

SMAI-JCM
SMAI JOURNAL OF
COMPUTATIONAL MATHEMATICS

On the efficiency of a posteriori
error estimators for parabolic
partial differential equations in the
energy norm

IAIN SMEARS

Volume 12 (2026), p. 27-42.

<https://doi.org/10.5802/jcm.142>

© The authors, 2026.



*The SMAI Journal of Computational Mathematics is a member
of the Centre Mersenne for Open Scientific Publishing*

<http://www.centre-mersenne.org/>

Submissions at <https://smai-jcm.centre-mersenne.org/ojs/submission>

e-ISSN: 2426-8399





On the efficiency of a posteriori error estimators for parabolic partial differential equations in the energy norm

IAIN SMEARS¹

¹ Department of Mathematics, University College London, Gower Street, London, WC1E 6BT, United Kingdom
E-mail address: i.smears@ucl.ac.uk.

Abstract. For the model problem of the heat equation discretized by an implicit Euler method in time and a conforming finite element method in space, we prove the efficiency of *a posteriori* error estimators with respect to the energy norm of the error, when considering the numerical solution as the average between the usual continuous piecewise affine-in-time and piecewise constant-in-time reconstructions. This illustrates how the efficiency of the estimators is not only possibly dependent on the choice of norm, but also on the choice of notion of numerical solution.

2020 Mathematics Subject Classification. 65M15, 65M60.

Keywords. A posteriori error analysis, parabolic partial differential equations, energy norm.

1. Introduction

The numerical approximation of partial differential equations (PDE) is crucial in many scientific and engineering fields. Ensuring the accuracy of the approximate solutions requires robust error control by *a posteriori* error estimators, which provide computable bounds on the discretization error without requiring knowledge of the exact solution. For time-dependent problems, such as parabolic PDE, there is a wide range of challenges that are not typically encountered for their steady-state counterparts. As a model problem, we consider the heat equation

$$\begin{aligned}\partial_t u - \Delta u &= f && \text{in } \Omega \times (0, T), \\ u &= 0 && \text{on } \partial\Omega \times (0, T), \\ u(0) &= u_0 && \text{in } \Omega,\end{aligned}\tag{1.1}$$

where Ω is a bounded, open polytopal domain in \mathbb{R}^d , $d \geq 1$, with Lipschitz boundary $\partial\Omega$, and where $T > 0$ is the final time. We assume in the following that $f \in L^2(0, T; H^{-1}(\Omega))$ and that $u_0 \in L^2(\Omega)$. We will consider below, for simplicity, an implicit Euler discretization in time with a conforming finite element approximation in space. One challenge that appears in the context of time dependent problems is that there is a plethora of norms that can be considered to measure the error, and the analytical properties of error estimators may vary depending on the choice of norm. Early works [6, 7, 8, 15], as well as some more recent ones [4, 17, 24], treat various norms such as $L^\infty(L^2)$ and $L^\infty(L^\infty)$, see also [27] regarding the $L^2(L^2)$ norm. Verfürth [28] and Picasso [21] obtained some of the first results on the efficiency of the estimators in the context of the $L^2(H^1)$ norm, under various restrictions on the relationship between the relative sizes of elements of the spatial mesh and the time-steps. In [28], the efficiency of the estimators was shown under a condition of the form $\tau \simeq h^2$, where h is the spatial mesh size and τ is the time-step size. In [21], it was shown that the spatial estimator can be bounded by the $L^2(H^1)$ -norm of the error *plus the temporal jump estimator*, under the condition

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/Y008758/1].

<https://doi.org/10.5802/jcm.142>

© The authors, 2026

$\tau \simeq h$. However, it is not yet known if the additional term of the temporal jump estimator can be removed from the right-hand side of these bounds. Later in [29], it was shown that the same estimators of [28] also provide, up to data oscillation terms, upper bounds on the $L^2(H^1) \cap H^1(H^{-1})$ norm of the error, and moreover that the estimators are then efficient, locally-in-time yet globally-in-space, without any restrictions between mesh and time-step sizes. Equilibrated flux estimators were also analysed for this functional setting in [11], covering a wide range of spatial discretization schemes. We also note that various alternative error quantifiers (also called error measures in the literature) have been considered, e.g. of energy norm type [2, 19, 23], for semi-discrete approximations; these quantifiers typically combine simultaneously several reconstructions of the numerical solution, which we shall discuss further below. More recently, a different functional setting was developed in [9], based on the extension of the $L^2(H^1) \cap H^1(H^{-1})$ norm to the nonconforming-in-time approximation space, and equilibrated flux estimators were developed that give guaranteed upper bounds on the error, as well as showing efficiency bounds that are local in both space and time. Then, in [10], it was shown that the equilibrated flux estimators are also efficient in the $L^2(H^1)$ -norm plus the temporal jump estimator under the improved one-sided condition $h^2 \lesssim \tau$, which is the practically relevant case for computations.

Overall, it is clear from these references that the choice of norm bears a significant influence on the question of efficiency of the estimators. Yet, as was observed in [9], the analysis of efficiency of the estimators is also influenced by another issue, which is rather more subtle and easily-overlooked, namely the matter of the choice of *notion of numerical solution*. Indeed, for many time-stepping methods, such as the implicit Euler method, there is an element of choice in how to reconstruct a discrete function that extends the computed values at the time-step points to the whole space-time domain. For instance, [21, 29] consider the numerical solution as a continuous piecewise affine-in-time function that interpolates the values at the time-step points; see also generalizations to higher-order temporal discretizations in [19], whereas [9] considered the solution as generally discontinuous-in-time function, in particular a piecewise constant-in-time approximation in the lowest-order case of an implicit Euler discretization. In this work, we aim to investigate further the intricate relationship between the estimators, norms and also possible reconstructions. We shall consider in particular the energy norm of the error

$$\|u - \tilde{u}_{h,\tau}\|_E^2 := \frac{1}{2} \|u(T) - \tilde{u}_{h,\tau}(T)\|_\Omega^2 + \int_0^T \|\nabla(u - \tilde{u}_{h,\tau})\|_\Omega^2 dt, \quad (1.2)$$

where $\|\cdot\|_\Omega$ denotes the usual norm on $L^2(\Omega)$ or $L^2(\Omega; \mathbb{R}^d)$, depending on the argument, and where the numerical approximation $\tilde{u}_{h,\tau}$ is some chosen reconstruction with regards to the temporal approximation. To motivate the ideas in this work, let us examine more closely the role of the choice of reconstruction by considering momentarily the semi-discrete setting.

Role of the reconstruction in the semi-discrete setting. Let us consider, in a first instance, the semi-discrete approximation of (1.1), using an implicit Euler discretization in time, but without spatial discretization. Let $\{t_n\}_{n=0}^N$, for some integer $N \geq 1$, denote a strictly increasing sequence of time-step points, with $t_0 = 0$ and $t_N = T$. For each $n \in \{1, \dots, N\}$, let $I_n := (t_{n-1}, t_n)$ denote the corresponding time-step interval, and let $\tau_n = t_n - t_{n-1} > 0$ denote the corresponding step length. For the sake of simplicity in this introduction, let us momentarily assume that the source term f is piecewise constant-in-time with respect to the time-step partition and that the initial datum $u_0 \in H_0^1(\Omega)$. For each $n \in \{1, \dots, N\}$, let $u_n \in H_0^1(\Omega)$ be defined inductively by

$$\left(\frac{u_n - u_{n-1}}{\tau_n}, v \right)_\Omega + (\nabla u_n, \nabla v)_\Omega = \langle f_n, v \rangle \quad \forall v \in H_0^1(\Omega), \quad (1.3)$$

where u_0 is the initial datum from (1.1), where $f_n := f|_{I_n}$, where $(\cdot, \cdot)_\Omega$, respectively $\langle \cdot, \cdot \rangle$, denotes the usual L^2 inner-product, respectively the duality pairing between $H_0^1(\Omega)$ and its dual $H^{-1}(\Omega)$, for either scalar or vector fields, depending on the arguments. Let $u_\tau: [0, T] \rightarrow H_0^1(\Omega)$ denote the unique left-continuous piecewise constant-in-time function given by $u_\tau|_{I_n} = u_n$ for all $n \in \{1, \dots, N\}$ as well as $u_\tau(0) = u_0$. Let $U_\tau: [0, T] \rightarrow H_0^1(\Omega)$ denote the continuous piecewise affine-in-time function that interpolates the $\{u_n\}_{n=0}^N$ at the time-step nodes. It is straightforward to check that $u_\tau(T) = U_\tau(T) = u_N$, and that $(\partial_t U_\tau(t), v)_\Omega + (\nabla u_\tau(t), \nabla v)_\Omega = \langle f(t), v \rangle$ for all $v \in H_0^1(\Omega)$. This identity implies that

$$\langle \partial_t(u - U_\tau)(t), v \rangle + (\nabla(u - u_\tau)(t), \nabla v) = 0 \quad \forall v \in H_0^1(\Omega), \quad (1.4)$$

for a.e. $t \in (0, T)$. It is then found that the errors $u - u_\tau$ and $u - U_\tau$ are orthogonal in the inner-product $(\cdot, \cdot)_E$ that relates to the energy norm, since

$$\begin{aligned} (u - u_\tau, u - U_\tau)_E &:= \frac{1}{2} (u - u_\tau(T), u - U_\tau(T))_\Omega + \int_0^T (\nabla(u - u_\tau), \nabla(u - U_\tau))_\Omega dt \\ &= \frac{1}{2} \|(u - U_\tau)(T)\|_\Omega^2 - \int_0^T \langle \partial_t(u - U_\tau), u - U_\tau \rangle dt \\ &= 0, \end{aligned} \quad (1.5)$$

where in passing to the second line above we used $u_\tau(T) = U_\tau(T)$ and (1.4), and in passing to the last line we used $u(0) = U_\tau(0) = u_0$. The orthogonality of the errors $u - u_\tau$ and $u - U_\tau$ implies the Pythagoras identity

$$\|u - u_\tau\|_E^2 + \|u - U_\tau\|_E^2 = \|u_\tau - U_\tau\|_E^2, \quad (1.6)$$

where, again using the fact that $u_\tau(T) = U_\tau(T)$, it is found that

$$\|u_\tau - U_\tau\|_E^2 = \int_0^T \|\nabla(u_\tau - U_\tau)\|_\Omega^2 dt = \frac{1}{3} \sum_{n=1}^N \tau_n \|\nabla(u_n - u_{n-1})\|_\Omega^2. \quad (1.7)$$

Thus the estimator $\|u_\tau - U_\tau\|_E$ is the well-known temporal jump estimator that appears in numerous analyses of *a posteriori* analysis of parabolic problems. The identity (1.6) is a parabolic analogue of the celebrated Prager–Synge identity [22], which plays a central role in the analysis of elliptic problems. In particular, it implies immediately the bounds $\|u - u_\tau\|_E \leq \|u_\tau - U_\tau\|_E$ and $\|u - U_\tau\|_E \leq \|u_\tau - U_\tau\|_E$. However, the problem is that, in general, the temporal jump estimator $\|u_\tau - U_\tau\|_E$ is not efficient with respect to either $\|u - u_\tau\|_E$ or $\|u - U_\tau\|_E$. For the sake of completeness, we include a short example on this issue in Appendix A below. Yet, following Prager and Synge [22], it is interesting to observe that (1.6) is equivalent to

$$\|u - \bar{u}_\tau\|_E = \frac{1}{2} \|u_\tau - U_\tau\|_E, \quad \bar{u}_\tau := \frac{1}{2}(u_\tau + U_\tau), \quad (1.8)$$

which expresses the fact that, in the geometry determined by the energy norm, the three functions u , u_τ , U_τ lie on a hypercircle of radius $\frac{1}{2}\|u_\tau - U_\tau\|_E$ that is centred on \bar{u}_τ , the midpoint between u_τ and U_τ . Note that considering the centre of the hypercircle as the appropriate notion of numerical solution is an idea that was already proposed by Prager and Synge [22, p. 248]. As we show below, it turns out that this third option offers an effective generalization to the fully discrete setting.

Contributions. We show that (1.8) can be generalized appropriately to the fully discrete setting, leading to upper and lower bounds on the error when considering the numerical solution as the midpoint between the two standard reconstructions. To be clear, we are not suggesting that this notion of numerical solution needs to be used in practical computations over the alternatives, but rather that it helps to better understand the behaviour of the estimators.

Our main results are as follows. For fully discrete approximation based on an implicit Euler method in time and conforming finite element method in space, resulting in a piecewise constant-in-time

approximation $u_{h,\tau}$ and continuous piecewise affine approximation $U_{h,\tau}$, we consider error bounds for $u - \bar{u}_{h,\tau}$ where $\bar{u}_{h,\tau} = \frac{1}{2}(u_{h,\tau} + U_{h,\tau})$. We show, in Theorem 3.1 below, an upper bound of the form

$$\|u - \bar{u}_{h,\tau}\|_E^2 \leq \int_0^T \left(\frac{1}{4} \|\nabla(u_{h,\tau} - U_{h,\tau})\|_\Omega^2 + \|\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_\Omega^2 \right) dt + \text{oscillation}, \quad (1.9)$$

where $\boldsymbol{\sigma}_{h,\tau} \in L^2(0, T; H(\text{div}, \Omega))$ is a locally computable equilibrated flux that is constructed in a similar manner as in [9, 10]. Furthermore, in Theorem 3.2 below, we obtain a global lower bound of the form

$$\int_0^T \left(\frac{1}{4} \|\nabla(u_{h,\tau} - U_{h,\tau})\|_\Omega^2 + \|\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_\Omega^2 \right) dt \lesssim \|u - \bar{u}_{h,\tau}\|_E^2 + \text{oscillation}, \quad (1.10)$$

where the hidden constant is independent of the discretization and problem parameters, under two additional hypotheses which we discuss more below. However, a crucial point that we wish to clarify immediately is that the efficiency of the estimators does not derive from the fact that one is considering the term $\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}$ in the flux estimator instead of either $\boldsymbol{\sigma}_{h,\tau} + \nabla u_{h,\tau}$ or $\boldsymbol{\sigma}_{h,\tau} + \nabla U_{h,\tau}$ since all of these would result in equivalent total estimators by virtue of the triangle inequality, so efficiency also holds for all of these possible choices. The additional hypotheses used to show the efficiency bound include the assumption that the L^2 -projection operator into the finite element space is H^1 -stable, and also a one-sided condition that the local mesh-size $h^2 \lesssim \tau$ the time-step size. Note that this latter condition enables large time-steps and is relevant to practical computations. We remark also that the H^1 -stability of the L^2 -projection is known to play an important role in numerical methods for parabolic problems [25], and also that this assumption is verified for a range of mesh refinements in practice [13]. As a significant original ingredient for the analysis, we derive a framework that characterizes the energy norm in terms of an inf-sup identity for a suitable weak form of the heat equation. This allows us to overcome various difficulties that would otherwise appear when trying the more usual approach to energy norm bounds, which is based on testing the residual with the error.

For the sake of simplicity, we do not go into detail into possible extensions of these results to other settings. First, we focus here only on the case of equilibrated flux estimators, since we make use of results from [10]. However, other spatial error estimators could most likely also be considered. One can also consider other time-stepping methods, such as higher-order approximations in time, following the ideas in [9, 10]. Yet another further extension involves mesh-modification between time-steps, as analysed in the previous works [9, 10], since this plays an important role in adaptive methods [3, 5, 7, 14, 16].

2. Notation and setting

Let $H_0^1(\omega)$ denote the closure of $C_0^\infty(\omega)$ in the space $H^1(\omega)$, where $C_0^\infty(\omega)$ denotes the space of real-valued infinitely differentiable compactly supported functions on ω . Note that for ω bounded, the Poincaré inequality implies that the mapping $v \mapsto \|\nabla v\|_\omega$ defines an equivalent norm on $H_0^1(\omega)$, see [1, Corollary 6.31, p. 184]. Let $H^{-1}(\omega)$ denote the dual space of $H_0^1(\omega)$, with norm

$$\|\Phi\|_{H^{-1}(\omega)} := \sup_{v \in H_0^1(\omega) \setminus \{0\}} \frac{\langle \Phi, v \rangle_{H^{-1}(\omega) \times H_0^1(\omega)}}{\|\nabla v\|_\omega} \quad \forall \Phi \in H^{-1}(\omega), \quad (2.1)$$

where $\langle \cdot, \cdot \rangle_{H^{-1}(\omega) \times H_0^1(\omega)}$ is the duality pairing between $H^{-1}(\omega)$ and $H_0^1(\omega)$. As noted above, in the case $\omega = \Omega$, we drop the subscript and denote the duality pairing simply by $\langle \cdot, \cdot \rangle$. The space $L^2(\omega)$ can be canonically embedded into $H^{-1}(\omega)$ through $\langle w, v \rangle_{H^{-1}(\omega) \times H_0^1(\omega)} = (w, v)_\omega$ for all $v \in H_0^1(\omega)$. Thus the spaces $H_0^1(\omega)$, $L^2(\omega)$ and $H^{-1}(\omega)$ form a *Gelfand triple*, with $H_0^1(\omega) \subset L^2(\omega) \subset H^{-1}(\omega)$, where each embedding is continuous, dense and injective.

The analysis will be formulated in terms of the following function spaces. First, let

$$X := L^2(0, T; H_0^1(\Omega)). \quad (2.2)$$

Since Ω is bounded, we shall equip X with the norm $\|\cdot\|_X$ defined by

$$\|v\|_X^2 := \int_0^T \|\nabla v\|_\Omega^2 dt \quad \forall v \in X. \quad (2.3)$$

Next, let

$$Y := L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)). \quad (2.4)$$

It is known that Y is continuously embedded in $C([0, T]; L^2(\Omega))$, see [12, p. 287], and that, for any $\varphi \in Y$ and any $t \in [0, T]$, we have the identity $\|\varphi(t)\|_\Omega^2 = \|\varphi(0)\|_\Omega^2 + 2 \int_0^t \langle \partial_t \varphi, \varphi \rangle ds$, where, here and throughout the following analysis, the values of functions in Y at specific times are understood in the sense of traces in $L^2(\Omega)$. Thus, we may consider the norm $\|\cdot\|_Y$ on the space Y defined by

$$\|\varphi\|_Y^2 := \int_0^T \left(\|\partial_t \varphi\|_{H^{-1}(\Omega)}^2 + \|\nabla \varphi\|_\Omega^2 \right) dt + \|\varphi(0)\|_\Omega^2 + \|\varphi(T)\|_\Omega^2 \quad \forall \varphi \in Y. \quad (2.5)$$

We remark that $\|\cdot\|_Y$ is then invariant with respect to reversal of the time variable, i.e. the map $Y \ni \varphi \mapsto \varphi(T - \cdot)$ is an isometry under the norm $\|\cdot\|_Y$. We also let $\mathbf{Z} := X \times L^2(\Omega)$ denote the product of X with $L^2(\Omega)$. A norm on the space \mathbf{Z} is defined by

$$\|\mathbf{v}\|_{\mathbf{Z}}^2 := \|v\|_X^2 + \frac{1}{2} \|v_T\|_\Omega^2 \quad \forall \mathbf{v} = (v, v_T) \in \mathbf{Z}. \quad (2.6)$$

Note that the choice of a factor of $\frac{1}{2}$ before the term $\|v_T\|_\Omega^2$ in (2.6) is motivated by the energy norm.

Inf-sup stability and connection to the energy norm. The following Lemma gives a useful alternative formula for the norm $\|\cdot\|_Y$ that was defined in (2.5) above. See also [9, 26] for similar identities.

Lemma 2.1. *We have the identity*

$$\|\varphi\|_Y^2 = 2\|\varphi(T)\|_\Omega^2 + \int_0^T \|(\partial_t + \Delta)\varphi\|_{H^{-1}(\Omega)}^2 dt \quad \forall \varphi \in Y. \quad (2.7)$$

Note that in (2.7), $\Delta: H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ denotes the Laplacian operator understood in a distributional sense.

Proof. Let $\varphi \in Y$ be arbitrary. Let $z \in X$ be the unique function that satisfies $(\nabla z(t), \nabla v)_\Omega = \langle \partial_t \varphi(t), v \rangle$ for all $v \in H_0^1(\Omega)$ and a.e. $t \in (0, T)$. It follows that $\|\nabla z\|_\Omega^2 = \|\partial_t \varphi\|_{H^{-1}(\Omega)}^2$ and $\|\partial_t \varphi + \Delta \varphi\|_{H^{-1}(\Omega)}^2 = \|\nabla(z - \varphi)\|_\Omega^2$ a.e. in $(0, T)$. So, by expanding the square, we get

$$\begin{aligned} \int_0^T \left(\|\partial_t \varphi + \Delta \varphi\|_{H^{-1}(\Omega)}^2 \right) dt &= \int_0^T \left(\|\nabla z\|_\Omega^2 - 2(\nabla z, \nabla \varphi)_\Omega + \|\nabla \varphi\|_\Omega^2 \right) dt \\ &= \int_0^T \left(\|\nabla z\|_\Omega^2 - 2\langle \partial_t \varphi, \varphi \rangle + \|\nabla \varphi\|_\Omega^2 \right) dt \\ &= \int_0^T \left(\|\partial_t \varphi\|_{H^{-1}(\Omega)}^2 + \|\nabla \varphi\|_\Omega^2 \right) dt + \|\varphi(0)\|_\Omega^2 - \|\varphi(T)\|_\Omega^2 \\ &= \|\varphi\|_Y^2 - 2\|\varphi(T)\|_\Omega^2, \end{aligned} \quad (2.8)$$

where in the penultimate line we used the identity $\int_0^T 2\langle \partial_t \varphi, \varphi \rangle dt = \|\varphi(T)\|_\Omega^2 - \|\varphi(0)\|_\Omega^2$. \blacksquare

Define the bilinear form $B_{\mathbf{Z}} : \mathbf{Z} \times Y \rightarrow \mathbb{R}$ by

$$B_{\mathbf{Z}}(\mathbf{v}, \varphi) := (v_T, \varphi(T))_{\Omega} + \int_0^T (-\langle \partial_t \varphi, v \rangle + (\nabla \varphi, \nabla v)_{\Omega}) dt, \quad (2.9)$$

for all $\mathbf{v} = (v, v_T) \in \mathbf{Z}$ and all $\varphi \in Y$.

The following theorem states the inf-sup identities for the bilinear form $B_{\mathbf{Z}}$.

Theorem 2.2 (Inf-sup identities). *We have the identities*

$$\sup_{\varphi \in Y \setminus \{0\}} \frac{B_{\mathbf{Z}}(\mathbf{v}, \varphi)}{\|\varphi\|_Y} = \|\mathbf{v}\|_{\mathbf{Z}} \quad \forall \mathbf{v} \in \mathbf{Z}, \quad (2.10a)$$

$$\sup_{\mathbf{v} \in \mathbf{Z} \setminus \{0\}} \frac{B_{\mathbf{Z}}(\mathbf{v}, \varphi)}{\|\mathbf{v}\|_{\mathbf{Z}}} = \|\varphi\|_Y \quad \forall \varphi \in Y. \quad (2.10b)$$

Proof. We start by showing that

$$|B_{\mathbf{Z}}(\mathbf{v}, \varphi)| \leq \|\mathbf{v}\|_{\mathbf{Z}} \|\varphi\|_Y \quad \forall \mathbf{v} \in \mathbf{Z}, \forall \varphi \in Y. \quad (2.11)$$

It follows from the definition of the $H^{-1}(\Omega)$ -norm in (2.1) that

$$\begin{aligned} \left| \int_0^T (-\langle \partial_t \varphi, v \rangle + (\nabla \varphi, \nabla v)_{\Omega}) dt \right| &= \left| \int_0^T -\langle \partial_t \varphi + \Delta \varphi, v \rangle dt \right| \\ &\leq \left(\int_0^T \|\partial_t \varphi + \Delta \varphi\|_{H^{-1}(\Omega)}^2 dt \right)^{\frac{1}{2}} \|v\|_X, \end{aligned} \quad (2.12)$$

for any $\varphi \in Y$ and $v \in X$, where we recall that $\|v\|_X = \left(\int_0^T \|\nabla v\|_{\Omega}^2 dt \right)^{\frac{1}{2}}$. Hence, the Cauchy–Schwarz inequality applied to the terms in $B_{\mathbf{Z}}(\mathbf{v}, \varphi)$ shows that

$$\begin{aligned} |B_{\mathbf{Z}}(\mathbf{v}, \varphi)| &\leq \left(\frac{1}{2} \|v_T\|_{\Omega}^2 + \|v\|_X^2 \right)^{\frac{1}{2}} \left(2 \|\varphi(T)\|_{\Omega}^2 + \int_0^T \|\partial_t \varphi + \Delta \varphi\|_{H^{-1}(\Omega)}^2 dt \right)^{\frac{1}{2}} \\ &= \|\mathbf{v}\|_{\mathbf{Z}} \|\varphi\|_Y, \end{aligned} \quad (2.13)$$

where we have used Lemma 2.1 in passing to the second line above. This yields (2.11). Next, let $\mathbf{v} = (v, v_T) \in \mathbf{Z}$ be arbitrary, and let $\varphi_* \in Y$ be the unique solution of the backward parabolic problem: find $\varphi_* \in Y$ such that $\varphi_*(T) = \frac{1}{2} v_T$ and $\partial_t \varphi_* + \Delta \varphi_* = \Delta v$ in $(0, T)$. Then, it is clear that $B_{\mathbf{Z}}(\mathbf{v}, \varphi_*) = \frac{1}{2} \|v_T\|_{\Omega}^2 + \|v\|_X^2 = \|\mathbf{v}\|_{\mathbf{Z}}^2$. Furthermore, using Lemma 2.1, we have

$$\|\varphi_*\|_Y^2 = 2 \|\varphi_*(T)\|_{\Omega}^2 + \int_0^T \|\partial_t \varphi_* + \Delta \varphi_*\|_{H^{-1}(\Omega)}^2 dt = \frac{1}{2} \|v_T\|_{\Omega}^2 + \|v\|_X^2 = \|\mathbf{v}\|_{\mathbf{Z}}^2. \quad (2.14)$$

Thus we obtain (2.10a) from the above identities and the upper bound (2.11). To show (2.10b), for a given $\varphi \in Y$, we take $\mathbf{v}_* = (v, v_T)$ with $v = (-\Delta)^{-1}(-\partial_t \varphi) + \varphi \in X$, and $v_T = 2\varphi(T) \in L^2(\Omega)$, and we perform similar computations to find that $\|\mathbf{v}_*\|_{\mathbf{Z}} = \|\varphi\|_Y$ and $B_{\mathbf{Z}}(\mathbf{v}_*, \varphi) = \|\varphi\|_Y^2$. This shows (2.10b). ■

We remark that we use the terminology of inf-sup identity in reference to the inf-sup condition [20, Theorem 3.2], since Theorem 2.2 implies that

$$\inf_{\mathbf{v} \in \mathbf{Z} \setminus \{0\}} \sup_{\varphi \in Y \setminus \{0\}} \frac{B_{\mathbf{Z}}(\mathbf{v}, \varphi)}{\|\mathbf{v}\|_{\mathbf{Z}} \|\varphi\|_Y} = \sup_{\mathbf{v} \in \mathbf{Z} \setminus \{0\}} \sup_{\varphi \in Y \setminus \{0\}} \frac{B_{\mathbf{Z}}(\mathbf{v}, \varphi)}{\|\mathbf{v}\|_{\mathbf{Z}} \|\varphi\|_Y} = 1,$$

i.e. the upper and lower constants in the usual inf-sup condition both coincide and equal one.

Heat equation. Under the hypotheses that $f \in L^2(0, T; H^{-1}(\Omega))$ and $u_0 \in L^2(\Omega)$, it is well-known that (1.1) admits a unique solution $u \in Y$, see [18]. We now show how to apply Theorem 2.2 to the heat equation. After testing the heat equation with a test function $\varphi \in Y$ and integrating-by-parts, we see that the solution u solves

$$(u(T), \varphi(T))_\Omega + \int_0^T (-\langle \partial_t \varphi, u \rangle + (\nabla u, \nabla \varphi)_\Omega) dt = \int_0^T \langle f, \varphi \rangle dt + (u_0, \varphi(0))_\Omega \quad \forall \varphi \in Y. \quad (2.15)$$

Upon defining $\mathbf{u} := (u, u(T)) \in \mathbf{Z}$, it follows that (2.15) can be written equivalently as

$$B_{\mathbf{Z}}(\mathbf{u}, \varphi) = \int_0^T \langle f, \varphi \rangle dt + (u_0, \varphi(0))_\Omega \quad \forall \varphi \in Y. \quad (2.16)$$

Hence (2.10a) implies that the energy norm $\|u\|_E$ of the solution satisfies

$$\|u\|_E = \|\mathbf{u}\|_{\mathbf{Z}} = \sup_{\varphi \in Y \setminus \{0\}} \frac{\int_0^T \langle f, \varphi \rangle dt + (u_0, \varphi(0))_\Omega}{\|\varphi\|_Y}. \quad (2.17)$$

Observe that this characterization provides an alternative to the usual energy identity based on testing (1.1) with the solution u , with the advantage of characterizing the energy norm in terms of a dual norm of the data. Thus, when it comes to the *a posteriori* error analysis below, we shall obtain an equivalence between the energy norm and a (generally noncomputable) dual norm of the residual, which can be bounded by the computable *a posteriori* error estimators.

2.1. Numerical Scheme

We now consider the numerical approximation of the solution of (1.1) by a conforming finite element method on a fixed spatial mesh coupled with the implicit Euler discretization in time. Recall from above that $\{t_n\}_{n=0}^N$ denotes a strictly increasing sequence of time-step points, with $t_0 = 0$ and $t_N = T$, and that $\tau_n = t_n - t_{n-1} > 0$ denotes the corresponding step length. Let \mathcal{T} be a conforming simplicial mesh on Ω . For each element $K \in \mathcal{T}$, we let h_K denote the diameter of K . Let $p \geq 1$ denote a fixed integer which will correspond to the polynomial degree of the finite element space to be defined below. The $H_0^1(\Omega)$ -conforming finite element space of order p is denoted by V_h , and is defined by

$$V_h := \{v \in H_0^1(\Omega), v|_K \in \mathcal{P}_p(K) \quad \forall K \in \mathcal{T}\}, \quad (2.18)$$

where $\mathcal{P}_p(K)$ denotes the space of real-valued polynomials of total degree at most p on K .

The analysis below will make use of the constant determining the H^1 -stability of the L^2 -orthogonal projection operator $\Pi_h: H_0^1(\Omega) \rightarrow V_h$. Let C_Π denote the optimal constant such that

$$\|\nabla \Pi_h v\|_\Omega \leq C_\Pi \|\nabla v\|_\Omega \quad \forall v \in H_0^1(\Omega). \quad (2.19)$$

The fact that a constant C_Π exists is simply a consequence of the finite dimensionality of V_h and the Poincaré inequality on Ω . It is known that the constant C_Π is independent of the mesh-size for a wide range of meshes used in practice, see [13] and the references therein. Note that (2.19) implies the following bound:

$$\|w_h\|_{H^{-1}(\Omega)} \leq C_\Pi \|w_h\|_{V_h^*} \quad \forall w_h \in V_h, \quad (2.20)$$

where $\|w_h\|_{V_h^*} := \sup_{v_h \in V_h \setminus \{0\}} \frac{(w_h, v_h)_\Omega}{\|\nabla v_h\|_\Omega}$ is the discrete dual norm. Conversely, we have $\|w_h\|_{V_h^*} \leq \|w_h\|_{H^{-1}(\Omega)}$ trivially, so necessarily $C_\Pi \geq 1$.

In the following, the notation for inequalities $a \lesssim b$ means $a \leq Cb$ for some constant C that is independent of the mesh and time-step sizes, but may depend on Ω , d , the shape-regularity parameter of \mathcal{T} , defined as $\theta_{\mathcal{T}} := \max_{K \in \mathcal{T}} \frac{h_K}{\rho_K}$, where ρ_K denotes the diameter of the largest inscribed balls in K , and also on the constant C_Π of (2.19). Thus, we are essentially supposing that C_Π is independent of the mesh size.

Numerical method. The numerical method consists of solving the following sequence of discrete problems: for each $n \in \{1, \dots, N\}$, find $u_{h,\tau,n} \in V_h$ such that

$$\left(\frac{u_{h,\tau,n} - u_{h,\tau,n-1}}{\tau_n}, v_h \right)_\Omega + (\nabla u_{h,\tau,n}, \nabla v_h)_\Omega = (f_{h,\tau,n}, v_h)_\Omega \quad \forall v_h \in V_h, \quad (2.21)$$

where $u_{h,\tau,0} \in V_h$ is some discrete approximation of the initial datum u_0 , and where $f_{h,\tau,n} \in L^2(\Omega)$ is some discrete approximation of $f(t_n)$. We do not prescribe here a specific choice of the data approximations $u_{h,\tau,0}$ and $f_{h,\tau,n}$. Instead, we allow for rather general approximations, with the only assumptions being that $u_{h,\tau,0} \in V_h$ as stated above, and also that $f_{h,\tau,n}|_K \in \mathcal{P}_p(K)$ for each $K \in \mathcal{T}$, for each $n \in \{1, \dots, N\}$. We remark here that for many standard choices, the various data oscillation terms that appear further below can be bounded via *a priori* estimates, see for instance [10, Lemma 6.2].

Piecewise-constant-in-time reconstruction. Let $\mathbb{V}_{h,\tau}^+$ denote the space of all left-continuous functions $v: [0, T] \rightarrow V_h$ that are piecewise constant with respect to the time partition $\{I_n\}_{n=1}^N$, i.e. $v|_{I_n} \in \mathcal{P}_0(I_n; V_h)$ for each $n \in \{1, \dots, N\}$, where we recall that $I_n = (t_{n-1}, t_n)$ is the n -th time-interval. Note that a function $v \in \mathbb{V}_{h,\tau}^+$ thus has a well-defined value $v(t) \in V_h$ for each time $t \in [0, T]$. Note that from this point of view, functions in $\mathbb{V}_{h,\tau}^+$ that agree for almost all $t \in [0, T]$ are not identified. Let $u_{h,\tau} \in \mathbb{V}_{h,\tau}^+$ be the unique function that satisfies

$$u_{h,\tau}(0) = u_{h,\tau,0}, \quad u_{h,\tau}|_{I_n} = u_{h,\tau,n} \quad \forall n \in \{1, \dots, N\}. \quad (2.22)$$

Observe also that left-continuity of $u_{h,\tau}$ ensures that $u_{h,\tau}(t_n) = u_{h,\tau,n}$ for each $n \in \{0, \dots, N\}$.

Continuous piecewise-affine-in-time reconstruction. Let $U_{h,\tau} \in C([0, T]; V_h)$ be the unique continuous piecewise-affine function that satisfies

$$U_{h,\tau}(t) = \frac{t - t_{n-1}}{\tau_n} u_{h,\tau,n} + \frac{t_n - t}{\tau_n} u_{h,\tau,n-1} \quad \forall t \in [t_{n-1}, t_n]. \quad (2.23)$$

Note that, by definition, $U_{h,\tau}(t_n) = u_{h,\tau,n}$ for each $n \in \{0, \dots, N\}$.

Reformulation of the fully discrete scheme. Let the piecewise constant-in-time function $f_{h,\tau} \in L^2(0, T; L^2(\Omega))$ be defined by $f_{h,\tau}|_{I_n} := f_{h,\tau,n}$ for $n \in \{1, \dots, N\}$, where $f_{h,\tau,n}$ is the chosen approximation of $f(t_n)$ appearing in (2.21) above. Under the assumptions above on the $f_{h,\tau,n}$, it follows that $f_{h,\tau}$ is not only piecewise constant in time with respect to the time-intervals $\{(t_{n-1}, t_n)\}_{n=1}^N$, but also a piecewise polynomial of degree at most p in space with respect to the mesh \mathcal{T} . It is clear that (2.21) implies that

$$(\partial_t U_{h,\tau}(t), v_h)_\Omega + (\nabla u_{h,\tau}(t), \nabla v_h)_\Omega = (f_{h,\tau}(t), v_h)_\Omega \quad \forall v_h \in V_h, \quad \text{a.e. } t \in (0, T). \quad (2.24)$$

2.2. Construction of the equilibrated flux

We consider now *a posteriori* error bounds for the error based on the construction of an equilibrated flux. The construction follows the one in [9, 10], which handles variable polynomial degrees in hp -FEM and also mesh modification between the time-steps. Since we are restricting ourselves here to the case of lowest-order approximations in time, with a single fixed polynomial degree across the spatial mesh elements, and also a fixed mesh for all time-steps, we shall give here a slightly simplified construction relative to the one in [9, 10].

Let \mathcal{V} denote the set of vertices of the mesh \mathcal{T} . Let $\mathcal{V}_\Omega = \mathcal{V} \cap \Omega$ denote the set of interior vertices and let $\mathcal{V}_{\partial\Omega} = \mathcal{V} \cap \partial\Omega$ denote the boundary vertices. For each vertex $\mathbf{a} \in \mathcal{V}$, let $\psi_{\mathbf{a}}$ denote the corresponding hat function. Note that $\{\psi_{\mathbf{a}}\}_{\mathbf{a} \in \mathcal{V}}$ forms a partition of unity of $\bar{\Omega}$, i.e. $\sum_{\mathbf{a} \in \mathcal{V}} \psi_{\mathbf{a}}(x) = 1$ for all $x \in \bar{\Omega}$.

Let $\mathcal{T}^{\mathbf{a}}$ denote the set of all elements of \mathcal{T} that contain \mathbf{a} , and let $\omega_{\mathbf{a}}$ denote the vertex patch around \mathbf{a} , i.e. the union of all elements of $\mathcal{T}^{\mathbf{a}}$. With the convention that elements of \mathcal{T} are considered to be closed sets, it follows that $\omega_{\mathbf{a}}$ is the support of $\psi_{\mathbf{a}}$. Let $\tilde{p} \geq p + 1$ denote a fixed choice of polynomial degree that will be used for the flux reconstruction, where it is recalled that p is the polynomial degree used for V_h above. For each $\mathbf{a} \in \mathcal{V}$, let the spaces $\mathcal{P}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}})$ and $\mathbf{RTN}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}})$ be defined by

$$\mathcal{P}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}}) := \{q_h \in L^2(\omega_{\mathbf{a}}) : q_h|_K \in \mathcal{P}_{\tilde{p}}(K) \quad \forall K \in \mathcal{T}^{\mathbf{a}}\}, \quad (2.25a)$$

$$\mathbf{RTN}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}}) := \{\mathbf{v}_h \in L^2(\omega_{\mathbf{a}}; \mathbb{R}^d) : \mathbf{v}_h|_K \in \mathbf{RTN}_{\tilde{p}}(K) \quad \forall K \in \mathcal{T}^{\mathbf{a}}\}, \quad (2.25b)$$

where $\mathbf{RTN}_{\tilde{p}}(K) := \mathcal{P}_{\tilde{p}}(K; \mathbb{R}^d) + \mathbf{x}\mathcal{P}_{\tilde{p}}(K)$ denotes the RTN space of order \tilde{p} on K . In other words, $\mathcal{P}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}})$ denotes the space of scalar functions that are piecewise polynomials of degree at most \tilde{p} with respect to $\mathcal{T}^{\mathbf{a}}$, and $\mathbf{RTN}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}})$ denotes the space of vector fields that are piecewise RTN of order \tilde{p} with respect to $\mathcal{T}^{\mathbf{a}}$. Note that there are no continuity conditions across mesh elements in the definition of these spaces. We now define the local mixed finite element spaces $Q_h^{\mathbf{a}}$ and $\mathbf{V}_h^{\mathbf{a}}$ by

$$Q_h^{\mathbf{a}} := \begin{cases} \{q_h \in \mathcal{P}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}}), \int_{\omega_{\mathbf{a}}} q_h \, dx = 0\} & \text{if } \mathbf{a} \in \mathcal{V}_{\Omega}, \\ \mathcal{P}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}}) & \text{if } \mathbf{a} \in \mathcal{V}_{\partial\Omega}, \end{cases} \quad (2.26a)$$

$$\mathbf{V}_h^{\mathbf{a}} := \begin{cases} \{\mathbf{v}_h \in H(\text{div}, \omega_{\mathbf{a}}) \cap \mathbf{RTN}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}}) \quad \mathbf{v}_h \cdot \mathbf{n} = 0 \text{ on } \partial\omega_{\mathbf{a}}\} & \text{if } \mathbf{a} \in \mathcal{V}_{\Omega}, \\ \{\mathbf{v}_h \in H(\text{div}, \omega_{\mathbf{a}}) \cap \mathbf{RTN}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}}) \quad \mathbf{v}_h \cdot \mathbf{n} = 0 \text{ on } \partial\omega_{\mathbf{a}} \setminus \partial\Omega\} & \text{if } \mathbf{a} \in \mathcal{V}_{\partial\Omega}, \end{cases} \quad (2.26b)$$

where \mathbf{n} denotes the unit outward normal on $\partial\omega_{\mathbf{a}}$. For each $\mathbf{a} \in \mathcal{V}$ and each $n \in \{1, \dots, N\}$, define the piecewise polynomial function $g_{h,\tau}^{\mathbf{a},n} : \omega_{\mathbf{a}} \rightarrow \mathbb{R}$ by

$$g_{h,\tau}^{\mathbf{a},n} := \psi_{\mathbf{a}} f_{h,\tau}|_{\omega_{\mathbf{a}} \times I_n} - \psi_{\mathbf{a}} \partial_t U_{h,\tau}|_{\omega_{\mathbf{a}} \times I_n} - \nabla \psi_{\mathbf{a}} \cdot \nabla u_{h,\tau}|_{\omega_{\mathbf{a}} \times I_n}. \quad (2.27)$$

Note in particular that the choice $\tilde{p} \geq p + 1$ is used above to ensure that the terms $\psi_{\mathbf{a}} f_{h,\tau,n}$ and $\psi_{\mathbf{a}} \partial_t U_{h,\tau}$ are piecewise polynomials of degree less than or equal to \tilde{p} with respect to the spatial mesh \mathcal{T} . Therefore $g_{h,\tau}^{\mathbf{a},n} \in \mathcal{P}_{\tilde{p}}(\mathcal{T}^{\mathbf{a}})$ for each $\mathbf{a} \in \mathcal{V}$. Note also that in the case of an interior vertex $\mathbf{a} \in \mathcal{V}_{\partial\Omega}$, choosing as a test function $v_h = \psi_{\mathbf{a}}$ in (2.24) above implies that

$$\int_{\omega_{\mathbf{a}}} g_{h,\tau}^{\mathbf{a},n} \, dx = (f_{h,\tau,n}, \psi_{\mathbf{a}})_{\Omega} - (\partial_t U_{h,\tau}, \psi_{\mathbf{a}})_{\Omega} - (\nabla u_{h,\tau}, \nabla \psi_{\mathbf{a}})_{\Omega} = 0, \quad (2.28)$$

where we recall that $f_{h,\tau}|_{I_n} = f_{h,\tau,n}$ by definition. Therefore, $g_{h,\tau}^{\mathbf{a},n}$ has zero-mean value in the case of an interior vertex \mathbf{a} , and thus we conclude that $g_{h,\tau}^{\mathbf{a},n} \in Q_h^{\mathbf{a}}$ for all vertices $\mathbf{a} \in \mathcal{V}$, where it is recalled that $Q_h^{\mathbf{a}}$ was defined in (2.26a) above. Next, for each time-step $n \in \{1, \dots, N\}$ and each vertex $\mathbf{a} \in \mathcal{V}$, we define the local flux contribution $\sigma_{h,\tau}^{\mathbf{a},n} \in \mathbf{V}_h^{\mathbf{a}}$ by

$$\sigma_{h,\tau}^{\mathbf{a},n} := \arg \min_{\substack{\mathbf{v}_{h,\tau} \in \mathbf{V}_h^{\mathbf{a}} \\ \nabla \cdot \mathbf{v}_{h,\tau} = g_{h,\tau}^{\mathbf{a},n}}} \|\mathbf{v}_{h,\tau} + \psi_{\mathbf{a}} \nabla u_{h,\tau}|_{I_n}\|_{\omega_{\mathbf{a}}}. \quad (2.29)$$

Crucially, the fact that $\sigma_{h,\tau}^{\mathbf{a},n}$ is well-defined in the case of an interior vertex is a consequence of (2.28). Next, we extend the individual local contributions by zero to the whole domain Ω . Note that boundary conditions imposed on vector fields in $\mathbf{V}_h^{\mathbf{a}}$ imply that the extension by zero of $\sigma_{h,\tau}^{\mathbf{a},n}$ to the rest of Ω , with same notation, is in $H(\text{div}, \Omega)$. Therefore, we may define the global flux $\sigma_{h,\tau} \in L^2(0, T; H(\text{div}, \Omega))$ as the unique piecewise constant-in-time vector field such that

$$\sigma_{h,\tau}|_{I_n} := \sum_{\mathbf{a} \in \mathcal{V}} \sigma_{h,\tau}^{\mathbf{a},n} \quad \forall n \in \{1, \dots, N\}. \quad (2.30)$$

Then, following the same argument as in [9, Theorem 4.2], we have the equilibration identity

$$\partial_t U_{h,\tau} + \nabla \cdot \sigma_{h,\tau} = f_{h,\tau} \quad \text{a.e. in } \Omega \times (0, T). \quad (2.31)$$

3. Main results

We now give the main results on the *a posteriori* error bound for the energy norm of the error. In Theorem 3.1, we show that the energy norm of the error $u - \bar{u}_{h,\tau}$ can be bounded, up to data oscillation, in terms of computable estimators derived from $u_{h,\tau} - U_{h,\tau}$ and $\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}$. Furthermore, the resulting estimators are globally efficient, as shown in Theorem 3.2 below.

Theorem 3.1 (Upper bound). *We have the upper bound*

$$\|u - \bar{u}_{h,\tau}\|_E \leq \left(\int_0^T \left(\frac{1}{4} \|\nabla(u_{h,\tau} - U_{h,\tau})\|_\Omega^2 + \|\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_\Omega^2 \right) dt \right)^{\frac{1}{2}} + \eta_{\text{osc},h,\tau,E}, \quad (3.1)$$

where the data oscillation term $\eta_{\text{osc},h,\tau,E}$ is defined by

$$\eta_{\text{osc},h,\tau,E} := \sup_{\varphi \in Y \setminus \{0\}} \frac{\int_0^T \langle f - f_{h,\tau}, \varphi \rangle dt + (u_0 - u_{h,\tau,0}, \varphi(0))_\Omega}{\|\varphi\|_Y}. \quad (3.2)$$

Proof. Theorem 2.2 implies that

$$\|u - \bar{u}_{h,\tau}\|_E = \sup_{\varphi \in Y \setminus \{0\}} \frac{\langle \mathcal{R}(\bar{u}_{h,\tau}), \varphi \rangle_{Y^* \times Y}}{\|\varphi\|_Y}, \quad (3.3)$$

where Y^* denotes the dual space of Y and the residual $\mathcal{R}(\bar{u}_{h,\tau})$ is defined by

$$\begin{aligned} \langle \mathcal{R}(\bar{u}_{h,\tau}), \varphi \rangle_{Y^* \times Y} &:= \int_0^T \langle f, \varphi \rangle dt + (u_0, \varphi(0))_\Omega \\ &\quad - (\bar{u}_{h,\tau}(T), \varphi(T))_\Omega - \int_0^T [-\langle \partial_t \varphi, \bar{u}_{h,\tau} \rangle + (\nabla \bar{u}_{h,\tau}, \nabla \varphi)_\Omega] dt, \end{aligned} \quad (3.4)$$

for all $\varphi \in Y$. Using the flux equilibration identity (2.31), integration-by-parts, and the identities $\bar{u}_{h,\tau}(T) = U_{h,\tau}(T)$ and $U_{h,\tau}(0) = u_{h,\tau,0}$, we see that, for any $\varphi \in Y$,

$$\int_0^T (f_{h,\tau}, \varphi)_\Omega dt = (\bar{u}_{h,\tau}(T), \varphi(T))_\Omega - (u_{h,\tau,0}, \varphi(0))_\Omega + \int_0^T [-\langle \partial_t \varphi, U_{h,\tau} \rangle - (\boldsymbol{\sigma}_{h,\tau}, \nabla \varphi)_\Omega] dt. \quad (3.5)$$

Therefore, the combination of (3.4) with (3.5) implies that

$$\begin{aligned} \langle \mathcal{R}(\bar{u}_{h,\tau}), \varphi \rangle_{Y^* \times Y} &= \int_0^T \langle f - f_{h,\tau}, \varphi \rangle dt + (u_0 - u_{h,\tau,0}, \varphi(0))_\Omega \\ &\quad + \int_0^T [\langle \partial_t \varphi, \bar{u}_{h,\tau} - U_{h,\tau} \rangle - (\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}, \nabla \varphi)_\Omega] dt. \end{aligned} \quad (3.6)$$

The triangle inequality and the Cauchy–Schwarz inequality then imply that

$$\sup_{\varphi \in Y \setminus \{0\}} \frac{\langle \mathcal{R}(\bar{u}_{h,\tau}), \varphi \rangle_{Y^* \times Y}}{\|\varphi\|_Y} \leq \left(\int_0^T \left(\|\nabla(\bar{u}_{h,\tau} - U_{h,\tau})\|_\Omega^2 + \|\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_\Omega^2 \right) dt \right)^{\frac{1}{2}} + \eta_{\text{osc},h,\tau,E}, \quad (3.7)$$

which shows the upper bound (3.1) upon noting that $\bar{u}_{h,\tau} - U_{h,\tau} = \frac{1}{2}(u_{h,\tau} - U_{h,\tau})$. \blacksquare

We now state the lower bound for the error in terms of the estimators, showing global efficiency of the estimators under the additional condition of the form $h^2 \lesssim \tau$. In the following, let $h_{\omega_{\mathbf{a}}} := \text{diam } \omega_{\mathbf{a}}$ denote the diameter of the vertex patch $\omega_{\mathbf{a}}$ for each $\mathbf{a} \in \mathcal{V}$.

Theorem 3.2 (Lower bound). *If there exists a constant $\gamma > 0$ such that $h_{\omega_{\mathbf{a}}}^2 \leq \gamma \tau_n$ for all $\mathbf{a} \in \mathcal{V}$ and all $n \in \{1, \dots, N\}$, then we also have the global lower bound*

$$\int_0^T \left(\frac{1}{4} \|\nabla(u_{h,\tau} - U_{h,\tau})\|_\Omega^2 + \|\boldsymbol{\sigma}_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_\Omega^2 \right) dt \lesssim \|u - \bar{u}_{h,\tau}\|_E^2 + [\eta_{\text{osc},h,\tau,E,\star}]^2, \quad (3.8)$$

where the data oscillation term $\eta_{\text{osc},h,\tau,E,\star}$ is defined by

$$[\eta_{\text{osc},h,\tau,E,\star}]^2 := [\tilde{\eta}_{\text{osc},h,\tau,E}]^2 + \sum_{n=1}^N \sum_{\mathbf{a} \in \mathcal{V}} [\eta_{\text{osc}}^{\mathbf{a},n}]^2, \quad (3.9a)$$

$$\tilde{\eta}_{\text{osc},h,\tau,E} := \sup_{\substack{\varphi_h \in H^1(0,T;V_h) \setminus \{0\} \\ \varphi_h(0)=0}} \frac{\int_0^T \langle f - f_{h,\tau}, \varphi_h \rangle dt}{\|\varphi_h\|_Y}, \quad [\eta_{\text{osc}}^{\mathbf{a},n}]^2 := \int_{I_n} \|f - f_{h,\tau}\|_{H^{-1}(\omega_{\mathbf{a}})}^2 dt. \quad (3.9b)$$

The hidden constant in (3.8) depends only on the shape-regularity of \mathcal{T} , the dimension d , the constant γ , and on the constant C_{Π} appearing in (2.19).

The proof of Theorem 3.2 is the subject of Section 4 below. Theorems 3.1 and 3.2 show global upper and lower bounds on the error. Therefore, the error estimator gives a guaranteed upper bound on the error up to data oscillation, i.e. there are no unknown constants in the upper bound. Furthermore, the error estimators are also globally efficient. However, the local efficiency is less clear; essentially this is due to the fact that the analysis of the jump estimator $\|u_{h,\tau} - U_{h,\tau}\|_X$ hinges on some nonlocal arguments both in space and in time. Note that the locality of the efficiency bounds of the estimator $\|u_{h,\tau} - U_{h,\tau}\|_X$ is also a challenge in the setting of Y -norm error estimators for $u - U_{h,\tau}$, see [29], where the lower bounds are global in space.

Remark 3.3 (Alternative error quantifiers). Various alternative error quantifiers have appeared in the literature in order to handle the problem of efficiency of the estimators. Often these quantifiers combine the norms of the errors for the two reconstructions $u_{h,\tau}$ and $U_{h,\tau}$ simultaneously, see for instance [2, 19, 23]. In particular, local efficiency of the estimators in space and time was shown in [10] for the setting of $L^2(H^1)$ norm. In the context of the energy norm, we can consider the quantity $\mathcal{E}_{h,\tau} := \|u - u_{h,\tau}\|_E + \|u - U_{h,\tau}\|_E$, in which case the analysis above and the results in [10] imply that, up to data oscillation terms, we have some form of global equivalence between $\mathcal{E}_{h,\tau}$ and $\|u - \bar{u}_{h,\tau}\|_E$, at least under the hypotheses of Theorem 3.2. This offers a further perspective on the global nature of these alternative error quantifiers that have been considered in the literature, although the localization of these error measures might still differ.

4. Proof of the lower bound

The starting point for the proof of the lower bound is to bound the temporal jump estimator in terms of the energy norm of the error. To this end, we derive an identity for the residual when the test functions are restricted to the subspace $H^1(0,T;V_h) \subset Y$, which consists of functions that are H^1 w.r.t. time with values in V_h .

Lemma 4.1 (Identity for the difference between reconstructions). *For all $\varphi_h \in H^1(0,T;V_h)$, we have the identity*

$$\begin{aligned} & \frac{1}{2} \int_0^T [(\partial_t \varphi_h, u_{h,\tau} - U_{h,\tau})_{\Omega} + (\nabla \varphi_h, \nabla(u_{h,\tau} - U_{h,\tau}))_{\Omega}] dt \\ & = B_{\mathbf{Z}}(\mathbf{u} - \bar{\mathbf{u}}_{h,\tau}, \varphi_h) - \int_0^T \langle f - f_{h,\tau}, \varphi_h \rangle dt - (u_0 - u_{h,\tau,0}, \varphi_h(0))_{\Omega}, \end{aligned} \quad (4.1)$$

where the bilinear form $B_{\mathbf{Z}}$ is defined above in (2.9), and where $\mathbf{u}, \bar{\mathbf{u}}_{h,\tau} \in \mathbf{Z}$ are defined by $\mathbf{u} := (u, u(T))$ and $\bar{\mathbf{u}}_{h,\tau} := (\bar{u}_{h,\tau}, \bar{u}_{h,\tau}(T))$.

I. SMEARS

Proof. Note that $H^1(0, T; V_h) \subset Y$ since $V_h \subset H_0^1(\Omega)$. Using the identity $\bar{u}_{h,\tau}(t_n) = U_{h,\tau}(t_n) = u_{h,\tau,n}$ for all $n \in \{0, \dots, N\}$ and using (2.24) (which holds in a pointwise sense in time), we find that

$$\begin{aligned} (\bar{u}_{h,\tau}(T), \varphi_h(T))_\Omega + \int_0^T [-\langle \partial_t \varphi_h, U_{h,\tau} \rangle + (\nabla u_{h,\tau}, \nabla \varphi_h)_\Omega] dt \\ = (u_{h,\tau,0}, \varphi_h(0))_\Omega + \int_0^T (f_{h,\tau}, \varphi_h)_\Omega dt, \end{aligned} \quad (4.2)$$

where we have used integration-by-parts on the temporal derivative term. After addition-subtraction of the terms in (4.2) above, we find that

$$\begin{aligned} B_Z(\bar{\mathbf{u}}_{h,\tau}, \varphi_h) &= (\bar{u}_{h,\tau}(T), \varphi_h(T))_\Omega + \int_0^T [-(\partial_t \varphi_h, \bar{u}_{h,\tau})_\Omega + (\nabla \bar{u}_{h,\tau}, \nabla \varphi_h)_\Omega] dt \\ &= (u_{h,\tau,0}, \varphi_h(0))_\Omega + \int_0^T [(f_{h,\tau}, \varphi_h)_\Omega + (\partial_t \varphi_h, U_{h,\tau} - \bar{u}_{h,\tau})_\Omega + (\nabla \varphi_h, \nabla(\bar{u}_{h,\tau} - u_{h,\tau}))_\Omega] dt \\ &= (u_{h,\tau,0}, \varphi_h(0))_\Omega + \int_0^T (f_{h,\tau}, \varphi_h)_\Omega dt - \frac{1}{2} \int_0^T [(\partial_t \varphi_h, u_{h,\tau} - U_{h,\tau})_\Omega + (\nabla \varphi_h, \nabla(u_{h,\tau} - U_{h,\tau}))_\Omega] dt, \end{aligned}$$

where we have simplified $\bar{u}_{h,\tau} - U_{h,\tau} = u_{h,\tau} - \bar{u}_{h,\tau} = \frac{1}{2}(u_{h,\tau} - U_{h,\tau})$. Recalling (2.16), after subtracting the identity above from $B_Z(\mathbf{u}, \varphi_h)$, we find that

$$\begin{aligned} B_Z(\mathbf{u} - \bar{\mathbf{u}}_{h,\tau}, \varphi_h) &= (u_0 - \bar{u}_{h,\tau}(0), \varphi_h(0))_\Omega + \int_0^T \langle f - f_{h,\tau}, \varphi_h \rangle dt \\ &\quad + \frac{1}{2} \int_0^T [(\partial_t \varphi_h, u_{h,\tau} - U_{h,\tau})_\Omega + (\nabla \varphi_h, \nabla(u_{h,\tau} - U_{h,\tau}))_\Omega] dt, \end{aligned} \quad (4.3)$$

which is equivalent to (4.1). ■

We now show that the temporal jump estimator is bounded by the energy norm of the error.

Lemma 4.2 (Efficiency of the jump estimator). *We have the bound*

$$\|u_{h,\tau} - U_{h,\tau}\|_X \lesssim \|u - \bar{u}_{h,\tau}\|_E + \tilde{\eta}_{\text{osc},h,\tau,E}, \quad (4.4)$$

where the hidden constant depends only on C_Π , and where $\tilde{\eta}_{\text{osc},h,\tau,E}$ is defined in (3.9b).

Proof. Since V_h is finite dimensional, we can find a unique $\varphi_h \in H^1(0, T; V_h) \cap C([0, T]; V_h)$ that satisfies $\varphi_h(0) = 0$ and, for a.e. $t \in (0, T)$,

$$(\partial_t \varphi_h(t), v_h)_\Omega + (\nabla \varphi_h(t), \nabla v_h)_\Omega = (\nabla(u_{h,\tau} - U_{h,\tau})(t), \nabla v_h)_\Omega \quad \forall v_h \in V_h. \quad (4.5)$$

Note that $\varphi_h \in Y$ since $V_h \subset H_0^1(\Omega)$. Furthermore, we can define $z_h \in L^2(0, T; V_h)$ as the unique solution of $(\nabla z_h(t), \nabla v_h)_\Omega = (\partial_t \varphi_h(t), v_h)_\Omega$ for all $v_h \in V_h$. Then, we have $(\nabla(z_h(t) + \varphi_h(t)), \nabla v_h)_\Omega = (\nabla(u_{h,\tau} - U_{h,\tau})(t), \nabla v_h)_\Omega$ for a.e. $t \in (0, T)$, which implies that $z_h + \varphi_h = u_{h,\tau} - U_{h,\tau}$ in Ω , for a.e. $t \in (0, T)$. Therefore, after choosing $v_h = u_{h,\tau} - U_{h,\tau}$ in (4.5) and expanding the square, we obtain

$$\begin{aligned} \int_0^T \|\nabla(u_{h,\tau} - U_{h,\tau})\|_\Omega^2 dt &= \int_0^T [(\partial_t \varphi_h, u_{h,\tau} - U_{h,\tau})_\Omega + (\nabla \varphi_h, \nabla(u_{h,\tau} - U_{h,\tau}))_\Omega] dt \\ &= \int_0^T \|\nabla(z_h + \varphi_h)\|_\Omega^2 dt \\ &= \int_0^T (\|\nabla z_h\|_\Omega^2 + \|\nabla \varphi_h\|_\Omega^2) dt + 2 \int_0^T (\nabla z_h, \nabla \varphi_h)_\Omega dt \\ &= \int_0^T (\|\partial_t \varphi_h\|_{V_h^*}^2 + \|\nabla \varphi_h\|_\Omega^2) dt + 2 \int_0^T (\partial_t \varphi_h, \varphi_h)_\Omega dt \\ &= \int_0^T (\|\partial_t \varphi_h\|_{V_h^*}^2 + \|\nabla \varphi_h\|_\Omega^2) dt + \|\varphi_h(T)\|_\Omega^2, \end{aligned} \quad (4.6)$$

where, in passing to the last line above, we have used the fact that $\varphi_h(0) = 0$. It then follows from (2.20) that $\|\partial_t \varphi_h\|_{H^{-1}(\Omega)} \leq C_{\Pi} \|\partial_t \varphi_h\|_{V_h^*}$ and thus

$$\|\varphi_h\|_Y^2 = \int_0^T \left(\|\partial_t \varphi_h\|_{H^{-1}(\Omega)}^2 + \|\nabla \varphi_h\|_{\Omega}^2 \right) dt + \|\varphi_h(T)\|_{\Omega}^2 \lesssim \|u_{h,\tau} - U_{h,\tau}\|_X^2, \quad (4.7)$$

for some constant depending only on C_{Π} . Combining the identity (4.1) of Lemma 4.1 (noting that $\varphi_h(0) = 0$) along with the first line of (4.6) above, it is found that

$$\frac{1}{2} \|u_{h,\tau} - U_{h,\tau}\|_X^2 = B_{\mathbf{Z}}(\mathbf{u} - \bar{\mathbf{u}}_{h,\tau}, \varphi_h) - \int_0^T \langle f - f_{h,\tau}, \varphi_h \rangle dt. \quad (4.8)$$

The triangle inequality and Theorem 2.2 then imply that

$$\frac{1}{2} \|u_{h,\tau} - U_{h,\tau}\|_X^2 \leq (\|u - \bar{u}_{h,\tau}\|_E + \tilde{\eta}_{\text{osc},h,\tau,E}) \|\varphi_h\|_Y. \quad (4.9)$$

Thus (4.4) readily follows from the inequality above and from (4.7). \blacksquare

In order to bound the equilibrated flux estimator in terms of the error, we will use here part of [10, Theorem 5.1], which showed that the flux estimator is efficient *up to the temporal jump terms* with respect to the X -norm of the error $u - u_{h,\tau}$, under the condition $h^2 \lesssim \tau$.

Lemma 4.3 ([10]). *Suppose that there exists a constant $\gamma > 0$ such that $h_{\omega_a}^2 \leq \gamma \tau_n$ for every $\mathbf{a} \in \mathcal{V}$ and every $n \in \{1, \dots, N\}$. Then, for each $K \in \mathcal{T}$ and each $n \in \{1, \dots, N\}$, we have*

$$\int_{I_n} \|\sigma_{h,\tau} + \nabla u_{h,\tau}\|_K^2 dt \lesssim \sum_{\mathbf{a} \in \mathcal{V}_K} \left[\int_{I_n} \left(\|\nabla(u - u_{h,\tau})\|_{\omega_a}^2 + \|\nabla(u_{h,\tau} - U_{h,\tau})\|_{\omega_a}^2 \right) dt + [\eta_{\text{osc}}^{\mathbf{a},n}]^2 \right], \quad (4.10)$$

and also

$$\int_0^T \|\sigma_{h,\tau} + \nabla u_{h,\tau}\|_{\Omega}^2 dt \lesssim \|u - u_{h,\tau}\|_X^2 + \|u_{h,\tau} - U_{h,\tau}\|_X^2 + \sum_{n=1}^N \sum_{\mathbf{a} \in \mathcal{V}} [\eta_{\text{osc}}^{\mathbf{a},n}]^2, \quad (4.11)$$

where the hidden constant depends only on the shape-regularity of \mathcal{T} , the dimension d and on γ , and where the data oscillation $\eta_{\text{osc}}^{\mathbf{a},n}$ is defined in (3.9b) above.

Proof of Theorem 3.2. To prove (3.8), we start by noting that, by the triangle inequality,

$$\int_0^T \|\sigma_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_{\Omega}^2 dt \lesssim \int_0^T \left(\|\sigma_{h,\tau} + \nabla u_{h,\tau}\|_{\Omega}^2 + \|\nabla(u_{h,\tau} - U_{h,\tau})\|_{\Omega}^2 \right) dt, \quad (4.12)$$

where we also used $\bar{u}_{h,\tau} - U_{h,\tau} = \frac{1}{2}(u_{h,\tau} - U_{h,\tau})$. Applying Lemma 4.3, in particular (4.11), we then find that

$$\begin{aligned} \int_0^T \|\sigma_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_{\Omega}^2 dt &\lesssim \|u - u_{h,\tau}\|_X^2 + \|u_{h,\tau} - U_{h,\tau}\|_X^2 + \sum_{n=1}^N \sum_{\mathbf{a} \in \mathcal{V}} [\eta_{\text{osc}}^{\mathbf{a},n}]^2 \\ &\lesssim \|u - \bar{u}_{h,\tau}\|_X^2 + \|u_{h,\tau} - U_{h,\tau}\|_X^2 + \sum_{n=1}^N \sum_{\mathbf{a} \in \mathcal{V}} [\eta_{\text{osc}}^{\mathbf{a},n}]^2, \end{aligned} \quad (4.13)$$

where we have used again the triangle inequality and the identity $\bar{u}_{h,\tau} - u_{h,\tau} = \frac{1}{2}(U_{h,\tau} - u_{h,\tau})$ in passing to the second-line above. We then use Lemma 4.2 above to bound $\|u_{h,\tau} - U_{h,\tau}\|_X$, which gives

$$\int_0^T \|\sigma_{h,\tau} + \nabla \bar{u}_{h,\tau}\|_{\Omega}^2 dt + \|u_{h,\tau} - U_{h,\tau}\|_X^2 \lesssim \|u - \bar{u}_{h,\tau}\|_E^2 + [\tilde{\eta}_{\text{osc},h,\tau,E}]^2 + \sum_{n=1}^N \sum_{\mathbf{a} \in \mathcal{V}} [\eta_{\text{osc}}^{\mathbf{a},n}]^2, \quad (4.14)$$

which shows (3.8). \blacksquare

Remark 4.4 (Connection to the semi-discrete setting). To make a connection to the discussion in the introduction regarding the semi-discrete setting, it is possible to give an alternative proof of the

hypercircle theorem (1.8) using entirely analogous arguments as those above. In particular, in the semi-discrete setting with vanishing data oscillation, it is found that

$$B_Z(\mathbf{u} - \bar{\mathbf{u}}_\tau, \varphi) = \frac{1}{2} \int_0^T [\langle \partial_t \varphi, u_\tau - U_\tau \rangle + (\nabla(u_\tau - U_\tau), \nabla \varphi)_\Omega] dt \quad \forall \varphi \in Y, \quad (4.15)$$

from which (1.8) follows immediately in view of Theorem 2.2 and the invariance of $\|\cdot\|_Y$ with respect to reversal of the time direction.

Appendix A. An example

We give a brief example to illustrate how the temporal jump estimator is not generally efficient with respect to either $u - u_{h,\tau}$ or $u - U_{h,\tau}$ in the energy norm setting. Indeed, to see this, it is enough to consider the case of an ordinary differential equation: find $u: [0, 1] \rightarrow \mathbb{R}$ such that

$$\partial_t u + \lambda u = 1 \quad \text{in } (0, 1], \quad u(0) = 0,$$

where $\lambda > 0$ is a parameter. This problem can be thought of as arising from some parabolic PDE after a transformation to Fourier modes, and suitable scaling of the parameters. In this context, the energy norm is understood as $\|u\|_E^2 = \frac{1}{2}|u(1)|^2 + \lambda \int_0^1 |u(t)|^2 dt$. Note that in this case the source term $f = 1$ is constant, so there is no data oscillation. The exact solution is $u(t) = \frac{1 - e^{-\lambda t}}{\lambda}$ for all $t \in [0, 1]$. Consider its discretization by the implicit Euler method on a single time-step of length $\tau_1 = 1$. Thus we find that $u_\tau(t) = \frac{1}{1+\lambda}$ for all $t \in (0, 1]$ and $U_\tau(t) = \frac{t}{1+\lambda}$ for all $t \in [0, 1]$. Then, direct computations show that

$$\lim_{\lambda \rightarrow 0} \frac{\|u_\tau - U_\tau\|_E}{\|u - U_\tau\|_E} = \infty, \quad (\text{A.1})$$

so the estimator is not efficient with respect to the energy norm of the error $u - U_\tau$ when the parameter λ is small. It is also found that

$$\lim_{\lambda \rightarrow \infty} \frac{\|u_\tau - U_\tau\|_E}{\|u - u_\tau\|_E} = \infty, \quad (\text{A.2})$$

so the estimator is not efficient with respect to the energy norm of the error $u - u_\tau$ when the parameter λ is large. More generally, the example above shows that we cannot generally expect to show the efficiency of the estimator with regards to the energy norm error for either of these reconstructions, independently of discretization and problem parameters. In particular, the loss of efficiency occurs in cases where one reconstruction is significantly closer to the true solution than the other. Incidentally, this also shows that there is no general answer to the question of which reconstruction is the more accurate one.

References

- [1] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics*. Academic Press Inc., second edition, 2003.
- [2] Georgios Akrivis, Charalambos Makridakis, and Ricardo H. Nochetto. Optimal order a posteriori error estimates for a class of Runge-Kutta and Galerkin methods. *Numer. Math.*, 114(1):133–160, 2009.
- [3] Zhiming Chen and Jia Feng. An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comput.*, 73(247):1167–1193, 2004.
- [4] Alan Demlow, Omar Lakkis, and Charalambos Makridakis. A posteriori error estimates in the maximum norm for parabolic problems. *SIAM J. Numer. Anal.*, 47(3):2157–2176, 2009.
- [5] Todd Dupont. Mesh modification for evolution equations. *Math. Comput.*, 39(159):85–107, 1982.
- [6] Kenneth Eriksson and Claes Johnson. Error estimates and automatic time step control for nonlinear parabolic problems. I. *SIAM J. Numer. Anal.*, 24(1):12–23, 1987.

- [7] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28(1):43–77, 1991.
- [8] Kenneth Eriksson and Claes Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.*, 32(3):706–740, 1995.
- [9] Alexandre Ern, Iain Smears, and Martin Vohralík. Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems. *SIAM J. Numer. Anal.*, 55(6):2811–2834, 2017.
- [10] Alexandre Ern, Iain Smears, and Martin Vohralík. Equilibrated flux *a posteriori* error estimates in $L^2(H^1)$ -norms for high-order discretizations of parabolic problems. *IMA J. Numer. Anal.*, 39(3):1158–1179, 2019.
- [11] Alexandre Ern and Martin Vohralík. A posteriori error estimation based on potential and flux reconstruction for the heat equation. *SIAM J. Numer. Anal.*, 48(1):198–223, 2010.
- [12] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1998.
- [13] Fernando D. Gaspoz, Claus-Justus Heine, and Kunibert G. Siebert. Optimal grading of the newest vertex bisection and H^1 -stability of the L_2 -projection. *IMA J. Numer. Anal.*, 36(3):1217–1241, 2016.
- [14] Fernando D. Gaspoz, Kunibert G. Siebert, Christian Kreuzer, and Daniel A. Ziegler. A convergent time-space adaptive dG(s) finite element method for parabolic problems motivated by equal error distribution. *IMA J. Numer. Anal.*, 39(2):650–686, 2019.
- [15] Claes Johnson, Yi Yong Nie, and Vidar Thomée. An a posteriori error estimate and adaptive timestep control for a backward Euler discretization of a parabolic problem. *SIAM J. Numer. Anal.*, 27(2):277–291, 1990.
- [16] Christian Kreuzer, Christian A. Möller, Alfred Schmidt, and Kunibert G. Siebert. Design and convergence analysis for an adaptive discretization of the heat equation. *IMA J. Numer. Anal.*, 32(4):1375–1403, 2012.
- [17] Omar Lakkis, Charalambos Makridakis, and Tristan Pryer. A comparison of duality and energy a posteriori estimates for $L_\infty(0, T; L_2(\Omega))$ in parabolic problems. *Math. Comput.*, 84(294):1537–1569, 2015.
- [18] Jacques-Louis Lions and Enrico Magenes. *Non-homogeneous boundary value problems and applications. Vol. I*, volume 181 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 1972. Translated from the French by P. Kenneth.
- [19] Charalambos Makridakis and Ricardo H. Nochetto. A posteriori error analysis for higher order dissipative methods for evolution problems. *Numer. Math.*, 104(4):489–514, 2006.
- [20] Jindřich Nečas. Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (3)*, 16:305–326, 1962.
- [21] Marco Picasso. Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Eng.*, 167(3-4):223–237, 1998.
- [22] William Prager and John L. Synge. Approximations in elasticity based on the concept of function space. *Q. Appl. Math.*, 5:241–269, 1947.
- [23] Dominik Schötzau and Thomas P. Wihler. A posteriori error estimation for *hp*-version time-stepping methods for parabolic partial differential equations. *Numer. Math.*, 115(3):475–509, 2010.
- [24] Oliver J. Sutton. Long-time $L^\infty(L^2)$ *a posteriori* error estimates for fully discrete parabolic problems. *IMA J. Numer. Anal.*, 40(1):498–529, 2020.
- [25] Francesca Tantardini and Andreas Veerer. The L^2 -projection and quasi-optimality of Galerkin methods for parabolic equations. *SIAM J. Numer. Anal.*, 54(1):317–340, 2016.
- [26] Karsten Urban and Anthony T. Patera. A new error bound for reduced basis approximation of parabolic partial differential equations. *Comptes Rendus. Mathématique*, 350(3-4):203–207, 2012.

I. SMEARS

- [27] Rudiger Verfürth. A posteriori error estimates for nonlinear problems. $L^r(0, T; L^p(\Omega))$ -error estimates for finite element discretizations of parabolic equations. *Math. Comput.*, 67(224):1335–1360, 1998.
- [28] Rudiger Verfürth. A posteriori error estimates for nonlinear problems: $L^r(0, T; W^{1,p}(\Omega))$ -error estimates for finite element discretizations of parabolic equations. *Numer. Methods Partial Differ. Equations*, 14(4):487–518, 1998.
- [29] Rudiger Verfürth. A posteriori error estimates for finite element discretizations of the heat equation. *Calcolo*, 40(3):195–212, 2003.