Numer. Math. (2018) 139:505–528
https://doi.org/10.1007/s00211-018-0949-3

Numerische Mathematik

Data assimilation for the heat equation using stabilized finite element methods

Erik Burman[1] · **Lauri Oksanen**[1]

Received: 23 September 2016 / Revised: 16 November 2017 / Published online: 10 February 2018
© The Author(s) 2018. This article is an open access publication

Abstract We consider data assimilation for the heat equation using a finite element space semi-discretization. The approach is optimization based, but the design of regularization operators and parameters rely on techniques from the theory of stabilized finite elements. The space semi-discretized system is shown to admit a unique solution. Combining sharp estimates of the numerical stability of the discrete scheme and conditional stability estimates of the ill-posed continuous pde-model we then derive error estimates that reflect the approximation order of the finite element space and the stability of the continuous model. Two different data assimilation situations with different stability properties are considered to illustrate the framework. Full detail on how to adapt known stability estimates for the continuous model to work with the numerical analysis framework is given in "Appendix".

Mathematics Subject Classification 65M12 · 65M15 · 65M30 · 65M32

1 Introduction

We consider two data assimilation problems for the heat equation

$$\partial_t u - \Delta u = f, \quad \text{in } (0, T) \times \Omega, \tag{1}$$

✉ Erik Burman
e.burman@ucl.ac.uk

Lauri Oksanen
l.oksanen@ucl.ac.uk

1 Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK

where $T > 0$ and $\Omega \subset \mathbb{R}^n$ is open. Let $\omega, B \subset \Omega$ be open and let $0 < T_1 < T_2 \leq T$. Both the data assimilation problems are of the following general form: determine the restriction $u|_{(T_1,T_2)\times B}$ of a solution to the heat equation (1) given f and the restriction $u|_{(0,T)\times\omega}$. The crucial difference between the two problems is that in the first problem we do not make any assumptions on the boundary condition satisfied by u, whereas in the second problem we assume that it satisfies the lateral Dirichlet boundary condition $u|_{(0,T)\times\partial\Omega} = 0$. We emphasize that, in both the problems, no information on the initial condition satisfied by u is assumed to be known.

For the first problem we assume the following geometry,

(A1) B is connected, $\omega \subset B$, and the closure $\overline{B}$ is compact and contained in Ω, and in the second case we may choose $B = \Omega$ while assuming that
(A2) $\overline{\Omega}$ is a compact, convex, polyhedral domain.

The first problem is conditionally Hölder stable, and this stability is optimal. More precisely, the following continuum stability estimate holds.

Theorem 1 *Let $\omega \subset \Omega$ be open and non-empty, and let $0 < T_1 < T_2 < T$. Suppose that an open set $B \subset \Omega$ satisfies* (A1). *Then there are $C > 0$ and $\kappa \in (0, 1)$ such that for all u in the space*

$$H^1(0, T; H^{-1}(\Omega)) \cap L^2(0, T; H^1(\Omega)), \tag{2}$$

it holds that

$$\begin{aligned}\|u\|_{L^2(T_1,T_2;H^1(B))} &\leq C(\|u\|_{L^2((0,T)\times\omega)} \\ &\quad + \|Lu\|_{(0,-1)})^{\kappa}(\|u\|_{L^2((0,T)\times\Omega)} + \|Lu\|_{(0,-1)})^{1-\kappa},\end{aligned}$$

where $L = \partial_t - \Delta$ and $\|\cdot\|_{(0,-1)} = \|\cdot\|_{L^2(0,T;H^{-1}(\Omega))}$.

In Theorem 1, the factor with the power κ can be viewed as the size of the data

$$q = u|_{(0,T)\times\omega}, \quad f = Lu, \tag{3}$$

and the norm $\|u\|_{L^2((0,T)\times\Omega)}$ as an apriori bound for the unknown u. Let us also emphasize that the assumption $\overline{B} \subset \Omega$ is essential, indeed, if $\overline{B} \cap \partial\Omega \neq \emptyset$ then the optimal stability is conditionally logarithmic. The second problem is stable.

Theorem 2 *Let $\omega \subset \Omega$ be open and non-empty, and let $0 < T_1 < T$. Suppose that* (A2) *holds. Then there is $C > 0$ such that for all u in the space*

$$H^1(0, T; H^{-1}(\Omega)) \cap L^2(0, T; H_0^1(\Omega)), \tag{4}$$

it holds that

$$\begin{aligned}&\|u\|_{C(T_1,T;L^2(\Omega))} + \|u\|_{L^2(T_1,T;H^1(\Omega))} + \|u\|_{H^1(T_1,T;H^{-1}(\Omega))} \\ &\quad \leq C(\|u\|_{L^2((0,T)\times\omega)} + \|Lu\|_{(0,-1)}).\end{aligned}$$

The estimates in Theorems 1 and 2 are tailored to work well with the finite element discretizations that are the main topic of the present paper. We review the literature on continuum stability estimates and give the proofs bridging the gap between the existing estimates and those in Theorems 1 and 2 in an "Appendix".

In what follows, we call the two data assimilation problems the unstable and the stable problem, respectively. We use the shorthand notation

$$a(u, z) = (\nabla u, \nabla z), \quad G_f(u, z) = (\partial_t u, z) + a(u, z) - \langle f, z \rangle, \quad G = G_0,$$

where $(\cdot, \cdot)$ is the inner product of $L^2((0, T) \times \Omega)$ and $\langle \cdot, \cdot \rangle$ is the dual pairing between $L^2(0, T; H^{-1}(\Omega))$ and $L^2(0, T; H_0^1(\Omega))$. Note that for $u \in H^1((0, T) \times \Omega)$, the equations

$$G_f(u, z) = 0, \quad z \in L^2(0, T; H_0^1(\Omega)), \tag{5}$$

give the weak formulation of $\partial_t u - \Delta u = f$. Our approach to solve the data assimilation problem, in both the cases, is based on minimizing the Lagrangian functional

$$L_{q,f}(u, z) = \frac{1}{2}\|u - q\|_\omega^2 + \frac{1}{2}s(u, u) - \frac{1}{2}s^*(z, z) + G_f(u, z), \tag{6}$$

where the data q and f are fixed. Here $\|\cdot\|_\omega$ is the norm of $L^2((0, T) \times \omega)$, and s and s^* are symmetric bilinear forms, that are chosen differently in the two cases, and that we call the primal and dual stabilizers, respectively. Note that minimizing $L_{q,f}$ can be seen as fitting $u|_{(0,T)\times\omega}$ to the data q under the constraint (5), z can be interpreted as a Lagrange multiplier, and $s/2$ and $s^*/2$ as regularizing penalty terms.

In this paper we consider only semi-discretizations, that is, we minimize $L_{q,f}$ on a scale of spaces that are discrete in the spatial variable but not in the time variable. The spatial mesh size $h > 0$ will be the only parameter controlling the convergence of the approximation, and we use piecewise affine finite elements.

An important feature of the present work is that the choice of the regularizing terms is driven by the analysis and designed to give error estimates that reflect the approximation properties of the finite element space and the stability of the continuous model. In particular, when the continuous model is Lipschitz stable, we obtain optimal error estimates in the usual sense in numerical analysis, that is, the estimates are optimal compared to interpolation of the exact solution. This is the case for the second model problem introduced above. The regularization is constructed on the discrete level, that is, s and s^* are defined on the semi-discrete spaces, and in some cases, they may not even make sense on the continuous level.

We show how the different stability properties of the two model problems lead to different regularization operators. The choice for a given problem is not unique, and the design of regularizing terms leading to optimal estimates can also be driven by computational considerations, such as computational cost or couplings in the discrete formulations. Therefore we present an abstract framework for the design of regularization operators. When measurement errors are present we also show how to quantify the damping of perturbations.

For the unstable problem, under suitable choices of the semi-discrete spaces and the regularization, we show that (6) has a unique minimizer (u_h, z_h), $h > 0$, satisfying the following convergence rate

$$\|u_h - u\|_{L^2(T_1,T_2;H^1(B))} \leq Ch^\kappa(\|u\|_* + \|f\|), \tag{7}$$

where u is the unique solution of the continuum data assimilation problem, $\kappa \in (0, 1)$ is the constant in Theorem 1, and the norms on the right-hand side capture the assumed apriori smoothness, see Theorem 3 below for the precise formulation. For the stable problem, we establish an analogous result but with $\kappa = 1$, that is, the convergence rate reflects now the stability estimate in Theorem 2. Moreover, in the stable case, we can replace the norm on the left-hand side of (7) with the stronger norm in Theorem 2, see Theorem 6 below for the precise formulation.

We consider the time discretization of the stable problem in a separate paper [10]. The approach there is an extension of the strategy in the present paper: a Lagrangian of the form (6) is first discretized and then the stabilizers are chosen in the discrete spaces. The main difference is that an additional stabilization in time is required. In [10] we describe also two computational implementations of the method: one using an off-the-self linear solver, and the other using an iterative scheme that is based on the interpretation of the Euler–Lagrange equation for (6) as a system of two coupled heat equations. Let us also remark that the convergence analysis of our method is unchanged if the stabilizers in (6) are rescaled. We provide a computational study of such rescaling in [10].

1.1 Previous literature

The classical approach to data assimilation and inverse problems is to introduce Tikhonov regularization on the continuous level [2,15,26] and then discretize the well-posed regularized model. The use of Carleman estimates for uniqueness of inverse problems and the quantification of stability was first proposed in [7]. The use of conditional stability to choose a Tikhonov regularization parameter has been explored in [13]. Nevertheless a common feature of all methods where regularization takes place on the continuous level is that the accuracy of the reconstruction will be determined by the regularization error, leaving no room to exploit the interplay between discretization and regularization.

The method we propose is an instance of the so-called 4DVAR method [21]. The analysis in the present paper, which draws on previous ideas in the stationary case [8,9,11], is to our best knowledge the first complete numerical analysis of a 4DVAR type method. The main characteristic of our approach is to separate the numerical stability and the stability of the continuous model, and use the continuous stability estimate in the perturbation analysis.

In contrast, another recent approach to parabolic data assimilation problems is to derive Carleman estimates directly on the discrete scheme [3,4], which may then be used for convergence analysis. Other methods for data assimilation include the so called back and forth nudging [1] and methods using null-controllability [18,23]. In

the former case, forward and backward solves are combined with filtering techniques (or penalty) to drive the approximate solution close to data, and the latter case leads to an approximation algorithm which uses auxiliary optimal control problems. For another example of a finite element method for identification of the initial data using Tikhonov regularization see [12] and for a discussion of different approaches to numerical approximation of control and inverse problems we refer to the monograph [17, Chapter 5].

2 Spatial discretization

We begin by observing that we may assume without loss of generality that $\Omega \subset \mathbb{R}^n$ is a compact and connected polyhedral domain also in the case (A1). Indeed, we can choose a compact and connected polygonal domain $\tilde{\Omega} \subset \Omega$ such that $\overline{B} \subset \tilde{\Omega}$ and then replace Ω by $\tilde{\Omega}$.

Consider a family $\mathcal{T} = \{\mathcal{T}_h;\ h > 0\}$ of triangulations of Ω consisting of simplices such that the intersection of any two distinct simplices is either a common vertex, a common edge or a common face. Moreover, assume that the family $\mathcal{T}$ is quasi uniform, see e.g. [16, Def. 1.140], and indexed by

$$h = \max_{K \in \mathcal{T}_h} \operatorname{diam}(K).$$

The family $\mathcal{T}$ satisfies the following trace inequality, see e.g. [6, Eq. 10.3.9],

$$\|u\|_{L^2(\partial K)} \leq C \left(h^{-\frac{1}{2}} \|u\|_{L^2(K)} + h^{\frac{1}{2}} \|\nabla u\|_{L^2(K)} \right), \quad u \in H^1(K), \tag{8}$$

where the constant C is independent of $K \in \mathcal{T}_h$ and $h > 0$.

Let V_h be the H^1-conformal approximation space based on the $\mathbb{P}_1$ finite element, that is,

$$V_h = \left\{ u \in C(\bar{\Omega}) :\ u|_K \in \mathbb{P}_1(K),\ K \in \mathcal{T}_h \right\}, \tag{9}$$

where $\mathbb{P}_1(K)$ denotes the set of polynomials of degree less than or equal to 1 on K. We recall that the family $\mathcal{T}$ satisfies the following discrete inverse inequality, see e.g. [16, Lem. 1.138],

$$\|\nabla u\|_{L^2(K)} \leq C h^{-1} \|u\|_{L^2(K)}, \quad u \in \mathbb{P}_1(K), \tag{10}$$

where the constant C is independent of $K \in \mathcal{T}_h$ and $h > 0$.

We choose an interpolator

$$\pi_h : H^1(\Omega) \to V_h, \quad h > 0,$$

that satisfies the following stability and approximation properties

$$\|\pi_h u\|_{H^1(\Omega)} \leq C\|u\|_{H^1(\Omega)}, \qquad u \in H^1(\Omega),\ h > 0, \tag{11}$$

$$\|u - \pi_h u\|_{H^m(\Omega)} \leq Ch^{k-m}\|u\|_{H^k(\Omega)}, \qquad u \in H^k(\Omega),\ h > 0, \tag{12}$$

where $k = 1, 2$ and $m = 0, k-1$, and that preserves vanishing Dirichlet boundary conditions, that is, using the notation $W_h = V_h \cap H_0^1(\Omega)$,

$$\pi_h : H_0^1(\Omega) \to W_h, \quad h > 0. \tag{13}$$

A possible choice is the Scott–Zhang interpolator [25].

2.1 Spatial jump stabilizer

In the unstable case, the regularization will be based on the spatial jump stabilizer that we will introduce next. We denote by $\mathcal{F}_h$ the set of internal faces of $\mathcal{T}_h$, and define for $F \in \mathcal{F}_h$,

$$[\![n \cdot u]\!]_F = n_1 \cdot u|_{K_1} + n_2 \cdot u|_{K_2},$$

where $K_1, K_2 \in \mathcal{T}_h$ are the two simplices satisfying $K_1 \cap K_2 = F$, and n_j is the outward unit normal vector of K_j, $j = 1, 2$. We define the jump stabilizer

$$\mathcal{J}_h(u_h, u_h) = \sum_{F \in \mathcal{F}_h} \int_F h[\![n \cdot \nabla u_h]\!]_F^2 \, \mathrm{d}s, \quad u_h \in V_h, \tag{14}$$

where $\mathrm{d}s$ is the Euclidean surface measure on F.

Lemma 1 (Discrete Poincaré inequality) *There is $C > 0$ such that all $u_h \in V_h$, and $h > 0$ satisfy*

$$\|u_h\|_{L^2(\Omega)} \leq Ch^{-1}\left(\mathcal{J}_h(u_h, u_h)^{1/2} + \|u_h\|_{L^2(\omega)}\right). \tag{15}$$

Proof The original results on Poincaré inequalities for piecewise H^1-functions may be found in [5]. For a detailed proof of (15) see [11]. □

Lemma 2 *There is $C > 0$ such that all $u_h \in V_h$, $v \in H_0^1(\Omega)$, $w \in H^2(\Omega)$ and $h > 0$ satisfy*

$$(\nabla u_h, \nabla v)_{L^2(\Omega)} \leq C\mathcal{J}_h(u_h, u_h)^{1/2}\left(h^{-1}\|v\|_{L^2(\Omega)} + \|v\|_{H^1(\Omega)}\right), \tag{16}$$

$$\mathcal{J}_h(\pi_h w, \pi_h w) \leq Ch^2\|w\|_{H^2(\Omega)}^2. \tag{17}$$

Proof Towards (16) we integrate by parts and recall that u_h is an affine function on each element to obtain

$$(\nabla u_h, \nabla v)_{L^2(\Omega)} = \sum_{K \in \mathcal{T}_h} \int_K \nabla u_h \cdot \nabla v \, \mathrm{d}x = h^{-1/2} \sum_{F \in \mathcal{F}_h} \int_F h^{1/2} [\![n \cdot \nabla u_h]\!]_F \, v \, \mathrm{d}s. \tag{18}$$

Thus by the Cauchy–Schwarz inequality

$$(\nabla u_h, \nabla v)_{L^2(\Omega)} \leq h^{-1/2} \mathcal{J}_h(u_h, u_h)^{1/2} \left(\sum_{F \in \mathcal{F}_h} \int_F v^2 \mathrm{d}s \right)^{1/2},$$

and the inequality (16) follows by applying (8) to the last factor.

Let us now turn to (17). As $[\![n \cdot \nabla w]\!]_F = 0$, we have using (8)

$$\begin{aligned} \int_F h [\![n \cdot \nabla \pi_h w]\!]_F^2 \, \mathrm{d}s &= h \int_F [\![n \cdot \nabla (\pi_h w - w)]\!]_F^2 \, \mathrm{d}s \\ &\leq C \sum_{j=1,2} \left(\| n_j \cdot \nabla (\pi_h w - w) \|_{L^2(K_j)}^2 + h^2 \| n_j \cdot \nabla (\pi_h w - w) \|_{H^1(K_j)}^2 \right). \end{aligned}$$

Summing over all internal faces and observing that

$$\| \nabla (\pi_h w - w) \|_{H^1(K_j)}^2 = \| \nabla (\pi_h w - w) \|_{L^2(K_j)}^2 + \| \nabla w \|_{H^1(K_j)}^2,$$

we obtain

$$\sum_{F \in \mathcal{F}_h} \int_F h [\![n \cdot \nabla \pi_h w]\!]_F^2 \, \mathrm{d}s \leq C \left(\| \nabla (\pi_h w - w) \|_{L^2(\Omega)}^2 + h^2 \| w \|_{H^2(\Omega)}^2 \right).$$

The first term on the right-hand side satisfies

$$\| \nabla (\pi_h w - w) \|_{L^2(\Omega)}^2 \leq C h^2 \| w \|_{H^2(\Omega)}^2,$$

by (12). □

3 The unstable problem

There are several possible choices for the primal and dual stabilizers s and s^* in (6), and different choices lead to numerical methods that may differ in terms of practical performance. To illustrate the main ideas of our approach, we begin by considering a concrete choice of the stabilizers in Sect. 3.1, and give an abstract framework in Sect. 4.

In what follows, we use the shorthand notations $H^{(k,m)} = H^k(0, T; H^m(\Omega))$,

$$\|u\|_{(k,m)} = \|u\|_{H^k(0,T;H^m(\Omega))}, \quad k, m \in \mathbb{R},$$

and $H_0^{(k,m)} = H^{(k,m)} \cap L^2(0, T; H_0^1(\Omega))$. We recall also that $\|u\| = \|u\|_{(0,0)}$ and that $\|u\|_\omega$ is the norm of $L^2((0, T) \times \omega)$.

3.1 A model case

We choose the following primal and dual stabilizers

$$s_h(u_h, v_h) = \int_0^T \mathcal{J}_h(u_h, v_h)\, dt + h^2(\partial_t u_h, \partial_t v_h), \quad s_h^*(z_h, w_h) = (\nabla z_h, \nabla w_h), \tag{19}$$

defined on the respective semi-discrete spaces

$$\mathcal{V}_h = H^1(0, T; V_h), \quad \mathcal{W}_h = L^2(0, T; W_h). \tag{20}$$

We define also the semi-norm and norm

$$|u_h|_{\mathcal{V}_h} = s_h(u_h, u_h)^{1/2}, \quad \|z_h\|_{\mathcal{W}_h} = s_h^*(z_h, z_h)^{1/2}, \quad u_h \in \mathcal{V}_h,\ z_h \in \mathcal{W}_h,\ h > 0. \tag{21}$$

Recall that $W_h = V_h \cap H_0^1(\Omega)$ and whence the Poincaré inequality

$$\|z\|_{L^2(\Omega)} \le C\|\nabla z\|_{L^2(\Omega)}, \quad z \in H_0^1(\Omega), \tag{22}$$

implies that $\|\cdot\|_{\mathcal{W}_h}$ is indeed a norm on $\mathcal{W}_h$.

Lemma 3 *The semi-norm $|\cdot|_{\mathcal{V}_h}$ and norm $\|\cdot\|_{\mathcal{W}_h}$ satisfy the following inequalities with a constant $C > 0$ that is independent from $h > 0$: lower bounds*

$$G(u_h, z) \le C|u_h|_{\mathcal{V}_h}\left(h^{-1}\|z\| + \|z\|_{(0,1)}\right), \quad u_h \in \mathcal{V}_h,\ z \in H_0^{(0,1)}, \tag{23}$$

$$G(u, z_h) \le C\|z_h\|_{\mathcal{W}_h}\left(\|u\|_{(1,0)} + \|u\|_{(0,1)}\right), \quad u \in H^{(1,0)} \cap H^{(0,1)},\ z_h \in \mathcal{W}_h, \tag{24}$$

upper bounds

$$|\pi_h u|_{\mathcal{V}_h} \le Ch(\|u\|_{(1,1)} + \|u\|_{(0,2)}), \quad u \in H^{(1,1)} \cap H^{(0,2)}, \tag{25}$$

$$\|\pi_h z\|_{\mathcal{W}_h} \le C\|z\|_{(0,1)}, \quad z \in H_0^{(0,1)}, \tag{26}$$

and finally, for the semi-norm $|\cdot|_{\mathcal{V}_h}$ only, the Poincaré type inequality

$$\|u_h\| \le Ch^{-1}(|u_h|_{\mathcal{V}_h} + \|u_h\|_\omega), \quad u_h \in \mathcal{V}_h. \tag{27}$$

Proof We have

$$G(u_h, z) = (\partial_t u_h, z) + a(u_h, z) \le \|h\partial_t u_h\| h^{-1}\|z\| \\ + C\int_0^T \mathcal{J}_h(u_h, u_h)dt\,(h^{-1}\|z\| + \|z\|_1),$$

where the bound for $a(u_h, z)$ follows from (18), analogously to (16), after integration in time. This implies that (23) holds.

The lower bound (24) follows from the Cauchy–Schwarz inequality

$$(\partial_t u, z_h) + a(u, z_h) \le \|\partial_t u\|\|z_h\| + a(u, u)^{1/2} a(z_h, z_h)^{1/2},$$

together with the Poincaré inequality (22).

Towards the upper bound (25), we have

$$|\pi_h u|^2_{\mathcal{V}_h} = \int_0^T \mathcal{J}_h(\pi_h u, \pi_h u) + h^2\|\partial_t \pi_h u\|^2 \le Ch^2\|u\|^2_{(0,2)} + Ch^2\|u\|^2_{(1,1)},$$

where the bound for the first term follows from (17). The bound for the second term holds by the H^1-stability (11) of the interpolator π_h.

The upper bound (26) follows immediately from the H^1-stability of π_h, and the lower bound (27) follows from (15). □

Observe that, by the lower bound (27),

$$|||(u_h, z_h)|||_h = |u_h|_{\mathcal{V}_h} + \|u_h\|_\omega + \|z_h\|_{\mathcal{W}_h},$$

is a norm on $\mathcal{V}_h \times \mathcal{W}_h$.

Let $q \in L^2((0, T) \times \omega)$ and $f \in H^{(0,-1)}$. The Lagrangian

$$L_{q,f,h}(u_h, z_h) = \frac{1}{2}\|u_h - q\|^2_\omega + \frac{1}{2}s_h(u_h, u_h) - \frac{1}{2}s^*_h(z_h, z_h) + G_f(u_h, z_h)$$

satisfies

$$D_{u_h} L_{q,f,h}\, v_h = (u_h - q, v_h)_\omega + s_h(u_h, v_h) + G(v_h, z_h),$$
$$D_{z_h} L_{q,f,h}\, w_h = -s^*_h(z_h, w_h) + G(u_h, w_h) - \langle f, w_h\rangle,$$

and therefore the critical points $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$ of $L_{q,f,h}$ satisfy

$$A_h[(u_h, z_h), (v_h, w_h)] = (q, v_h)_\omega + \langle f, w_h\rangle, \quad (v_h, w_h) \in \mathcal{V}_h \times \mathcal{W}_h, \tag{28}$$

where A_h is the symmetric bilinear form

$$A_h[(u_h, z_h), (v_h, w_h)] \\ = (u_h, v_h)_\omega + s_h(u_h, v_h) + G(v_h, z_h) - s^*_h(z_h, w_h) + G(u_h, w_h). \tag{29}$$

Note that

$$A_h[(u_h, z_h), (u_h, -z_h)] = |u_h|^2_{\mathcal{V}_h} + \|u_h\|^2_\omega + \|z_h\|^2_{\mathcal{W}_h},$$

and therefore A_h is weakly coercive in the following sense

$$|||(u_h, z_h)|||_h \leq C \sup_{(v_h, w_h)\in \mathcal{V}_h \times \mathcal{W}_h} \frac{A_h[(u_h, z_h), (v_h, w_h)]}{|||(v_h, w_h)|||_h}, \quad (u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h. \tag{30}$$

The Babuska-Lax-Milgram theorem implies that the equation (28) has a unique solution in $\mathcal{V}_h \times \mathcal{W}_h$.

3.2 Error estimates

In this section we show that the solution (u_h, z_h) of (28) satisfies $u_h \to u$ in $(T_1, T_2)\times B$ as $h \to 0$, with the convergence rate (7). Here u is a smooth enough solution of the unstable data assimilation problem in the continuum. We use the shorthand notation

$$\|u\|_* = \|u\|_{(1,1)} + \|u\|_{(0,2)}.$$

Lemma 4 *Let $u \in H^{(1,1)} \cap H^{(0,2)}$ and define $f = \partial_t u - \Delta u$ and $q = u|_{(0,T)\times\omega}$. Let $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$ be the solution of (28). Then there exists $C > 0$ such that for all $h \in (0, 1)$,*

$$|||(u_h - \pi_h u, z_h)|||_h \leq Ch\|u\|_*.$$

Proof The equations $\partial_t u - \Delta u = f$ and $u|_\omega = q$ are equivalent with

$$\begin{aligned} G(u, w) &= \langle f, w\rangle, \quad w \in L^2(0, T; H^1_0(\Omega)), \\ (q - u, v)_\omega &= 0, \quad v \in L^2((0, T) \times \omega), \end{aligned} \tag{31}$$

and the Eqs. (28) and (31) imply for all $v_h \in \mathcal{V}_h$ and $w_h \in \mathcal{W}_h$ that

$$A_h[(u_h - \pi_h u, z_h), (v_h, w_h)] = (u - \pi_h u, v_h)_\omega + G(u - \pi_h u, w_h) - s_h(\pi_h u, v_h). \tag{32}$$

By (30) it is enough to show that

$$A_h[(u_h - \pi_h u, z_h), (v_h, w_h)] \leq Ch\|u\|_*|||(v_h, w_h)|||_h.$$

We use (12) to bound the first term in (32),

$$(u - \pi_h u, v_h)_\omega \leq \|u - \pi_h u\|_\omega \|v_h\|_\omega \leq Ch\|u\|_{(0,1)}\|v_h\|_\omega,$$

for the second term we use (24) and (12),

$$G(u-\pi_h u, w_h) \le C\left(\|u-\pi_h u\|_{(1,0)} + \|u-\pi_h u\|_{(0,1)}\right)\|w_h\|_{\mathcal{W}_h} \le Ch\|u\|_*\|w_h\|_{\mathcal{W}_h},$$

and for the third term we use (25),

$$s_h(\pi_h u, v_h) \le |\pi_h u|_{\mathcal{V}_h}|v_h|_{\mathcal{V}_h} \le Ch\|u\|_*|v_h|_{\mathcal{V}_h}.$$

□

Theorem 3 *Let $\omega, B \subset \Omega$, $0 < T_1 < T_2 < T$ and $\kappa \in (0,1)$ be as in Theorem 1. Let $u \in H^{(1,1)} \cap H^{(0,2)}$ and define $f = \partial_t u - \Delta u$ and $q = u|_{(0,T)\times\omega}$. Let $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$ be the solution of (28). Then there exists $C > 0$ such that for all $h \in (0,1)$*

$$\|u_h - u\|_{L^2(T_1,T_2;H^1(B))} \le Ch^\kappa(\|u\|_* + \|f\|).$$

Proof We define the functional $\langle r, w\rangle = G(u_h - u, w)$, $w \in H_0^{(0,1)}$. That is, r is the residual $r = (\partial_t - \Delta)(u_h - u)$ in the weak sense. The equation $\partial_t u - \Delta u = f$ implies that

$$\langle r, w\rangle = G(u_h, w) - \langle f, w\rangle.$$

By taking $v_h = 0$ in (28) we get $G(u_h, w_h) = \langle f, w_h\rangle + s_h^*(z_h, w_h)$, $w_h \in \mathcal{W}_h$, and therefore

$$\begin{aligned}\langle r, w\rangle &= G(u_h, w) - \langle f, w\rangle - G(u_h, \pi_h w) + G(u_h, \pi_h w)\\ &= G(u_h, w-\pi_h w) - \langle f, w-\pi_h w\rangle + s_h^*(z_h, \pi_h w), \quad w \in H_0^{(0,1)}.\end{aligned} \tag{33}$$

We arrive to the estimate

$$\|r\|_{(0,-1)} \le C(|u_h|_{\mathcal{V}_h} + \|z_h\|_{\mathcal{W}_h} + h\|f\|) \tag{34}$$

after we bound the first term in (33) by using (23) and (11)–(12)

$$\begin{aligned}G(u_h, w-\pi_h w) &\le C|u_h|_{\mathcal{V}_h}\left(h^{-1}\|w-\pi_h w\| + \|w-\pi_h w\|_{(0,1)}\right)\\ &\le C|u_h|_{\mathcal{V}_h}\|w\|_{(0,1)},\end{aligned} \tag{35}$$

the second term by using (12)

$$(f, w-\pi_h w) \le \|f\|\|w-\pi_h w\| \le Ch\|f\|\|w\|_{(0,1)},$$

and the last term by using (26)

$$s_h^*(z_h, \pi_h w) \leq \|z_h\|_{\mathcal{W}_h} \|\pi_h w\|_{\mathcal{W}_h} \leq C\|z_h\|_{\mathcal{W}_h} \|w\|_{(0,1)}.$$

We bound the first term in (34) by using Lemma 4 and (25)

$$|u_h|_{\mathcal{V}_h} \leq |u_h - \pi_h u|_{\mathcal{V}_h} + |\pi_h u|_{\mathcal{V}_h} \leq Ch\|u\|_*,$$

and observe that the analogous bound for the second term follows immediately from Lemma 4. Thus

$$\|r\|_{(0,-1)} \leq Ch(\|u\|_* + \|f\|).$$

By applying Theorem 1 to $u_h - u$, and using Lemma 4 and (12) to bound the term

$$\|u_h - u\|_\omega \leq \|u_h - \pi_h u\|_\omega + \|\pi_h u - u\|_\omega \leq Ch\|u\|_* + Ch\|u\|_{(0,1)}$$

we see that

$$\|u_h - u\|_{L^2(T_1,T_2;H^1(B))} \leq Ch^\kappa(\|u\|_* + \|f\|)^\kappa(\|u_h - u\| + h(\|u\|_* + \|f\|))^{1-\kappa}.$$

It remains to show the bound

$$\|u_h - u\| \leq C\|u\|_*. \tag{36}$$

The Poincaré type inequality (27) implies

$$\|u_h - \pi_h u\| \leq Ch^{-1}|u_h - \pi_h u|_{\mathcal{V}_h} + Ch^{-1}\|u_h - \pi_h u\|_\omega \leq C\|u\|_*,$$

where we used again Lemma 4. The inequality (36) follows since $\|\pi_h u - u\|$ can be bounded by $\|u\|_*$, in fact, even by $Ch^2\|u\|_*$. □

3.3 The effect of perturbations in data

In practice the data f, q in (3) are typically polluted by measurement errors. Such effects can easily be included in the analysis and we show in this section how the above results must be modified to account for this case. We consider the case where f, q in (28) are replaced by the perturbed counterparts

$$\tilde{q} = u|_{(0,T)\times\omega} + \delta q, \quad \tilde{f} = f + \delta f, \tag{37}$$

where u is the solution to the underlying unpolluted problem, that is, $\partial_t u - \Delta u = f$. We assume that the perturbations $\delta q, \delta f$ satisfy $\delta f \in H^{-1}(\Omega)$ and $\delta q \in L^2(\omega)$, and use the following measure of the size of the perturbation

$$\delta(\tilde{q}, \tilde{f}) = \|\delta q\|_{L^2(0,T;L^2(\omega))} + \|\delta f\|_{(0,-1)}.$$

The critical points $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$ of the perturbed Lagrangian $L_{\tilde{q},\tilde{f},h}$ satisfy

$$A_h[(u_h, z_h), (v_h, w_h)] = (\tilde{q}, v_h)_\omega + \langle \tilde{f}, w_h \rangle, \quad (v_h, w_h) \in \mathcal{V}_h \times \mathcal{W}_h. \tag{38}$$

We consider a perturbation of q in L^2 since this is used in the data fitting term in the Lagrangian. It should be noted that the formulation (38) remains well defined for rougher perturbations δq, but they will lead to worse estimates in h.

First we consider again the residual quantities as in Lemma 4.

Lemma 5 *Let $u \in H^{(1,1)} \cap H^{(0,2)}$ and define $\tilde{f}$ and $\tilde{q}$ by (37). Let $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$ be the solution of (38). Then there exists $C > 0$ such that for all $h \in (0, 1)$*

$$|||(u_h - \pi_h u, z_h)|||_h \leq C(h\|u\|_* + \delta(\tilde{q}, \tilde{f})).$$

Proof Proceeding as in the proof of Lemma 4, the Eqs. (38) and (31) imply for all $v_h \in \mathcal{V}_h$ and $w_h \in \mathcal{W}_h$ that

$$\begin{aligned} A_h[(u_h - \pi_h u, z_h), (v_h, w_h)] &= (u - \pi_h u, v_h)_\omega + (\delta q, v_h)_\omega \\ &+ G(u - \pi_h u, w_h) + \langle \delta f, w_h \rangle - s_h(\pi_h u, v_h). \end{aligned} \tag{39}$$

The only terms that are not covered by the previous analysis are those with the perturbations, and we have

$$(\delta q, v_h)_\omega + \langle \delta f, w_h \rangle \leq \delta(\tilde{q}, \tilde{f}) |||(v_h, w_h)|||_h$$

The choice (19) of the dual stabilizer s^* guarantees that $\|w\|_{(0,1)} \leq \|w\|_{\mathcal{W}_h}$. □

The error estimate of Theorem 3 may now be modified to include the perturbations.

Theorem 4 *Let $\omega, B \subset \Omega$, $0 < T_1 < T_2 < T$ and $\kappa \in (0, 1)$ be as in Theorem 1. Let $u \in H^{(1,1)} \cap H^{(0,2)}$ and define $\tilde{f}$ and $\tilde{q}$ by (37). Let $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$ be the solution of (38). Then there exists $C > 0$ such that for all $h \in (0, 1)$*

$$\begin{aligned} &\|u_h - u\|_{L^2(T_1,T_2;H^1(B))} \\ &\quad \leq C(h(\|u\|_* + \|f\|) + \delta(\tilde{q}, \tilde{f}))^\kappa (\|u\|_* + h\|f\| + h^{-1}\delta(\tilde{q}, \tilde{f}))^{(1-\kappa)}. \end{aligned}$$

Proof We proceed as in the proof of Theorem 3, but due to the perturbation, the Galerkin orthogonality (33) no longer holds exactly and we obtain

$$\langle r, w \rangle = G(u_h, w - \pi_h w) - \langle f, w - \pi_h w \rangle + \langle \delta f, \pi_h w \rangle + s_h^*(z_h, \pi_h w),$$

for $w \in H_0^{(0,1)}$. Analogously to the proof of Theorem 3, we obtain

$$\|r\|_{(0,-1)} \leq C(|u_h|_{\mathcal{V}_h} + \|z_h\|_{\mathcal{W}_h} + h\|f\| + \|\delta f\|_{(0,-1)}). \tag{40}$$

We bound the first term in (40) by using Lemma 5 and (25)

$$|u_h|_{\mathcal{V}_h} \leq |u_h - \pi_h u|_{\mathcal{V}_h} + |\pi_h u|_{\mathcal{V}_h} \leq C(h\|u\|_* + \delta(\tilde{q}, \tilde{f})),$$

and the second term by using again Lemma 5. Thus

$$\|r\|_{(0,-1)} \leq Ch(\|u\|_* + \|f\|) + C\delta(\tilde{q}, \tilde{f}).$$

As before, applying Theorem 1 to $u_h - u$, and using Lemma 5 and (12) to bound the term $\|u_h - u\|_\omega$ leads to

$$\begin{aligned}&\|u_h - u\|_{L^2(T_1,T_2;H^1(B)}\\ &\quad\leq C(h(\|u\|_* + \|f\|) + \delta(\tilde{q}, \tilde{f}))^\kappa(\|u_h - u\| + h(\|u\|_* + \|f\|) + \delta(\tilde{q}, \tilde{f}))^{1-\kappa}.\end{aligned}$$

It remains to bound $\|u_h - u\|$. The Poincaré type inequality (27) implies

$$\begin{aligned}\|u_h - \pi_h u\| &\leq Ch^{-1}|u_h - \pi_h u|_{\mathcal{V}_h} + Ch^{-1}\|u_h - \pi_h u\|_\omega\\ &\leq C(\|u\|_* + h^{-1}\delta(\tilde{q}, \tilde{f})),\end{aligned}$$

where we used again Lemma 5. □

According to Theorem 4, the error of the numerical approximation can diverge if no function $\tilde{u}$ exists such that $\partial_t \tilde{u} - \Delta\tilde{u} = \tilde{f}$ and $\tilde{u}|_{(0,T)\times\omega} = \tilde{q}$. On the other hand, in the context of a specific application, an estimate on the noise level might be available, and this can be used to choose a suitable mesh size. That is, assuming that there exists $h_0 > 0$ such that

$$\delta(\tilde{q}, \tilde{f}) \leq h_0(\|u\|_* + \|f\|),$$

then for $h > h_0$ there holds

$$\|u_h - u\|_{L^2(T_1,T_2;H^1(B))} \leq Ch^\kappa(\|u\|_* + \|f\|)^\kappa.$$

4 A framework for stabilization

Before proceeding to the stable problem, we introduce an abstract stabilization framework based on the essential features of the model case in Sect. 3.1.

Let s_h and s_h^* be bilinear forms on the spaces $\mathcal{V}_h$ and $\mathcal{W}_h$, respectively. Let $|\cdot|_{\mathcal{V}_h}$ be a semi-norm on $\mathcal{V}_h$ and let $\|\cdot\|_{\mathcal{W}_h}$ be a norm on $\mathcal{W}_h$. We relax (21) by requiring only that s_h and s_h^* are continuous with respect to $|\cdot|_{\mathcal{V}_h}$ and $\|\cdot\|_{\mathcal{W}_h}$, that is,

$$s_h(u_h, u_h) \leq C|u_h|^2_{\mathcal{V}_h}, \quad s_h^*(z_h, z_h) \leq C\|z_h\|^2_{\mathcal{W}_h}, \quad u_h \in \mathcal{V}_h,\ z_h \in \mathcal{W}_h,\ h > 0. \tag{41}$$

Let $\|\cdot\|_*$ be the norm of a continuously embedded subspace H^* of the energy space (2). The space H^* encodes the apriori smoothness. We relax the lower bounds (23) and (24) as follows

$$G(u_h, z - \pi_h z) \le C|u_h|_{\mathcal{V}_h}\|z\|_{(0,1)}, \qquad u_h \in \mathcal{V}_h,\ z \in H_0^{(0,1)}, \tag{42}$$

$$G(u - \pi_h u, z_h) \le Ch\|z_h\|_{\mathcal{W}_h}\|u\|_*, \qquad u \in H^*,\ z_h \in \mathcal{W}_h, \tag{43}$$

where π_h is an interpolator satisfying (11)–(13). We replace the upper bound (25) by its abstract analogue

$$|\pi_h u|_{\mathcal{V}_h} \le Ch\|u\|_*, \qquad u \in H^*, \tag{44}$$

and require that (26) and (27) hold as before. Finally, in the abstract setting, we assume that the weak coercivity (30) holds where A_h and $|||\cdot|||_h$ are defined as above.

It can be verified that the following analogue of Theorem 3 holds under these assumptions.

Theorem 5 *Let $\omega, B \subset \Omega$, $0 < T_1 < T_2 < T$ and $\kappa \in (0,1)$ be as in Theorem 1. Let $u \in H^*$ and define $f = \partial_t u - \Delta u$ and $q = u|_\omega$. Suppose that the primal and dual stabilizers satisfy (41)–(44), (26), (27) and (30). Then (28) has a unique solution $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$, and there exists $C > 0$ such that for all $h \in (0,1)$*

$$\|u_h - u\|_{L^2(T_1,T_2;H^1(B))} \le Ch^\kappa(\|u\|_* + \|f\|).$$

5 The stable problem

Let us now consider the stable problem with the additional lateral boundary condition $u|_{(0,T)\times\partial\Omega} = 0$. We may impose the boundary condition on the discrete level by changing the space $\mathcal{V}_h$, defined previously by (20), to

$$\mathcal{V}_h = H^1(0, T; W_h).$$

In the stable case we do not need an inequality analogous to (36) since the estimate in Theorem 2 does not contain an apriori term, that is, a term analogous to $\|u\|_{L^2((0,T)\times\Omega)}$ in Theorem 1. In the previous section, the Poincaré type inequality (27) was used only to obtain (36), and whence we can relax the conditions there by dropping (27). However, we still need to require that $|||\cdot|||_h$ is a norm on $\mathcal{V}_h \times \mathcal{W}_h$.

We will now proceed to a concrete case. The choices in Sect. 4 work also in the stable case, however, the additional structure allows us to choose a weaker stabilization. The use of a weaker stabilization is motivated by the fact that it leads to a weaker coupling of the two heat equations in (28), which again can be exploited when solving (28) in practice. See [10] for a computational implementation based on this.

The stabilizers and semi-norms are chosen as follows,

$$s_h(u_h, u_h) = \|h\nabla u_h(0,\cdot)\|^2_{L^2(\Omega)}, \quad s_h^*(z_h, z_h) = (\nabla z_h, \nabla z_h), \tag{45}$$

$$|u_h|_{\mathcal{V}_h} = s_h(u_h, u_h)^{1/2} + \|h\partial_t u_h\|, \quad \|z_h\|_{\mathcal{W}_h} = s_h^*(z_h, z_h)^{1/2}, \tag{46}$$

and we define $H^* = H_0^{(1,1)}$. To counter the lack of primal stabilization on most of the cylinder $(0,T)\times\Omega$, we choose π_h to be the orthogonal projection $\pi_h : H_0^1(\Omega) \to W_h$ with respect to the inner product $(\nabla u, \nabla v)_{L^2(\Omega)}$. That is, π_h is defined by

$$(\nabla \pi_h u, \nabla v_h)_{L^2(\Omega)} = (\nabla u, \nabla v_h)_{L^2(\Omega)}, \quad u \in H_0^1(\Omega), \ v_h \in W_h. \tag{47}$$

As Ω is a convex polyhedron, this choice satisfies the estimates (11)–(12), see e.g. [16, Th. 3.12-18].

Lemma 6 *The choices* (45)–(47) *satisfy the properties* (41)–(44), (26) *and* (30). *Moreover,* $|||\cdot|||_h$ *is a norm on* $\mathcal{V}_h \times \mathcal{W}_h$.

Proof It is clear that the continuities (41) hold. We begin with the lower bound (42). By the orthogonality (47),

$$G(u_h, z - \pi_h z) = (\partial_t u_h, z - \pi_h z) \le \|h\partial_t u_h\| h^{-1} \|z - \pi_h z\| \le C\|h\partial_t u_h\| \|z\|_{(0,1)}.$$

Towards the lower bound (43), we use the orthogonality (47) as above,

$$G(u - \pi_h u, z_h) = (\partial_t u - \pi_h \partial_t u, z_h) \le Ch\|u\|_{(1,1)} \|z_h\|.$$

The bound (43) follows from the Poincaré inequality (22).

The bound (44) follows from the continuity of the trace

$$\|\nabla u(0, \cdot)\|_{L^2(\Omega)} \le \|u\|_{(1,1)}, \tag{48}$$

and the continuity of the projection π_h. The bound (26) follows immediately from the continuity of π_h.

We turn to the weak coercivity (30). The essential difference between the unstable and the stable case is that in the latter case $\partial_t u_h \in \mathcal{W}_h$ when $u_h \in \mathcal{V}_h$. We have

$$\begin{aligned} A_h[(u_h, z_h), (0, \partial_t u_h)] &= -s_h^*(z_h, \partial_t u_h) + G(u_h, \partial_t u_h) \\ &= \|\partial_t u_h\|^2 + a(u_h, \partial_t u_h) - a(z_h, \partial_t u_h), \end{aligned}$$

and thus using bilinearity of A_h,

$$\begin{aligned} A_h[(u_h, z_h), (u_h, \alpha h^2 \partial_t u_h - z_h)] &= s_h(u_h, u_h) + \alpha \|h\partial_t u_h\|^2 + \|u_h\|_\omega^2 + s_h^*(z_h, z_h) \\ &\quad + \alpha h^2 a(u_h, \partial_t u_h) - \alpha h^2 a(z_h, \partial_t u_h), \end{aligned} \tag{49}$$

where $\alpha > 0$. We will establish (30) by showing that there is $\alpha \in (0, 1)$ such that

$$|||(u_h, w_h - z_h)|||_h \le C|||(u_h, z_h)|||_h, \tag{50}$$

$$|||(u_h, z_h)|||_h^2 \le C A_h[(u_h, z_h), (u_h, w_h - z_h)], \tag{51}$$

where $w_h = \alpha h^2 \partial_t u_h$.

Towards (50) we observe that

$$\begin{aligned} |||(u_h, w_h - z_h)|||_h^2 &= |||(u_h, z_h)|||_h^2 - 2s_h^*(z_h, w_h) + s_h^*(w_h, w_h) \\ &\le 2|||(u_h, z_h)|||_h^2 + 2s_h^*(w_h, w_h). \end{aligned}$$

We use the discrete inverse inequality (10) to bound the second term

$$s_h^*(w_h, w_h) = \alpha^2 h^4 \|\partial_t \nabla u_h\|^2 \le C\alpha^2 h^2 \|\partial_t u_h\|^2 \le C\alpha^2 |||(u_h, z_h)|||^2, \quad \alpha > 0.$$

It remains to show (51). Towards bounding the first cross term in (49) we observe that

$$2a(u_h, \partial_t u_h) = \int_0^T \partial_t \|\nabla u_h(t,\cdot)\|^2_{L^2(\Omega)} dt = \|\nabla u_h(T,\cdot)\|^2_{L^2(\Omega)} - \|\nabla u_h(0,\cdot)\|^2_{L^2(\Omega)}.$$

Hence $\alpha h^2 a(u_h, \partial_t u_h) \ge -\alpha s_h(u_h, u_h)/2$. We use the arithmetic-geometric inequality,

$$ab \le (4\epsilon)^{-1} a^2 + \epsilon b^2, \quad a, b \in \mathbb{R},\ \epsilon > 0,$$

and the discrete inverse inequality (10) to bound the second cross term in (49),

$$\begin{aligned} \alpha h^2 a(z_h, \partial_t u_h) &\le \alpha (4\epsilon)^{-1} a(z_h, z_h) + \alpha\epsilon h^4 \|\partial_t \nabla u_h\|^2 \\ &\le \alpha (4\epsilon)^{-1} a(z_h, z_h) + C\alpha\epsilon \|h\partial_t u_h\|^2. \end{aligned}$$

Choosing $\epsilon = 1/(2C)$ we obtain

$$\begin{aligned} &A_h[(u_h, z_h), (u_h, w_h - z_h)] \\ &\quad \ge (1 - \alpha/2) s_h(u_h, u_h) + \alpha \|h\partial_t u_h\|^2/2 + \|u_h\|^2_\omega + (1 - C\alpha/2) s_h^*(z_h, z_h), \end{aligned}$$

and therefore (51) holds with small enough $\alpha > 0$.

Finally, using the Poincaré inequality (22), we see that $|||(u_h, z_h)|||_h = 0$ implies that $z_h = 0$ and $u_h(0,\cdot) = 0$. As also $\partial_t u_h = 0$, we have $u_h = 0$, and therefore $|||\cdot|||_h$ is a norm. □

In the stable case we have the following convergence rate.

Theorem 6 *Let $\omega \subset \Omega$ be open and non-empty and let $0 < T_1 < T$. Suppose that (A2) holds. Let $u \in H^*$ and define $f = \partial_t u - \Delta u$ and $q = u|_\omega$. Suppose that the primal and dual stabilizers satisfy* (41)–(44), (26) *and* (30). *Then* (28) *has a unique solution $(u_h, z_h) \in \mathcal{V}_h \times \mathcal{W}_h$, and there exists $C > 0$ such that for all $h \in (0, 1)$*

$$\begin{aligned} &\|u_h - u\|_{C(T_1,T;L^2(\Omega))} + \|u_h - u\|_{L^2(T_1,T;H^1(\Omega))} + \|u_h - u\|_{H^1(T_1,T;H^{-1}(\Omega))} \\ &\quad \le Ch(\|u\|_* + \|f\|). \end{aligned}$$

The proof is analogous to that of Theorem 3, however, we give it here for the sake of completeness.

Proof We begin again by showing the estimate

$$|||(u_h - \pi_h u, z_h)|||_h \le Ch\|u\|_*. \tag{52}$$

The weak coercivity (30) implies that in order to show (52) it is enough bound the three terms on the right hand side of (32). For the first of them, that is, $(u - \pi_h u, v_h)_\omega$, we use (12) as in the proof of Lemma 4. The lower bound (43) applies to the second term $G(u - \pi_h u, w_h)$, and for the third one we use the continuity (41) and the upper bound (44),

$$s_h(\pi_h u, v_h) \leq C|\pi_h u|_{\mathcal{V}_h}|v_h|_{\mathcal{V}_h} \leq Ch\|u\|_*|v_h|_{\mathcal{V}_h}.$$

We define the residual r as in the proof of Theorem 3, and show next that (34) holds. It is enough to bound the three terms on the right hand side of (33). The lower bound (42) applies to the first term $G(u_h, w - \pi_h w)$, for the second term $(f, w - \pi_h w)$ we use (12) as in the proof of Theorem 3, and for the third term we use the continuity (41) and the upper bound (26)

$$s_h^*(z_h, \pi_h w) \leq C\|z_h\|_{\mathcal{W}_h}\|\pi_h w\|_{\mathcal{W}_h} \leq C\|z_h\|_{\mathcal{W}_h}\|w\|_{(0,1)}.$$

The inequalities (34), (52) and (44) imply

$$\|r\|_{(0,-1)} \leq C(|u_h - \pi_h u|_{\mathcal{V}_h} + |\pi_h u|_{\mathcal{V}_h} + \|z_h\|_{\mathcal{W}_h} + h\|f\|) \leq Ch(\|u\|_* + \|f\|).$$

Finally, Theorem 2 implies that

$$\begin{aligned}&\|u_h - u\|_{C(T_1,T;L^2(\Omega))} + \|u_h - u\|_{L^2(T_1,T;H^1(\Omega))} + \|u_h - u\|_{H^1(T_1,T;H^{-1}(\Omega))}\\ &\quad \leq C\|u_h - u\|_\omega + Ch(\|u\|_* + \|f\|).\end{aligned}$$

The claim follows by using (52) and (12),

$$\|u_h - u\|_\omega \leq \|u_h - \pi_h u\|_\omega + \|\pi_h u - u\|_\omega \leq Ch\|u\|_*.$$

Here we used also the assumption that H^* is a continuously embedded subspace of the energy space (2), namely, this implies that the embedding $H^* \subset H^{(0,1)}$ is continuous. □

Remark 1 If the data q, f is perturbed in the stable case, the data assimilation problem behaves like a typical well posed problem, that is, the term $\delta(\tilde{q}, \tilde{f})$ needs to be added on the right-hand side of the estimate in Theorem 6, but this time without any negative power of h.

6 Conclusion

In the present paper our aim was to show how methods known from the theory of stabilized finite element methods can be applied to the design of computational methods for non-stationary data assimilation problems by first discretizing and then regularizing the corresponding 4DVAR optimization system.

A key feature of this framework is that the error analysis is based on numerical stability of residual quantities that are independent of the stability properties of the continuous model. The residual quantities are then bounded by using conditional stability estimates, based on Carleman estimates, to derive error estimates that reflect the approximation properties of the finite element space and the stability of the continuous problem in an optimal way. An upshot of this approach is that it gives a clear lead on how to design regularization for a given problem in order to obtain the best accuracy of the approximation with the least effect of perturbations in data.

Observe that the stabilization operators proposed herein are not unique, for instance it is straightforward to show that the dual stabilizer in Eq. (19) may be chosen as the first part of the primal stabilizer, leading to similar error estimates for unperturbed data. The error estimates also gives an indication on what form Carleman estimates should take to make them immediately applicable for error analysis.

Acknowledgements The authors thank J. Le Rousseau and S. Ervedoza for useful discussions. LO was partly supported by EPSRC Grant EP/L026473/1. EB was partly supported by EPSRC Grants EP/J002313/2 and EP/P01576X/1.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Continuum estimates

Theorem 1 is a consequence of the so-called three cylinders inequality that goes back to [19], see [27,28] for an overview of the related literature. The variant of the inequality, needed in the above convergence analysis, seems not to appear in the literature, and we prove it here by using the Carleman estimate [20].

Theorem 2 is a variant of [14], the difference being that in [14] the domain Ω is assumed to have C^2-smooth boundary. We outline below the modifications needed in the case that Ω is a convex polyhedron.

We use the notation $B(x,r) = \{y \in \Omega;\ d(y,x) < r\}$, $x \in \Omega$, $r > 0$, where d is the Euclidean distance.

Theorem 7 (Three cylinders inequality) *Let $x_0 \in \Omega$ and $0 < r_1 < r_2 < d(x_0, \partial\Omega)$. Write $B_j = B(x_0, r_j)$, $j = 1, 2$. Let $T > 0$ and $0 < \epsilon < T/2$. Then there are $C > 0$ and $\kappa \in (0,1)$ such that for all $u \in C^2(\mathbb{R} \times \Omega)$*

$$\|u\|_{L^2(\epsilon, T-\epsilon; H^1(B_2))} \le C(\|u\|_{L^2(0,T;H^1(B_1))} + \|Lu\|_{L^2((0,T)\times\Omega)})^{\kappa} \|u\|^{1-\kappa}_{L^2(0,T;H^1(\Omega))}.$$

Proof Let $0 < r_0 < r_1$ and $r_2 < r_3 < r_4 < d(x_0, \partial\Omega)$. Define $B_j = B(x_0, r_j)$, $j = 0, 3, 4$. We choose non-positive $\rho_1 \in C^\infty(\Omega)$ such that $\rho_1 > -r_0$ in B_0 and that $\rho_1(x) = -d(x, x_0)$ outside B_0. Define $I_1 = (\epsilon, T-\epsilon)$ and $I_2 = (\epsilon/2, T-\epsilon/2)$, and choose non-positive $\rho_2 \in C^\infty(\mathbb{R})$ such that $\rho_2 \le -r_3$ outside I_2 and $\rho_2 = 0$ in I_1. We define $\rho(t,x) = \rho_1(x) + \rho_2(t)$ and $\phi = e^{\alpha\rho}$ where $\alpha > 0$.

We will apply [20, Th. 1.1] to the heat operator $L = \partial_t - \Delta$, with the above weight function ϕ. Condition (1.5) in the theorem reduces to

$$D^2\phi(\zeta, \overline{\zeta}) > 0, \quad \zeta = \xi + i\tau\nabla\phi, \ |\xi| = \tau|\nabla\phi|, \ \tau > 0, \ \xi \in \mathbb{R}^n, \tag{53}$$

where $D^2\phi$ is the Hessian of ϕ with respect to x. We have $\nabla\phi = \alpha\phi\nabla\rho$ and

$$\begin{aligned} D^2\phi(\zeta, \overline{\zeta}) &= \alpha\phi(\alpha|\nabla\rho\cdot\zeta|^2 + D^2\rho(\zeta, \overline{\zeta})) \\ &\geq \alpha\phi(\alpha(\tau\alpha\phi|\nabla\rho|^2)^2 - \mu(\tau\alpha\phi|\nabla\rho|)^2), \end{aligned}$$

where μ is a constant depending only on ρ. As $|\nabla\rho| = 1$ outside B_0, condition (53) holds in $\overline{B_4} \setminus B_0$ for large enough $\alpha > 0$. Let $\chi \in C_0^\infty((0,T) \times (B_4 \setminus B_0))$ satisfy $\chi = 1$ in $I_2 \times (B_3 \setminus B_1)$, and set $w = \chi u$. Then [20, Th. 1.1] implies that for large $\tau > 0$,

$$\int_0^T \int_{B_4\setminus B_0} (\tau|\nabla w|^2 + \tau^3|w|^2)e^{2\tau\phi}\, dxdt \leq C \int_0^T \int_{B_4\setminus B_0} |Lw|^2 e^{2\tau\phi}\, dxdt. \tag{54}$$

We define $\Phi(r) = e^{-\alpha r}$, $I = (0, T)$ and

$$Q_1 = I_2 \times B_1, \quad Q_2 = ((I \setminus I_2) \times B_4) \cup (I \times (B_4 \setminus B_3)).$$

Observe that $\phi \leq \Phi(r_3)$ in Q_2, and that the commutator $[L, \chi]$ vanishes outside $Q_1 \cup Q_2$. Hence the right-hand side of (54) is bounded by a constant times

$$\begin{aligned} &\int_{(0,T)\times B_4} |Lu|^2 e^{2\tau\phi} dxdt + \int_{Q_1\cup Q_2} |[L, \chi]u|^2 e^{2\tau\phi} dxdt \\ &\quad \leq e^{2\tau}\|Lu\|^2_{L^2((0,T)\times B_4)} + e^{2\tau}\|u\|^2_{L^2(0,T;H^1(B_1))} + e^{2\tau\Phi(r_3)}\|u\|^2_{L^2(0,T;H^1(B_4))}. \end{aligned} \tag{55}$$

The left-hand side of (54) is bounded from below by

$$\int_{I_1\times(B_2\setminus B_1)} \left(\tau|\nabla u|^2 + \tau^3|u|^2\right) e^{2\tau\phi}\, dxdt \geq e^{2\tau\Phi(r_2)}\|u\|^2_{L^2(\epsilon,T-\epsilon;H^1(B_2\setminus B_1))}. \tag{56}$$

The inequalities (54)–(56) imply

$$\begin{aligned} \|u\|_{L^2(\epsilon,T-\epsilon;H^1(B_2))} &\leq e^{\tau}\left(\|Lu\|_{L^2((0,T)\times B_4)} + \|u\|_{L^2(0,T;H^1(B_1))}\right) \\ &\quad + e^{-p\tau}\|u\|_{L^2(0,T;H^1(B_4))}. \end{aligned}$$

where $p = \Phi(r_2) - \Phi(r_3) > 0$. The claim follows from [22, Lemma 5.2]. □

Proof of Theorem 1 We use notation from the proof of Theorem 7. By replacing ω with a smaller set we may assume without loss of generality that it is a ball of the above form B_1. We will show a local version of the claimed estimate where B is replaced by

a ball of the form B_2. The general case follows by covering B by finite chains of balls starting from ω, and by iterating the local result.

Let us first consider the case $u \in C^\infty(\mathbb{R} \times \Omega)$. We choose $0 < r_0 < r_1$ and $r_2 < r_3 < d(x_0, \Omega)$ and write $B_j = B(x_0, r_j)$, $j = 0, 3$. Let $0 < \epsilon_0 < \epsilon$ and choose $\eta \in C^\infty(\mathbb{R})$ such that $\eta = 0$ near the origin and $\eta(t) = 1$ for $t > \epsilon_0$. Let w be the solution of

$$\begin{aligned} Lw &= \eta Lu \quad \text{in } (0, T) \times B_3, \\ w|_{x \in \partial B_3} &= 0, \quad w|_{t=0} = 0, \end{aligned}$$

and set $v = u - w$. Note that $w \in C^\infty([\epsilon_0, T] \times \Omega)$. Theorem 7 with u replaced by v, B_1 by B_0, Ω by B_3, and $t = 0$ by $t = \epsilon_0$ implies that

$$\begin{aligned} \|v\|_{L^2(\epsilon, T-\epsilon; H^1(B_2))} &\le C\|v\|^{\kappa}_{L^2(\epsilon_0, T; H^1(B_0))} \|v\|^{1-\kappa}_{L^2(\epsilon_0, T; H^1(B_3))} \\ &\le C(\|u\|_{L^2(\epsilon_0, T; H^1(B_0))} + \|Lu\|_{0,-1})^{\kappa} (\|u\|_{L^2(\epsilon_0, T; H^1(B_3))} + \|Lu\|_{0,-1})^{1-\kappa}. \end{aligned}$$

Here we have applied the energy estimate for the heat equation to w, see (57) below. Moreover,

$$\|u\|_{L^2(\epsilon, T-\epsilon; H^1(B_2))} \le \|v\|_{L^2(\epsilon, T-\epsilon; H^1(B_2))} + C\|Lu\|_{0,-1}.$$

We choose $\chi \in C_0^\infty((0, T) \times B_1)$ such that $\chi = 1$ in $(\epsilon_0, T - \epsilon_0) \times B_0$. Then χu satisfies

$$L(\chi u) = \chi Lu + [\chi, L]u, \quad (\chi u)|_{x \in \partial B_1} = 0, \quad (\chi u)|_{t=0} = 0,$$

and as the commutator $[\chi, L]$ is of first order in space and zeroth order in time,

$$\|u\|_{L^2(\epsilon_0, T-\epsilon_0; H^1(B_0))} \le \|\chi u\|_{L^2(0, T; H^1(B_1))} \le C\|Lu\|_{0,-1} + C\|u\|_{L^2((0,T) \times B_1)}.$$

Analogously,

$$\|u\|_{L^2(\epsilon_0, T-\epsilon_0; H^1(B_3))} \le C\|Lu\|_{0,-1} + C\|u\|_{L^2((0,T) \times \Omega)}.$$

and we have shown the local estimate in the case that u is smooth. To conclude we observe that smooth functions are dense in the space (2). □

Let us now turn to Theorem 2. In the case of convex, polygonal Ω, the following simple lemma gives a weight function satisfying Condition 1.1 of [14].

Lemma 7 *Suppose that $\Omega \subset \mathbb{R}^n$ is open and convex set and that $\partial\Omega$ is piecewise smooth. Let $x_0 \in \Omega$ and $\epsilon > 0$ satisfy $\overline{B(x_0, \epsilon)} \subset \Omega$. Then there is $\rho \in C^\infty(\Omega)$ satisfying $\rho < 0$ in Ω, $|\nabla\rho| = 1$ in $\Omega \setminus B(x_0, \epsilon)$ and $\partial_\nu \rho \le 0$ on $\partial\Omega$.*

Proof We choose a strictly negative $\rho \in C^\infty(\Omega)$ such that $\rho(x) = -d(x, x_0)$ outside $B(x_0, \epsilon)$. Then $\nabla\rho = -(x - x_0)/|x - x_0|$ outside $B(x_0, \epsilon)$. By convexity, $(x - x_0) \cdot \nu \ge 0$ on $\partial\Omega$, and therefore $\partial_\nu \rho \le 0$ on the boundary. □

Now [14, Lemma 1.3] can be written in the following form:

Lemma 8 (Global Carleman estimate by Imanuvilov) *Let $\Omega \subset \mathbb{R}^n$, $x_0 \in \Omega$ and $\epsilon > 0$ be as in Lemma 7. Let ρ be the function given by Lemma 7. Let $T > 0$ and define $\ell = \lambda\hbar^{-1}(e^{\alpha\rho} - 1)$, where $\hbar = t(T-t)$, $\alpha, \lambda > 0$. Then there are $C, \alpha, \lambda_0 > 0$ such that for all $w \in C^2([0,T] \times \overline{\Omega})$ satisfying $w = 0$ on $(0,T) \times \partial\Omega$ and all $\lambda > \lambda_0$ it holds that*

$$\int_0^T \int_\Omega \left(\tau|\nabla w|^2 + \tau^3|w|^2\right) e^{2\ell} dxdt$$
$$\leq C \int_0^T \int_\Omega |Lw|^2 e^{2\ell} dxdt + C \int_0^T \int_{B(x_0,\epsilon)} \tau^3 |w|^2 e^{2\ell} dxdt.$$

where $\tau = \lambda\hbar^{-1}$.

Proof of Theorem 2 We write $\mathcal{C} = C^\infty(0,T; C_0^\infty(\Omega))$ and denote by $\mathcal{H}$ the energy space (4). Recall that $\mathcal{H}$ is continuously embedded in $C(0,T; L^2(\Omega))$, see e.g. [24, Lemma 11.4], and that $\mathcal{C}$ is dense in $\mathcal{H}$. By density it is enough to consider the case $u \in \mathcal{C}$. Recall also the energy estimate for $v \in \mathcal{H}$, see e.g. [24, Theorem 11.3],

$$\|v\|_{\mathcal{H}} \leq C\|v(0)\|_{L^2(\Omega)} + C\|Lv\|_{L^2(0,T;H^{-1}(\Omega))}. \tag{57}$$

We write $f = \partial_t - \Delta u$, let $v \in \mathcal{H}$ be the solution of

$$\partial_t v - \Delta v = f, \quad v|_{t=0} = 0,$$

and define $w = u - v$. Then w satisfies $\partial_t w - \Delta w = 0$. Choose $x_0 \in \omega$ and $\epsilon > 0$ such that $B(x_0, \epsilon) \subset \omega$, and apply Lemma 8 on w. Note that $\tau^3 e^{2\ell} \to 0$ as $t \to 0$ or $t \to T$. Let $0 < \delta_0 < \delta$, then $e^{2\ell}$ is strictly positive on $[\delta_0, T - \delta_0] \times \Omega$. In particular,

$$\|w\|_{L^2((\delta_0, T-\delta_0)\times\Omega)} \leq C\|w\|_{L^2((0,T)\times\omega)}.$$

The estimate (57) implies that

$$\|w(\delta)\|^2_{L^2(\Omega)} \leq C\|w(s)\|^2_{L^2(\Omega)}, \quad s \in (0,\delta).$$

We integrate this over the interval (δ_0, δ) to obtain

$$(\delta - \delta_0)\|w(\delta)\|^2_{L^2(\Omega)} \leq C\|w\|^2_{L^2((\delta_0, T-\delta_0)\times\Omega)}.$$

By using (57) again, we have

$$\|v(\delta)\|_{L^2(\Omega)} + \|v\|_{L^2((0,T)\times\Omega)} \leq C\|f\|_{L^2(0,T;H^{-1}(\Omega))}.$$

Hence

$$\begin{aligned}\|u(\delta)\|_{L^2(\Omega)} &\le C\|w\|_{L^2((0,T)\times\omega)} + C\|f\|_{L^2(0,T;H^{-1}(\Omega))}\\ &\le C\|u\|_{L^2((0,T)\times\omega)} + C\|f\|_{L^2(0,T;H^{-1}(\Omega))}.\end{aligned}$$

The claim follows by applying (57) once more. □

References

1. Auroux, D., Blum, J.: Back and forth nudging algorithm for data assimilation problems. C. R. Math. Acad. Sci. Paris **340**(12), 873–878 (2005)
2. Baumeister, J.: Stable Solution of Inverse Problems. Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig (1987)
3. Boyer, F., Hubert, F., Le Rousseau, J.: Discrete Carleman estimates for elliptic operators and uniform controllability of semi-discretized parabolic equations. J. Math. Pures Appl. (9) **93**(3), 240–276 (2010)
4. Boyer, F., Hubert, F., Le Rousseau, J.: Uniform controllability properties for space/time-discretized parabolic equations. Numer. Math. **118**(4), 601–661 (2011)
5. Brenner, S.C.: Poincaré–Friedrichs inequalities for piecewise H^1 functions. SIAM J. Numer. Anal. **41**(1), 306–324 (2003)
6. Brenner, S .C., Scott, L .R.: The Mathematical Theory of Finite Element Methods, Texts in Applied Mathematics, vol. 15, 3rd edn. Springer, New York (2008)
7. Bukhgeĭm, A.L., Klibanov, M.V.: Uniqueness in the large of a class of multidimensional inverse problems. Dokl. Akad. Nauk SSSR **260**(2), 269–272 (1981)
8. Burman, E.: Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: elliptic equations. SIAM J. Sci. Comput. **35**(6), A2752–A2780 (2013)
9. Burman, E.: Error estimates for stabilized finite element methods applied to ill-posed problems. C. R. Math. Acad. Sci. Paris **352**(7–8), 655–659 (2014)
10. Burman, E., Ish-Horowicz, J., Oksanen, L.: Fully discrete finite element data assimilation method for the heat equation (2017, preprint). arXiv:1707.06908
11. Burman, E., Larson, M.G., Hansbo, P.: Solving ill-posed control problems by stabilized finite element methods: an alternative to Tikhonov regularization. Inverse Probl **34**(3), (2018)
12. Casas, E., Vexler, B., Zuazua, E.: Sparse initial data identification for parabolic PDE and its finite element approximations. Math. Control Relat. Fields **5**(3), 377–399 (2015)
13. Cheng, J., Yamamoto, M.: One new strategy for a priori choice of regularizing parameters in Tikhonov's regularization. Inverse Probl. **16**(4), L31 (2000)
14. Èmanuilov, O.Y.: Controllability of parabolic equations. Mat. Sb. **186**(6), 109–132 (1995)
15. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of Inverse Problems. Mathematics and its Applications, vol. 375. Kluwer Academic Publishers Group, Dordrecht (1996)
16. Ern, A., Guermond, J.-L.: Theory and Practice of Finite Elements. Applied Mathematical Sciences, vol. 159. Springer, New York (2004)
17. Ervedoza, S., Zuazua, E.: Numerical Approximation of Exact Controls for Waves. Springer Briefs in Mathematics. Springer, New York (2013)
18. García, G.C., Osses, A., Puel, J.P.: A null controllability data assimilation methodology applied to a large scale ocean circulation model. ESAIM Math. Model. Numer. Anal. **45**(2), 361–386 (2011)
19. Glagoleva, R.J.: The three cylinder theorem and its applications. Dokl. Akad. Nauk SSSR **163**, 801–804 (1965)
20. Isakov, V.: Carleman type estimates in an anisotropic case and applications. J. Differ. Equ. **105**(2), 217–238 (1993)
21. Le Dimet, F.-X., Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus A **38A**(2), 97–110 (1986)
22. Le Rousseau, J., Lebeau, G.: On Carleman estimates for elliptic and parabolic operators. Applications to unique continuation and control of parabolic equations. ESAIM Control Optim. Calc. Var. **18**(3), 712–747 (2012)

23. Puel, J.-P.: A nonstandard approach to a data assimilation problem and Tychonov regularization revisited. SIAM J. Control Optim. **48**(2), 1089–1111 (2009)
24. Renardy, M., Rogers, R .C.: An Introduction to Partial Differential Equations. Texts in Applied Mathematics, vol. 13, 2nd edn. Springer, New York (2004)
25. Scott, L.R., Zhang, S.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math. Comput. **54**(190), 483–493 (1990)
26. Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-Posed Problems. V. H. Winston & Sons, Washington, D.C.: Wiley, New York-Toronto, Ont.-London (1977)
27. Vessella, S.: Quantitative estimates of unique continuation for parabolic equations, determination of unknown time-varying boundaries and optimal stability estimates. Inverse Probl. **24**(2), 023001, 81 (2008)
28. Yamamoto, M.: Carleman estimates for parabolic equations and applications. Inverse Probl. **25**(12), 123013, 75 (2009)