobloch for fruitful discussions. Moreover, he acknowledges support of the CSF project 24–10366S. We warmly thank the anonymous reviewer for the very careful reading of our manuscript and for the insightful comments and suggestions, which have helped us to improve the paper considerably.

References

[1] Soane, A.M., Gobbert, M.K., Seidman, T.I.: Numerical exploration of a system of reaction-diffusion equations with internal and transient layers. Nonlinear Anal.: Real World Appl. **6**(5), 914–934 (2005)

[2] Gobbert, M.K., Churchill, A., Wang, G., Seidman, T.I.: COMSOL Multiphysics for efficient solution of a transient reaction-diffusion system with fast reaction. In: Rao, Y. (ed.) Proceedings of the COMSOL Conference 2009, Boston, MA (2009)

[3] Gobbert, M.K., Hoffman, K.A., Shen, J.: Advances in control of partial differential equations in honor of Thomas I. Seidman. IEEE Control Systems Magazine **27**(2), 92–93 (2007)

[4] Hanhart, A.L.: *Coarse-Grained Parallel Solution of a Three-Dimensional Model for Calcium Concentration in Human Heart Cells*. M.S. Thesis, Department of Mathematics and Statistics, University of Maryland, Baltimore County (2002)

[5] Hanhart, A.L., Gobbert, M.K., Izu, L.T.: A memory-efficient finite element method for systems of reaction-diffusion equations with non-smooth forcing. J. Comput. Appl. Math. **169**(2), 431–458 (2004)

[6] Gobbert, M.K.: Long-time simulations on high resolution meshes to model calcium waves in a heart cell. SIAM J. Sci. Comput. **30**(6), 2922–2947 (2008)

[7] Trott, D.W., Gobbert, M.K.: Finite element convergence for time-dependent PDEs with a point source in COMSOL 4.2. In: Rao, Y. (ed.) Proceedings of the COMSOL Conference 2011, Boston, MA (2011)

[8] Seidman, T.I., Gobbert, M.K., Trott, D.W., Kružík, M.: Finite element approximation for time-dependent diffusion with measure-valued source. Numer. Math. **122**(4), 709–723 (2012)

[9] Schäfer, J., Huang, X., Kopecz, S., Birken, P., Gobbert, M.K., Meister, A.: A memory-efficient finite volume method for advection-diffusion-reaction systems with non-smooth sources. Numer. Methods Partial Differential Equations **31**(1), 143–167 (2015)

[10] Petitta, F.: Renormalized solutions of nonlinear parabolic equations with general measure data. Annali di Matematica Pura ed Applicata **187**(4), 563–604 (2008) https://doi.org/10.1007/s10231-007-0050-9

[11] Gallouët, T., Larcher, A., Latché, J.-C.: Convergence of a finite volume scheme for the convection–diffusion equation with L^1 data. Mathematics of Computation **81**(279), 1429–1454 (2012) https://doi.org/10.1090/S0025-5718-2012-02564-3

[12] Guibé, O., Leclavier, S.: Finite volume scheme and renormalized solutions for a non-coercive and nonlinear parabolic problem with L^1 data. Journal of Mathematical Analysis and Applications **475**(2), 1321–1349 (2019) https://doi.org/10.1016/j.jmaa.2019.03.025

[13] Droniou, J., Gallouët, T., Herbin, R.: A finite volume scheme for a noncoercive elliptic equation with measure data. SIAM Journal on Numerical Analysis **41**(6), 1997–2031 (2003) https://doi.org/10.1137/S0036142902405205

[14] Kružík, M., Roubíček, T.: Mathematical Methods in Continuum Mechanics of Solids. Springer, Cham, Switzerland (2019)

[15] Barrenechea, G.R., John, V., Knobloch, P.: Monotone Discretizations for Elliptic Second Order Partial Differential Equations vol. 61. Springer, Cham, Switzerland (2025)

[16] Eymard, R., Gallouët, T., Herbin, R.: Finite volume methods. In: Ciarlet, P.G., Lions, J.-L. (eds.) Handbook of Numerical Analysis, Vol. VII, pp. 713–1020. North-Holland, Amsterdam (2000)

[17] Higham, N.J.: Accuracy and Stability of Numerical Algorithms, 2nd edn. SIAM, Philadelphia (2002)

[18] Bramble, J.H.: A second order finite difference analog of the first biharmonic boundary value problem. Numerische Mathematik **9**, 236–249 (1966)

[19] Scott, R.: Finite element convergence for singular data. Numerische Mathematik **21**, 317–327 (1973)

[20] Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems, 2nd edn. Springer Series in Computational Mathematics, vol. 25. Springer, Berlin (2006)