ESAIM: M2AN 52 (2018) 2065–2082
https://doi.org/10.1051/m2an/2018030

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

FULLY DISCRETE FINITE ELEMENT DATA ASSIMILATION METHOD FOR THE HEAT EQUATION

Erik Burman[1,*], Jonathan Ish-Horowicz[2] and Lauri Oksanen[3]

Abstract. We consider a finite element discretization for the reconstruction of the final state of the heat equation, when the initial data is unknown, but additional data is given in a sub domain in the space time. For the discretization in space we consider standard continuous affine finite element approximation, and the time derivative is discretized using a backward differentiation. We regularize the discrete system by adding a penalty on the H^1-semi-norm of the initial data, scaled with the mesh-parameter. The analysis of the method uses techniques developed in E. Burman and L. Oksanen [*Numer. Math.* **139** (2018) 505–528], combining discrete stability of the numerical method with sharp Carleman estimates for the physical problem, to derive optimal error estimates for the approximate solution. For the natural space time energy norm, away from $t = 0$, the convergence is the same as for the classical problem with known initial data, but contrary to the classical case, we do not obtain faster convergence for the L^2-norm at the final time.

Mathematics Subject Classification. 65M12,65M15, 65M30, 65M32

Received August 15, 2017. Accepted April 30, 2018.

1. Introduction

Time discretization of parabolic problems, discretized in space using finite element methods, is a well studied topic, see for example the monograph by Thomée [34]. The analysis for all such methods relies on the satisfaction of the hypothesis of the Lions theorem [26], stating the existence, uniqueness and stability properties of the problem.

The classical problem can be cast in the abstract form, find $u \in V$ such that

$$(\partial_t u, v)_H + a(u, v) = \langle f, v \rangle_{V',V}\,, \tag{1.1}$$

$$u(0) = u_0 \in H, \tag{1.2}$$

where V, H are some Hilbert spaces, with V dense in H and imbedded with continuous identity, $\langle \cdot, \cdot \rangle_{V',V}$ denotes the duality pairing between V and its dual, and $a(u, v) : V \times V \mapsto \mathbb{R}$ a symmetric bilinear form representing the weak form of a second order differential operator. A key ingredient of the theory is that the spatial operator

Keywords and phrases: Heat equation, inverse problem, data assimilation, stabilized finite elements.

[1] Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK.
[2] MRC Biostatistics Unit, University of Cambridge, Cambridge Biomedical Campus, Cambridge CB2 0SR, UK.
[3] Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK.

* Corresponding author: e.burman@ucl.ac.uk

satisfies the the Gårding's inequality, there are $\alpha > 0$ and $\beta \geq 0$ such that for all $v \in V$ there holds

$$a(v,v) \geq \alpha \|v\|_V^2 - \beta \|v\|_H^2. \tag{1.3}$$

In many situations for instance in environmental science and meteorology the initial data is not available, instead some other data in the space time domain have been collected through measurements. This leads to a data assimilation problem, that is, a problem to incorporate the observations of the physical system into the state of a computational model of the system. Computations can not be based on the classical theory, since the equation (1.2) can not be enforced when u_0 is not known.

It is then an interesting problem in computational mathematics what quantities can be approximated and what is the effect of measurement errors on such an approximation. The approximation methods need to take into account the fact that these data assimilation problems are ill-posed in the sense that a necessary condition for them to be solvable is that the observations indeed come from the system. That is, the theory of these problems concerns uniqueness and stability, but not existence of solutions. From computational point of view, discretization causes already a perturbation in the exact solution and the stability of the data assimilation problem is our main concern. In particular, we want to show that an approximate, discrete solution to the problem converges to the exact one at an optimal rate. We will also discuss the case of noisy data.

In [7], we studied finite element methods for two data assimilation problems with unknown u_0. The two problems differ in the sense that the lateral boundary data for u is either known or unknown. In the first case (1.3) holds, whereas unknown lateral boundary data leads to a failure of (1.3). This again gives rise to very different stability properties. When the lateral boundary data is known, the data assimilation problem to recover u in $(\delta, T) \times \Omega$, with $\delta > 0$, is Lipschitz stable in suitable spaces, but the optimal stability is of conditional Hölder type when no information is given on the lateral boundary. The recovery of the initial condition $u(0,\cdot)$ is exponentially unstable in both cases [21]. Here we restrict our attention to the case with known lateral boundary data, and extend the corresponding results of [7] to a fully discrete method. In [7] discretization only in space was considered.

The fully discrete analysis does not reduce straightforwardly to the semi-discrete case, as demonstrated by the fact that, in order to achieve the optimal convergence rate with respect to the size of the time step, an additional regularization term is needed, see Theorem 3.3 below. There we consider two different asymptotic rates, $\tau = \mathcal{O}(h)$ and $\tau = \mathcal{O}(h^2)$, between the size of the finite element mesh h and the time step τ, and the analysis under the less restrictive rate $\tau = \mathcal{O}(h)$ is valid only when additional regularization is present (the case $\gamma_1 > 0$ in the theorem). In Section 4, we give a computational example showing that the additional regularization is necessary.

To keep the exposition simple, we assume that the physical system is modelled by the heat equation

$$\partial_t u - \Delta u = f \quad \text{in } (0,T) \times \Omega, \tag{1.4}$$

with $u = 0$ on the boundary $\partial\Omega$. Here $\Omega \subset \mathbb{R}^d$ is a connected polyhedral domain. Of course, in the absence of additional information, the equation (1.4) does not have a unique solution. We assume that measurements of u, denoted by q, are available in the space time domain $(0,T) \times \omega$, where ω is a non-empty, open subset of Ω. We want to solve (1.4) under the additional constraint that

$$u = q \quad \text{in } (0,T) \times \omega. \tag{1.5}$$

It is known that if there exists a solution u to the equations (1.4) and (1.5), then the solution is unique.

A convenient way of solving the problem (1.4)–(1.5) is through optimization. Methods where the distance to the measured data in some norm over a space-time domain, plus some regularizing term, are minimised under the constraint of the partial differential equation are commonly referred to as 4DVAR. The abbreviation refers to the four dimensional character (time plus three space dimensions) of the variational problem. Such methods are important in data assimilation for meteorology and environmental science and we refer to [1, 14, 25, 29, 31–33]

for some formulations and results in the applied sciences. Although these methods are widely used and popular tools, there appears to be no rigorous numerical analysis assessing discretisation errors for them. One objective of the present publication is to start filling this gap. To make the presentation as accessible as possible we only consider space discretization using piecewise affine H^1-conforming finite elements and time discretization using the backward Euler method, the approach however generalizes to higher order scheme drawing on the ideas from [6].

We will now discuss the previous mathematical literature on the problem (1.4)–(1.5). We focus on techniques that work in dimensions $1 + d$ with $d > 1$, and refer to the papers [18, 35] and references therein for the $1 + 1$-dimensional case. Our finite element method builds on the stability estimate [15], and in a wider context, the literature on continuum stability estimates for parabolic data assimilation (or unique continuation) problems is reviewed in [17, 36].

Computational methods for the problem (1.4)–(1.5) go back to [24] where the quasi-reversibility method was introduced. Variations of this method for parabolic problems were developed in [20, 23, 30] and in [2], and we refer to [21, 22] for a review of the quasi-reversibility method outside the parabolic context. Although for example the papers [2, 20] consider convergence with respect to a Tikhonov type regularization parameter, none of the above papers prove convergence rates with respect to the refinement of a discretization. Proving such a convergence rate is the main novelty of the present paper. Moreover, compared to the previous literature, an attractive feature of our method is that no auxiliary Tikhonov type regularization parameters need to be introduced, the only asymptotic parameters are the size of the finite element mesh in space and the size of the time step.

Both the quasi-reversibility method and our method are based on Carleman estimates for the continuous problem. In our case, the proof of the stability estimate in Theorem 3.2 below is based on the Carleman estimate in [15]. An alternative approach is to derive Carleman estimates directly on the discrete level, see for example [3] where such an approach was used for the closely related null controllability problem for the heat equation.

The approach in the present paper has grown out of the study of stabilized finite element methods for unique continuation problems for elliptic equations [4, 5, 8]. Another line of research that appears to be converging to a similar optimization based approach originates from the numerical analysis of the exact controllability of the wave equation [9, 11, 13]. The approach has been applied to stable unique continuation problems for the wave equation [10, 12] and to the null controllability problem for the heat equation.

2. Discrete optimization problem

Following [7], we first discretize (1.4) in space only. Let $\mathcal{T}_h$ be a conforming triangulation of the polyhedral domain Ω. Let $h_K = \operatorname{diam}(K)$ be the local mesh parameter and $h = \max_{K\in\mathcal{T}_h} h_K$ the mesh size. We assume that the family of triangulations $\{\mathcal{T}_h\}_h$ is quasi-uniform in the sense that there exists a constant c_1 such that for all $K \in \mathcal{T}_h$ it holds that $h_K \le h \le c_1 h_K$. Let V_h be the standard space of piecewise affine continuous finite elements satisfying the zero boundary condition,

$$V_h = \{v \in H_0^1(\Omega);\ v|_K \in \mathbb{P}_1(K),\ \forall K \in \mathcal{T}_h\}.$$

We may then write a semi-discrete finite element formulation of (1.4) as follows, find $u \in C^1(0, T; V_h)$ such that

$$(\partial_t u, v) + a(u, v) = (f, v), \quad v \in V_h, \tag{2.1}$$

where

$$(u, v) = \int_\Omega uv \,\mathrm{d}x, \quad a(u, v) = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x.$$

The idea is then to minimize the distance to the data (1.5) under the constraint of this dynamical system.

In order to outline this idea, let us consider the following preliminary Lagrangian functional,

$$\mathcal{L}_0(u,z) := \frac{1}{2}\|u-q\|^2_{L^2((0,T)\times\omega)} + \int_0^T (\partial_t u, z) + a(u,z) - (f,z)\,\mathrm{d}t, \tag{2.2}$$

where $u \in \mathcal{U} = H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H^1_0(\Omega))$ and $z \in \mathcal{Z} = L^2(0,T;H^1_0(\Omega))$. Writing the Euler–Lagrange equations for $\mathcal{L}_0$ we arrive to the following problem, find $(u,z) \in \mathcal{U}\times\mathcal{Z}$ such that

$$\begin{aligned}\langle \partial_u \mathcal{L}_0(u,z), v\rangle &= \int_0^T (\partial_t v, z) + a(v,z) + (u-q,v)_\omega\,\mathrm{d}t = 0,\\ \langle \partial_z \mathcal{L}_0(u,z), w\rangle &= \int_0^T (\partial_t u, w) + a(u,w) - (f,w)\,\mathrm{d}t = 0\end{aligned}$$

for all $(v,w) \in \mathcal{U}\times\mathcal{Z}$. Here $(\cdot,\cdot)_\omega$ is the inner product on $L^2(\omega)$. Clearly, if $z=0$ and u solves (2.1) with $u|_{(0,T)\times\omega} = q$, then these equations are satisfied, and hence they are consistent with the data assimilation problem that we set. This leads to a first possible approach: discretize this system in time and find the stationary points of the discrete system. A numerical analysis however shows that this approach is unlikely to be successful as the term $(u-q,v)_\omega$ does not seem to give enough stability for the problem to converge, and indeed, our computational examples in Section 4 verify this. Instead we add certain regularization terms in the fully discrete context that we will describe next.

Let $N \in \mathbb{N}$ and $\tau > 0$ satisfy $N\tau = T$, and define $t_n = n\tau$. Furthermore, define for $u = (u^n)_{n=0}^N \in V_h^{N+1}$,

$$\partial_\tau u^n = \frac{u^n - u^{n-1}}{\tau}, \quad n = 1,\dots,N.$$

Consider the Lagrangian $\mathcal{L} : V_h^{N+1}\times V_h^N \to \mathbb{R}$ defined by

$$\begin{aligned}\mathcal{L}(u,z) = {}& \frac{1}{2}\gamma_M\tau\sum_{n=1}^N \|u^n - q^n\|^2_\omega + \frac{1}{2}\gamma_0\left\|h\nabla u^0\right\|^2 + \frac{1}{2}\gamma_1\tau\sum_{n=1}^N \|\tau\nabla\partial_\tau u^n\|^2\\ &+ \tau\sum_{n=1}^N \left((\partial_\tau u^n, z^n) + a(u^n,z^n) - (f^n,z^n)\right),\end{aligned} \tag{2.3}$$

where, for fixed functions $f \in C(0,T;L^2(\Omega))$ and $q \in C(0,T;L^2(\omega))$,

$$f^n = f(t_n), \quad q^n = q(t_n), \quad n = 1,\dots,N,$$

and γ_M, γ_0 and γ_1 are fixed constants satisfying

$$\gamma_M, \gamma_0 > 0 \quad \text{and} \quad \gamma_1 \geq 0. \tag{2.4}$$

Observe that the first term in (2.3) is a discrete, rescaled version of the first term in (2.2), that is, the data fitting term, and the sum on the second line is a discrete version of the integral in (2.2), that is, the constraint term corresponding to the heat equation. The terms containing γ_j, $j = 0,1$, correspond to additional regularization.

We emphasize that the constants γ_M, γ_0 and γ_1 are not Tikhonov type regularization parameters, since they do not converge to zero. The only asymptotic parameters in this paper are the spatial and temporal mesh sizes

h and τ. For theoretical purposes, we could simply take $\gamma_M = \gamma_1 = \gamma_0 = 1$, however, from the point of view of practical computations the size of these constants matters. This is discussed further in Section 4.3 below. Moreover, the choice $\gamma_1 = 0$ gives a method that converges with a slower rate, see Theorem 3.3 and Figure 1 below.

Defining the bilinear forms

$$\begin{aligned} A_1(u,w) &= \tau \sum_{n=1}^{N} \left((\partial_\tau u^n, w^n) + a(u^n, w^n) \right), \\ A_2((u,z),v) &= \gamma_M \tau \sum_{n=1}^{N} (u^n, v^n)_\omega + \gamma_0 (h\nabla u^0, h\nabla v^0) + \gamma_1 \tau \sum_{n=1}^{N} (\tau \nabla \partial_\tau u^n, \tau \nabla \partial_\tau v^n) \\ &\quad + \tau \sum_{n=1}^{N} \left((\partial_\tau v^n, z^n) + a(v^n, z^n) \right), \end{aligned}$$

the Euler–Lagrange equations for $\mathcal{L}$ are

$$A_1(u,w) = \tau \sum_{n=1}^{N} (f^n, w^n), \quad A_2((u,z),v) = \gamma_M \tau \sum_{n=1}^{N} (q^n, v^n)_\omega. \tag{2.5}$$

We define the seminorms

$$\begin{aligned} |||u|||_R^2 &= \gamma_M \tau \sum_{n=1}^{N} \|u^n\|_\omega^2 + \gamma_0 \left\| h\nabla u^0 \right\|^2 + \gamma_1 \tau \sum_{n=1}^{N} \|\tau \nabla \partial_\tau u^n\|^2, \\ |||u,z|||_D^2 &= \left\| z^1 \right\|^2 + \left\| z^N \right\|^2 + \tau^2 \sum_{n=2}^{N} \|\partial_\tau z^n\|^2 + \tau \sum_{n=1}^{N} \|\nabla z^n\|^2 \\ &\quad + \left\| h\nabla u^N \right\|^2 + h^2 \tau \sum_{n=1}^{N} \|\partial_\tau u^n\|^2 + h^2 \sum_{n=1}^{N} \|\tau \nabla \partial_\tau u^n\|^2, \\ |||v,w|||_C^2 &= |||v|||_R^2 + \tau \sum_{n=1}^{N} \|w^n\|^2. \end{aligned}$$

Note that $|||\cdot|||_D$ is, in fact, a norm on V_h^{2N+1}. Also, if $\gamma_1 > 0$ then $|||\cdot|||_R$ and $|||\cdot|||_C$ are norms on V_h^{N+1} and V_h^{2N+1}, respectively. The system (2.5) has the following coercivity property.

Proposition 2.1. *There is $C > 0$ such that for all $N \in \mathbb{N}$, $h > 0$ and (u,z) in V_h^{2N+1} there is (v,w) in V_h^{2N+1} satisfying*

$$|||u|||_R^2 + |||u,z|||_D^2 \le C \left(A_1(u,w) + A_2((u,z),v) \right), \quad |||v,w|||_C \le C|||u|||_R + C|||u,z|||_D.$$

Proof. We will show first that there is $\alpha > 0$ such that for all $(u,z) \in V_h^{2N+1}$

$$\frac{1}{2} \left(|||u|||_R^2 + \alpha |||u,z|||_D^2 \right) \le A_1(u, -z + \alpha h^2 \partial_\tau u) + A_2((u,z), u + \alpha \hat{z}), \tag{2.6}$$

where $\partial_\tau u = (\partial_\tau u^n)_{n=1}^N \in V_h^N$ and $\hat{z} = (\hat{z}^n)_{n=0}^N \in V_h^{N+1}$ is defined by $\hat{z}^0 = 0$ and $\hat{z}^n = z^n$, $n = 1, \ldots, N$. Observe that

$$|||u|||_R^2 = A_1(u, -z) + A_2((u, z), u).$$

The identity

$$\tau \sum_{n=1}^{N} (\partial_\tau u^n, u^n) = \frac{1}{2} \left(\left\| u^N \right\|^2 - \left\| u^0 \right\|^2 \right) + \frac{\tau^2}{2} \sum_{n=1}^{N} \left\| \partial_\tau u^n \right\|^2 \tag{2.7}$$

is the discrete analogue of

$$\int_0^T (\partial_t u, u) \, \mathrm{d}t = \frac{1}{2} \left(\|u(T)\|^2 - \|u(0)\|^2 \right).$$

To derive (2.7) we employ the polarization identity

$$\tau(\partial_\tau u^n, u^n) = \|u^n\|^2 - (u^{n-1}, u^n) = \|u^n\|^2 - \frac{1}{2} \left(\|u^n\|^2 + \left\| u^{n-1} \right\|^2 - \left\| u^n - u^{n-1} \right\|^2 \right),$$

and observe that there is a telescoping type cancellation. Using the identity (2.7) with the bilinear form $(\cdot, \cdot)$ replaced by $a(\cdot, \cdot)$, we have

$$\begin{aligned} A_1(u, \partial_\tau u) &= \tau \sum_{n=1}^{N} \left(\left\| \partial_\tau u^n \right\|^2 + a(u^n, \partial_\tau u^n) \right) \\ &= \tau \sum_{n=1}^{N} \left\| \partial_\tau u^n \right\|^2 + \frac{1}{2} \left(\left\| \nabla u^N \right\|^2 - \left\| \nabla u^0 \right\|^2 \right) + \frac{\tau^2}{2} \sum_{n=1}^{N} \left\| \nabla \partial_\tau u^n \right\|^2. \end{aligned}$$

Observe that if $\alpha \leq \gamma_0$ then $-\alpha h^2 \left\| \nabla u^0 \right\|^2 / 2$ is absorbed by $|||u|||_R^2$.

We have

$$\begin{aligned} A_2((u, z), \hat{z}) = \gamma_M \tau \sum_{n=1}^{N} (u^n, z^n)_\omega + \gamma_1 \tau \sum_{n=1}^{N} (\tau \nabla \partial_\tau u^n, \tau \nabla \partial_\tau \hat{z}^n) \\ + \tau \sum_{n=1}^{N} \left((\partial_\tau \hat{z}^n, z^n) + \left\| \nabla z^n \right\|^2 \right). \end{aligned}$$

The identity (2.7) gives

$$\tau \sum_{n=1}^{N} (\partial_\tau \hat{z}^n, z^n) = \frac{1}{2} \left\| z^N \right\|^2 + \frac{\tau^2}{2} \sum_{n=1}^{N} \left\| \partial_\tau \hat{z}^n \right\|^2 = \frac{1}{2} \left\| z^N \right\|^2 + \frac{1}{2} \left\| z^1 \right\|^2 + \frac{\tau^2}{2} \sum_{n=2}^{N} \left\| \partial_\tau z^n \right\|^2.$$

Let us now consider the cross terms. The Poincaré inequality gives

$$(u^n, z^n)_\omega \leq (4\delta)^{-1} \left\| u^n \right\|_\omega^2 + C\delta \left\| \nabla z^n \right\|^2,$$

and the second term can be absorbed by $\|\nabla z^n\|^2$ for small $\delta > 0$. The first term is absorbed by $|||u|||_R^2$ for small $\alpha > 0$. For the second cross term,

$$\tau \sum_{n=1}^{N} (\tau \nabla \partial_\tau u^n, \tau \nabla \partial_\tau \hat{z}^n) \leq (2\delta)^{-1} \tau \sum_{n=1}^{N} \|\tau \nabla \partial_\tau u^n\|^2 + \delta \tau \sum_{n=1}^{N} \|\nabla z^n\|^2,$$

and we see that these two terms are absorbed analogously with the above. This finishes the proof of (2.6).

It remains to show that

$$|||v, w|||_C \leq C|||u|||_R + C|||u, z|||_D.$$

when $v = u + \alpha \hat{z}$ and $w = -z + \alpha h^2 \partial_\tau u$. We have

$$|||\hat{z}|||_R^2 = \gamma_M \tau \sum_{n=1}^{N} \|z^n\|_\omega^2 + \gamma_1 \tau \sum_{n=1}^{N} \|\tau \nabla \partial_\tau \hat{z}^n\|^2 \leq C\tau \sum_{n=1}^{N} \|\nabla z^n\|^2 \leq C|||0, z|||_D^2,$$

where the Poincaré inequality and the triangle inequality was used for the first and the second term, respectively. Using the Poincaré inequality again, we have

$$\tau \sum_{n=1}^{N} \|z^n\|^2 \leq C|||0, z|||_D^2.$$

The bounds for the terms containing u are trivial. □

Denote by N_h the dimension of V_h. Equations (2.5) define a square linear system of $(2N + 1)N_h$ unknowns, and taking $f^n = 0$ and $q^n = 0$, $n = 1, \ldots, N$, it follows from Proposition 2.1 that $(u, z) = 0$ is the only solution of the corresponding homogeneous system. Thus (2.5) has a unique solution.

3. *A PRIORI* ERROR ESTIMATES

Proposition 3.1. *Suppose that Ω is a convex polyhedral domain and that u is in*

$$H^1(0, T; H_0^1(\Omega)) \cap H^2(0, T; L^2(\Omega)). \tag{3.1}$$

Denote by $\|\cdot\|_*$ the norm in (3.1). Let $(u_h, z_h) \in V_h^{2N+1}$ be the solution of (2.5) with $f = \partial_t u - \Delta u$ and $q = u|_{(0,T)\times\omega}$, and suppose that $f \in C(0, T; L^2(\Omega))$. Then

$$|||\pi_h u - u_h|||_R + |||\pi_h u - u_h, z_h|||_D \leq C(h + \tau) \|u\|_*,$$

where $\pi_h u$ is the orthogonal projection defined by

$$a(\pi_h u, w) = a(u, w), \quad w \in V_h. \tag{3.2}$$

Proof. We use the shorthand notation $\xi_h = \pi_h u - u_h$. By Proposition 2.1 it is enough to show that

$$A_1(\xi_h, w) + A_2((\xi_h, z_h), v) \leq C(h + \tau)|||v, w|||_C \|u\|_*, \quad (v, w) \in V_h^{2N+1}.$$

The point values $u^n = u(t_n)$ satisfy

$$(\partial_t u^n, \phi) + a(u^n, \phi) = (f^n, \phi), \quad n = 1, \dots, N, \ \phi \in H_0^1(\Omega).$$

This implies the following consistency relation

$$\begin{aligned} A_1(u - u_h, w) &= \tau \sum_{n=1}^{N} ((\partial_\tau u^n, w^n) + a(u^n, w^n)) - \tau \sum_{n=1}^{N} (f^n, w^n) \\ &= \tau \sum_{n=1}^{N} (\partial_\tau u^n - \partial_t u^n, w^n), \quad \forall w \in V_h^{N+1}. \end{aligned} \tag{3.3}$$

Using also the orthogonality (3.2), we get

$$\begin{aligned} A_1(\xi_h, w) &= A_1(\pi_h u - u, w) + A_1(u - u_h, w) \\ &= \tau \sum_{n=1}^{N} ((\pi_h - 1)\partial_\tau u^n, w^n) + \tau \sum_{n=1}^{N} (\partial_\tau u^n - \partial_t u^n, w^n). \end{aligned}$$

The Cauchy–Schwarz inequality implies that $A_1(\xi_h, w) \le 2(I_1 + I_2)^{1/2} |||0, w|||_C$ where

$$I_1 = \tau \sum_{n=1}^{N} \|(\pi_h - 1)\partial_\tau u^n\|^2, \quad I_2 = \tau \sum_{n=1}^{N} \|\partial_\tau u^n - \partial_t u^n\|^2.$$

We estimate I_1 by using the approximation properties of π_h, see *e.g.* Theorems 3.16 and 3.18 of [16],

$$\begin{aligned} I_1 &= \tau^{-1} \sum_{n=1}^{N} \left\| \int_{t_{n-1}}^{t_n} (\pi_h - 1)\partial_t u \, \mathrm{d}t \right\|^2 \le \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|(\pi_h - 1)\partial_t u \, \mathrm{d}t\|^2 \\ &\le C h^2 \int_0^T \|\nabla \partial_t u\|^2 \, \mathrm{d}t. \end{aligned}$$

For I_2 we use Taylor's theorem with the integral form of the remainder,

$$\begin{aligned} I_2 &= \tau^{-1} \sum_{n=1}^{N} \left\| \int_{t_{n-1}}^{t_n} \frac{t_n - t}{2} \partial_t^2 u \, \mathrm{d}t \right\|^2 \le \tau^{-1} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (t_n - t)^2 \, \mathrm{d}t \int_{t_{n-1}}^{t_n} \left\| \partial_t^2 u \right\|^2 \mathrm{d}t \\ &\le \tau^2 \int_0^T \left\| \partial_t^2 u \right\|^2 \mathrm{d}t. \end{aligned}$$

Let us now turn to the second bilinear form. We have

$$\begin{aligned} A_2((\xi_h, z_h), v) &= \gamma_M \tau \sum_{n=1}^{N} (\pi_h u^n - u^n, v^n)_\omega + \gamma_0 (h \nabla \pi_h u^0, h \nabla v^0) \\ &\quad + \gamma_1 \tau \sum_{n=1}^{N} (\tau \nabla \partial_\tau \pi_h u^n, \tau \nabla \partial_\tau v^n), \quad \forall v \in V_h^N. \end{aligned} \tag{3.4}$$

Thus $A_2((\xi_h, z_h), v) \leq C(I_3 + I_4 + I_5)^{1/2} |||v, 0|||_C$, where

$$
\begin{aligned}
I_3 &= \tau \sum_{n=1}^{N} \|\pi_h u^n - u^n\|_\omega^2 \leq h^2 \tau \sum_{n=1}^{N} \|\nabla u^n\|^2 \leq Ch^2 \|\nabla u\|^2_{H^1(0,T;L^2(\Omega))}, \\
I_4 &= \left\|h \nabla \pi_h u^0\right\|^2 \leq Ch^2 \|\nabla u\|^2_{H^1(0,T;L^2(\Omega))}, \\
I_5 &= \tau \sum_{n=1}^{N} \|\nabla \pi_h \tau \partial_\tau u^n\|^2 = \tau \sum_{n=1}^{N} \left\| \int_{t_{n-1}}^{t_n} \nabla \pi_h \partial_t u \, \mathrm{d}t \right\|^2 \leq \tau^2 \int_0^T \|\nabla \partial_t u\|^2 \, \mathrm{d}t.
\end{aligned} \tag{3.5}
$$

Here we used the trace inequality in time and the continuity of π_h. □

We recall the following variation of a stability estimate from [15] that was proven in [7].

Theorem 3.2. *Let $\Omega \subset \mathbb{R}^d$ be a convex polyhedron, let $\omega \subset \Omega$ be open and non-empty, and let $0 < \delta < T$. Then there is $C > 0$ such that for all u in the space*

$$
H^1(0,T;H^{-1}(\Omega)) \cap L^2(0,T;H_0^1(\Omega)), \tag{3.6}
$$

it holds that

$$
\|u\|_\delta \leq C(\|u\|_{L^2((0,T)\times\omega)} + \|\partial_t u - \Delta u\|_{L^2(0,T;H^{-1}(\Omega))}),
$$

where $\|\cdot\|_\delta$ is the norm in $C(\delta, T; L^2(\Omega)) \cap L^2(\delta, T; H^1(\Omega)) \cap H^1(\delta, T; H^{-1}(\Omega))$.

For $u_h = (u_h^n)_{n=0}^N \in V_h^{2N+1}$ we define the linear interpolation

$$
\tilde{u}_h(t) = \tau^{-1} \left((t - t_{n-1}) u_h^n + (t_n - t) u_h^{n-1} \right), \quad t \in [t_{n-1}, t_n], \ n = 1, \ldots, N. \tag{3.7}
$$

Observe that $\tilde{u}_h$ is in the space (3.6) and also in $C(0,T;H_0^1(\Omega))$. We are now ready to prove our main result on the convergence of the stabilized finite element method.

Theorem 3.3. *Let $\omega \subset \Omega \subset \mathbb{R}^d$ and $\delta > 0$ be as in Theorem 3.2. Let u, f and (u_h, z_h) be as in Proposition 3.1 and define $\tilde{u}_h$ by (3.7). Suppose that $f \in H^1(0,T;L^2(\Omega))$. Furthermore, in the case $\gamma_1 > 0$ suppose that $\tau = \mathcal{O}(h)$, and in the case $\gamma_1 = 0$ suppose that $\tau = \mathcal{O}(h^2)$. Then*

$$
\|u - \tilde{u}_h\|_\delta \leq Ch \left(\|u\|_* + \|f\|_{H^1(0,T;L^2(\Omega))} \right).
$$

Recall that $\|\cdot\|_*$ is the norm in the space (3.1).

Proof. Let $e = u - \tilde{u}_h$, and define the linear form

$$
\langle r, w \rangle = \int_0^T (\partial_t e, w) + a(e, w) \, \mathrm{d}t, \quad w \in L^2(0,T;H_0^1(\Omega)).
$$

By Theorem 3.2 it is enough to show the following two inequalities

$$
\|e\|_{L^2((0,T)\times\omega)} \leq Ch \|u\|_*, \tag{3.8}
$$

$$
\langle r, w \rangle \leq Ch \left(\|u\|_* + \|f\|_{L^2((0,T)\times\Omega)} \right) \|w\|_{L^2(0,T;H_0^1(\Omega))}. \tag{3.9}
$$

Let us begin with (3.8). We define the projection on the piecewise constant functions

$$\pi_0 v(t) = v(t^n), \quad t \in (t_{n-1}, t_n],\ n = 1, \dots, N.$$

Observe that

$$\|\pi_0 v - v\|_{L^2(0,T)} \leq \tau \|\partial_t v\|_{L^2(0,T)}, \quad v \in H^1(0,T).$$

We have

$$\|e\|^2_{L^2((0,T)\times\omega)} \leq C(h^2+\tau^2)\|u\|^2_* + \int_0^T \|\pi_0\pi_h u - \tilde{u}_h\|^2_\omega \,\mathrm{d}t,$$

and

$$\begin{aligned}\int_0^T \|\pi_0\pi_h u - \tilde{u}_h\|^2_\omega \,\mathrm{d}t &\leq \int_0^T \|\pi_0\pi_h u - \pi_0\tilde{u}_h\|^2_\omega \,\mathrm{d}t + \int_0^T \|\pi_0\tilde{u}_h - \tilde{u}_h\|^2_\omega \,\mathrm{d}t \\ &= \tau\sum_{n=1}^N \|\pi_h u^n - u_h^n\|^2_\omega + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\pi_0\tilde{u}_h - \tilde{u}_h\|^2_\omega \,\mathrm{d}t.\end{aligned}$$

Here the first term is bounded by $|||\pi_h u - u_h|||^2_R$, and we use the identity

$$\tilde{u}_h = u_h^n + (t-t_n)\partial_\tau u_h^n \tag{3.10}$$

to estimate the second one as follows

$$\begin{aligned}\sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\pi_0\tilde{u}_h - \tilde{u}_h\|^2 \,\mathrm{d}t &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|(t_n - t)\partial_\tau u_h^n\|^2 \,\mathrm{d}t \leq \tau\sum_{n=1}^N \|\tau\partial_\tau u_h^n\|^2 \\ &\leq \tau\sum_{n=1}^N \|\tau\partial_\tau(\pi_h u^n - u_h^n)\|^2 + \tau\sum_{n=1}^N \|\tau\partial_\tau\pi_h u^n\|^2.\end{aligned}$$

As $\tau = \mathcal{O}(h)$, the first term above is bounded by $|||\pi_h u - u_h, 0|||^2_D$, and the second term is bounded by $\tau^2\|u\|^2_*$. The inequality (3.8) follows from Proposition 3.1.

We turn to (3.9), and define the piecewise constant function defined by local time averages

$$\overline{w}(t) = \tau^{-1}\int_{t_{n-1}}^{t_n} w \,\mathrm{d}t, \quad t \in (t_{n-1}, t_n],\ n = 1, \dots, N.$$

We have

$$\int_0^T (\partial_t u, w) + a(u,w) \,\mathrm{d}t = \int_0^T (f, w) \,\mathrm{d}t = \int_0^T (f - \pi_0 f, w) \,\mathrm{d}t + \tau\sum_{n=1}^N (f^n, \overline{w}),$$

and using the identity (3.10) and the orthogonality (3.2),

$$
\begin{aligned}
-\int_0^T (\partial_t \tilde{u}_h, w) + a(\tilde{u}_h, w)\,\mathrm{d}t &= -\tau \sum_{n=1}^{N} (\partial_\tau u_h^n, \overline{w}) - \int_0^T a(\tilde{u}_h, \pi_h w)\,\mathrm{d}t \\
&= -\tau \sum_{n=1}^{N} (\partial_\tau u_h^n, \overline{w}) - \tau \sum_{n=1}^{N} a(u_h^n, \pi_h \overline{w}) - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (t - t_n)\, a(\partial_\tau u_h^n, \pi_h w)\,\mathrm{d}t.
\end{aligned}
$$

As u_h satisfies (2.5), it holds that

$$
\begin{aligned}
\langle r, w \rangle = &\int_0^T (f - \pi_0 f, w)\,\mathrm{d}t + \tau \sum_{n=1}^{N} (f^n, \overline{w} - \pi_h \overline{w}) - \tau \sum_{n=1}^{N} (\partial_\tau u_h^n, \overline{w} - \pi_h \overline{w}) \\
&- \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (t - t_n)\, a(\partial_\tau u_h^n, \pi_h w)\,\mathrm{d}t. \qquad (3.11)
\end{aligned}
$$

We have

$$
\begin{aligned}
\int_0^T (f - \pi_0 f, w)\,\mathrm{d}t &\le \tau \left\| f \right\|_{H^1(0,T;L^2(\Omega))} \left\| w \right\|_{L^2((0,T)\times\Omega)}, \\
\tau \sum_{n=1}^{N} (f^n, \overline{w} - \pi_h \overline{w}) &\le Ch \left\| f \right\|_{H^1(0,T;L^2(\Omega))} \left\| w \right\|_{L^2(0,T;H^1(\Omega))}.
\end{aligned}
$$

Moreover,

$$
\tau \sum_{n=1}^{N} (\partial_\tau u_h^n, \overline{w} - \pi_h \overline{w}) \le Ch \left\| u \right\|_{H^2(0,T;L^2(\Omega))} \left\| w \right\|_{L^2(0,T;H^1(\Omega))},
$$

where we used Proposition 3.1, after observing that

$$
h^2 \tau \sum_{n=1}^{N} \left\| \partial_\tau u_h^n \right\|^2 \le \left| \left| \left| u_h - \pi_h u, 0 \right| \right| \right|_D^2 + h^2 \left\| u \right\|_*^2.
$$

Finally,

$$
\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (t - t_n)\, a(\partial_\tau u_h^n, \pi_h w)\,\mathrm{d}t \le \tau \left(\tau \sum_{n=1}^{N} \left\| \nabla \partial_\tau u_h^n \right\|^2 \right)^{\frac{1}{2}} \left\| w \right\|_{L^2(0,T;H^1(\Omega))},
$$

and using the triangle inequality and (3.5),

$$
\tau \sum_{n=1}^{N} \left\| \tau \nabla \partial_\tau u_h^n \right\|^2 \le \tau \sum_{n=1}^{N} \left\| \tau \nabla \partial_\tau (u_h^n - \pi_h u^n) \right\|^2 + C\tau^2 \int_0^T \left\| \nabla \partial_t u \right\|^2 \mathrm{d}t.
$$

Observe that

$$\tau \sum_{n=1}^{N} \|\tau \nabla \partial_\tau (u_h^n - \pi_h u^n)\|^2 \leq C \begin{cases} |||u_h - \pi_h u|||_R^2, & \gamma_1 > 0, \\ |||u_h - \pi_h u, 0|||_D^2, & \tau = \mathcal{O}(h^2). \end{cases}$$

The inequality (3.9) follows from Proposition 3.1. □

If $\gamma_1 = 0$ and $\tau = \mathcal{O}(h)$ then Theorem 3.3 does not predict optimal convergence. Indeed, in this case the bound in the last step becomes

$$\tau \sum_{n=1}^{N} \|\tau \nabla \partial_\tau (u_h^n - \pi_h u^n)\|^2 \leq C h^{-1} |||u_h - \pi_h u, 0|||_D^2.$$

This then leads to a convergence of order $O(h^{\frac{1}{2}} + \tau^{\frac{1}{2}})$ using Proposition 3.1.

3.1. The case of perturbations in data

Thanks to the Lipschitz stability of Theorem 3.2 the extension of the above analysis to the case where the data is perturbed is straightforward. Indeed, assume that instead of $(q^n, f^n)_{n=1}^N$ in (2.3) we have at are disposal the perturbed data $(\tilde{q}^n, \tilde{f}^n)_{n=1}^N$,

$$\tilde{q}^n = q^n + e_q^n, \quad \tilde{f}^n = f^n + e_f^n$$

with $e_q^n \in L^2(\omega)$ and $e_f^n \in H^{-1}(\Omega)$. Then augmenting the proofs of Proposition 3.1 and Theorem 3.3 with a standard perturbation argument, we obtain the following result

Theorem 3.4. *Let $\omega \subset \Omega \subset \mathbb{R}^d$ and $\delta > 0$ be as in Theorem 3.2. Let u, f be as in Proposition 3.1, let (u_h, z_h) be the solution to (2.5) with q^n and f^n replaced by $\tilde{q}^n$ and $\tilde{f}^n$, and define $\tilde{u}_h$ by (3.7). Suppose that $f \in H^1(0,T;L^2(\Omega))$. Furthermore, in the case $\gamma_1 > 0$ suppose that $\tau = \mathcal{O}(h)$, and in the case $\gamma_1 = 0$ suppose that $\tau = \mathcal{O}(h^2)$. Then*

$$\|u - \tilde{u}_h\|_\delta \leq Ch \left(\|u\|_* + \|f\|_{H^1(0,T;L^2(\Omega))} \right) + \mathcal{E}_{q,f},$$

where

$$\mathcal{E}_{q,f} := C \left(\tau \sum_{n=1}^{N} \left(\|e_q^n\|_\omega^2 + \|e_f^n\|_{H^{-1}(\Omega)}^2 \right) \right)^{\frac{1}{2}}.$$

Proof. The proof follows from minor modifications of the proofs of Proposition 3.1 and Theorem 3.3. We will only give some pointers to the modifications necessary to include the perturbations. First we note that the perturbed data modifies the consistency in equations (3.3) and (3.4) to

$$\begin{aligned} A_1(u - u_h, w) &= \tau \sum_{n=1}^{N} \left((\partial_\tau u^n, w^n) + a(u^n, w^n) \right) - \tau \sum_{n=1}^{N} (f^n + e_f^n, w^n) \\ &= \tau \sum_{n=1}^{N} \left[(\partial_\tau u^n - \partial_t u^n, w^n) - (e_f^n, w^n) \right], \quad \forall w \in V_h^{N+1} \end{aligned} \tag{3.12}$$

and

$$A_2((\xi_h, z_h), v) = \gamma_M \tau \sum_{n=1}^{N} (\pi_h u^n - u^n - e_q^n, v^n)_\omega + \gamma_0 (h\nabla \pi_h u^0, h\nabla v^0) + \gamma_1 \tau \sum_{n=1}^{N} (\tau \nabla \partial_\tau \pi_h u^n, \tau \nabla \partial_\tau v^n), \quad \forall v \in V_h^N. \tag{3.13}$$

The Cauchy–Schwarz inequality implies that

$$\tau \sum_{n=1}^{N} (e_f^n, w^n) + \gamma_M \tau \sum_{n=1}^{N} (e_q^n, v^n)_\omega \leq \mathcal{E}_{q,f} |||v, w|||_C.$$

Proceeding as in Proposition 3.1 then leads to the bound

$$|||\pi_h u - u_h|||_R + |||\pi_h u - u_h, z_h|||_D \leq C(h + \tau) \|u\|_* + \mathcal{E}_{q,f}. \tag{3.14}$$

In Theorem 3.3 this leads to modifications of the bounds (3.8)–(3.9),

$$\|e\|_{L^2((0,T)\times\omega)} \leq Ch \|u\|_* + \mathcal{E}_{q,f}, \tag{3.15}$$

$$\langle r, w \rangle \leq \left(Ch \left(\|u\|_* + \|f\|_{L^2((0,T)\times\Omega)} \right) + \mathcal{E}_{q,f} \right) \|w\|_{L^2(0,T;H_0^1(\Omega))}. \tag{3.16}$$

Inequality (3.15) is an immediate consequence of (3.14). For the residual bound (3.16) we must once again take into account the lack of consistency, leading to a modification in equation (3.11),

$$\langle r, w \rangle = \int_0^T (f - \pi_0 f, w) \, \mathrm{d}t + \tau \sum_{n=1}^{N} ((f^n, \overline{w} - \pi_h \overline{w}) - (e_f^n, \pi_h \overline{w})) - \tau \sum_{n=1}^{N} (\partial_\tau u_h^n, \overline{w} - \pi_h \overline{w}) - \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (t - t_n) \, a(\partial_\tau u_h^n, \pi_h w) \, \mathrm{d}t. \tag{3.17}$$

We proceed using the Poincaré and Cauchy-Schwarz inequalities and stability of the projection in the perturbation term,

$$\tau \sum_{n=1}^{N} (e_f^n, \pi_h \overline{w}) \leq \mathcal{E}_{0,f} \|w\|_{L^2(0,T;H^1(\Omega))},$$

followed by an application of the perturbed bound (3.14). The claim follows by applying Theorem 3.2 to the error and the associate perturbation equation, using (3.15) and (3.16). □

This is a similar result as one would obtain for a well-posed problem. In particular, the mesh sizes h and τ can be chosen independently of the size of the perturbations e_q^n and e_f^n, the constants γ_M, γ_0 and γ_1 do not depend on the size of e_q^n and e_f^n, and the method converges to the exact solution u when the mesh sizes and the perturbations converge to zero.

4. Computational examples

The main objectives of the computational examples are twofold.

TABLE 1. Convergence with $\gamma_M = \gamma_0 = 1$ and $\gamma_1 = 0$ using the MINRES method. The error is $\|u(T) - u_h^N\|_{L^2(\Omega)}$. *Left.* Order 1 convergence in h with $N = 16$. *Right.* Order 1/2 convergence in τ with $N_h = 200$.

h	0.02	0.01	0.005
Error	0.224	0.119	0.043

τ	0.004	0.002	0.001
Error	0.104	0.073	0.048

(1) First we verify that the predicted reduction in convergence order to $O(h^{\frac{1}{2}} + \tau^{\frac{1}{2}})$ for $\gamma_1 = 0$ and $\tau = \mathcal{O}(h)$ indeed takes place, even in a simple model case.
(2) Then we confirm that the situation is rectified for $\gamma_1 > 0$.

The Euler–Lagrange equation (2.5) form a non-singular, symmetric system of $(2N + 1)N_h$ linear equations. We emphasize that the system is not positive definite. In principle, it can be solved using off-the-shelf methods, for example the MINRES method [28].

We implemented this straightforward strategy in the case that $\gamma_1 = 0$, and verified that the convergence order in space is that predicted by Theorem 3.3. For the convergence order in time we verify that failure to meet the condition $\tau = \mathcal{O}(h^2)$ indeed leads to suboptimal convergence. We observe $\mathcal{O}(\tau^{\frac{1}{2}})$ convergence under refinement of τ in the regime where $\tau = \mathcal{O}(h)$. In all our computational examples Ω is the unit interval $(0, 1)$, $\omega = (a, 1 - a)$, $a = 0.2$, and we use a regular mesh on Ω. Moreover, the function u is of the form

$$u(t, x) = e^{-\pi^2 k^2 t} \sin(\pi k x), \quad k = 1, 2. \tag{4.1}$$

Computations for $k = 2$ and $T = 0.02$ are summarized in Table 1. We also verified that the computations diverge when no regularization is introduced, that is, when $\gamma_0 = 0$. In these computations we used the MINRES implementation of SciPy with the default parameters [19], and the initial guess was set to zero. The convergence is typically slow, requiring thousands of iterations.

The remaining examples will exploit the structure of (2.5) to reduce the memory requirements of the solution algorithm. The classical steepest descent approach will be applied, using the adjoint to evaluate the gradient (see for instance [33] for a discussion of the approach in the context of 4DVAR).

4.1. The Euler–Lagrange equations as a system of two coupled heat equations

An attractive feature of the regularization in (2.3) is that it acts only on the primal variable u. This leads to the one-way coupling in (2.5), that is, the dual variable z does not appear in the equation involving A_1. We present next a method solving (2.5) that is based on the one-way coupling.

Note that the first equation in (2.5), that is,

$$\tau \sum_{n=1}^{N} ((\partial_\tau u^n, w^n) + a(u^n, w^n)) = \tau \sum_{n=1}^{N} (f^n, w^n) \tag{4.2}$$

is simply a discretization of the heat equation (1.4). Let us next interpret the second equation in (2.5) as a discretization of a heat equation for z. Observe that, setting $z^{N+1} = 0$, we obtain

$$\tau \sum_{n=1}^{N} (\partial_\tau v^n, z^n) = -\tau \sum_{n=1}^{N} (v^n, \partial_\tau z^{n+1}) - (v^0, z^1).$$

Thus choosing $v^0 = 0$ in (2.5) for the moment, we see that z satisfies

$$\tau \sum_{n=1}^{N} \left(-(v^n, \partial_\tau z^{n+1}) + a(v^n, z^n) \right) = \gamma_M \tau \sum_{n=1}^{N} (q^n - u^n, v^n)_\omega - \gamma_1 \tau \sum_{n=1}^{N} (\tau \nabla \partial_\tau u^n, \tau \nabla \partial_\tau v^n), \tag{4.3}$$

and this can be interpreted as a discretization of

$$-\partial_t z - \Delta z = \gamma_M (q - u) 1_\omega.$$

Here 1_ω is the indicator function of ω, that is, $1_\omega(x) = 1$ if $x \in \omega$ and $1_\omega(x) = 0$ otherwise. Note that, when rescaled by τ^{-2}, the second term on the right-hand side of (4.3) is a discretization of $\int_0^T (\nabla \partial_t u, \nabla \partial_t v)\, \mathrm{d}t$. Taking now $v^n = 0$, $n = 1, \ldots, N$, in (2.5) we get the additional constraint

$$\gamma_0 (h \nabla u^0, h \nabla v^0) - \gamma_1 \tau (\tau \nabla \partial_\tau u^1, \nabla v^0) - (z^1, v^0) = 0.$$

Define $U(\phi)$ to be the solution of (4.2) with $u^0 = \phi$, and $Z(\phi)$ the solution of (4.3) with $z^{N+1} = 0$ and $u = U(\phi)$. Observe that these can be easily computed by using time stepping. Furthermore, define the function

$$\mathcal{C}(\phi, \psi) = \gamma_0 (h \nabla U^0(\phi), h \nabla \psi) - \gamma_1 \tau (\tau \nabla \partial_\tau U^1(\phi), \nabla \psi) - (Z^1(\phi), \psi), \quad \psi \in V_h.$$

Then $(u, z) = (U(\phi), Z(\phi))$ solves (2.5) if and only if

$$\mathcal{C}(\phi, \psi) = 0, \quad \psi \in V_h. \tag{4.4}$$

We will use a gradient descent type method to solve (4.4). Starting from an initial guess $\phi_0 \in V_h$, we define the iteration

$$(\phi_{m+1}, \psi) = (\phi_m, \psi) - \alpha \mathcal{C}(\phi_m, \psi), \quad \psi \in V_h, \tag{4.5}$$

where $\alpha > 0$ is a step size. The system (4.5) is a discretization of the differential equation

$$\Phi(0) = \phi_0, \quad (\partial_s \Phi(s), \psi) = -\mathcal{C}(\Phi(s), \psi), \quad \psi \in V_h, \tag{4.6}$$

and its use to solve (4.4) is justified by the following lemma.

Lemma 4.1. *Let $\phi_0 \in V_h$ and define a one parameter family $\Phi(s)$, $s \geq 0$, in V_h by (4.6). Let (u_h, z_h) be the solution of (2.5). Then $\Phi(s)$ converges to u_h^0 as $s \to \infty$.*

Proof. For each $s \geq 0$ it holds by definition that $u(s) = U(\Phi(s))$ and $z(s) = Z(\Phi(s))$ satisfy (4.2) and (4.3), respectively. Hence

$$\partial_s \mathcal{L}(u, z) = (\partial_u L, \partial_s u) + (\partial_z L, \partial_s z) = \mathcal{C}(\Phi, \partial_s u^0) = \mathcal{C}(\Phi, \partial_s \Phi) = -\left\| \partial_s \Phi \right\|^2.$$

Equation (4.2) implies also that

$$\mathcal{L}(u, z) = \frac{1}{2} \gamma_M \tau \sum_{n=1}^{N} \left\| u^n - q^n \right\|_\omega^2 + \frac{1}{2} \gamma_0 \left\| h \nabla \Phi \right\|^2 + \frac{1}{2} \gamma_1 \tau \sum_{n=1}^{N} \left\| \tau \nabla \partial_\tau u^n \right\|^2.$$

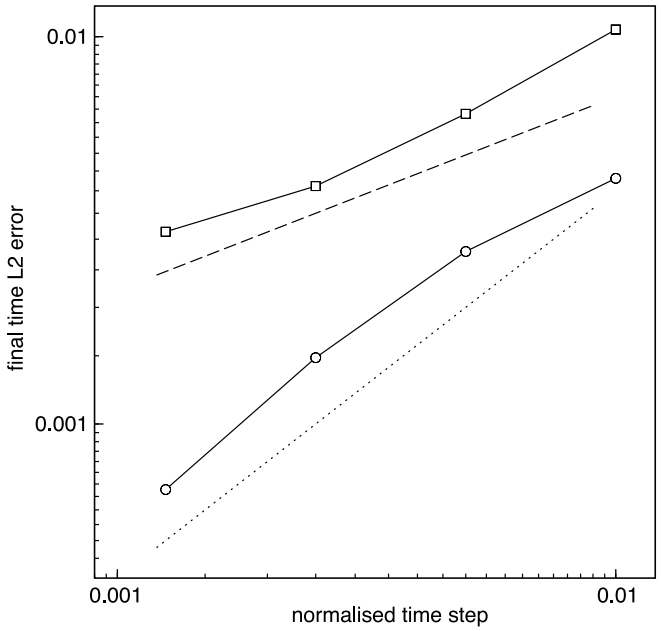

FIGURE 1. The effect of regularization on the convergence in τ. The convergence is of order $1/2$ (slope of dashed reference line) when $\gamma_1 = 0$ (data with square markers) and of order 1 (slope of dotted reference line) when $\gamma_1 = 1$ (data with circle markers). Here $\gamma_M = \gamma_0 = 1$, $h = 10^{-2}$, and the error is $\left\|u(T) - u_h^N\right\|_{L^2(\Omega)}$.

As $\mathcal{L}$ is non-negative and decreasing along the family $(u(s), z(s))$, it follows that $\partial_s \mathcal{L}(u, z) \to 0$ as $s \to \infty$. Hence also $\partial_s \Phi \to 0$ as $s \to \infty$, and the differential equation (4.6) implies that the limit $\phi_\infty = \lim_{s\to\infty} \Phi(s)$ exists and satisfies (4.4). By the discussion preceding the proof, we have $\phi_\infty = u_h^0$. □

We will use the above gradient descent method in the computational examples below and assume that the initial guess ϕ_0 is a small perturbation of $u(0)$. Such an assumption can be relevant for many data assimilation applications. Indeed, it is typical that new observations need to be incorporated into the state of the system, and the current state can then be used as an initial guess.

4.2. The effect of regularization on the convergence in τ

We verified that the presence of the additional regularization in the case $\gamma_1 > 0$ leads to the improved convergence rate in τ as predicted by Theorem 3.3. Indeed, in the computations summarized in Figure 1, the convergence is of order $1/2$ when $\gamma_1 = 0$ and of order 1 when $\gamma_1 = 1$. Here $\gamma_M = \gamma_0 = 1$, $h = 10^{-2}$, u is of the form (4.1) with $k = 1$, and $T = 0.1$. We used the gradient descent method with the initial guess $\phi_0 = v + h$ where v is the interpolation of $u(0)$ on V_h. The step size in (4.5) was taken $\alpha = 0.1$ and the iteration (4.5) was terminated when $\left\|z^1\right\|$ started to increase.

4.3. Sensitivity to the choice of γ_0 and γ_1

In all the numerical experiments above we have taken the parameters γ_0 and γ_1 to be either one or zero. This was to avoid special effects that can appear due to parameter tuning. In a final numerical experiment we verified that the method is not sensitive to the particular choices of the constants $\gamma_0, \gamma_1 > 0$. The conclusion of the study is that the method is robust for a wide range of choices of γ_0 and γ_1, including $\gamma_0 = \gamma_1 = 1$. We observed that choosing both parameters large resulted in solutions that were over regularized and yielded

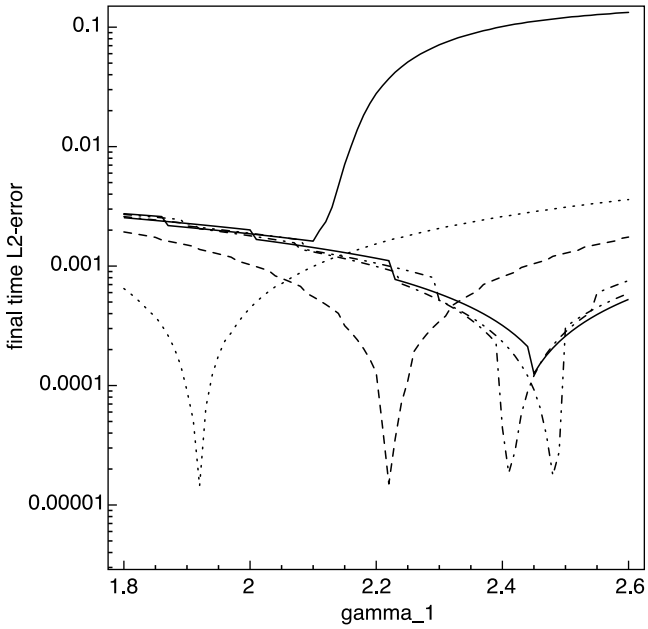

FIGURE 2. The error for various choices of the constants γ_0, γ_1. Here $\gamma_M = 1$, $h = \tau = 10^{-2}$ and the error is $\left\|u(T) - u_h^N\right\|_{L^2(\Omega)}$. For each $0.1 \leq \gamma_0 \leq 1.2$, the method is robust for a large range in γ_1. There also is an optimal value of γ_1 for each such γ_0. However, this is mesh dependent and it is not clear if the phenomenon can be exploited in practice. ($\gamma_0 = 0.1$ – dotted line; $\gamma_0 = 0.2$ – dashed line; $\gamma_0 = 0.6$ – dash/dotted line; $\gamma_0 = 1.0$ – dash/doubledotted line; $\gamma_0 = 1.2$ – doubledash/doubledotted line; $\gamma_0 = 1.5$ – filled line.)

suboptimal accuracy compared to lower values of the parameters. See the filled line of Figure 2 for an example. We also observed that there are certain "sweet spot" combinations of values of γ_0 and γ_1 for which the errors are orders of magnitude smaller than for the neighbouring parameter combinations. These optimal parameter combinations however did not appear to be stable under mesh refinement and it is unclear if this effect can be of any use in practice. The computations are summarized in Figure 2, with particular focus on the parameter interval where the optimal parameter choices appeared. Here $h = \tau = 10^{-2}$ and the other choices are as in the previous example.

Acknowledgements. LO was supported in part by the EPSRC grants EP/L026473/1 and EP/P01593X/1. EB was supported in part by the EPSRC grant EP/P01576X/1.

REFERENCES

[1] R.N. Bannister, A review of forecast error covariance statistics in atmospheric variational data assimilation. II: modelling the forecast error covariance statistics. *Quaterly J. Roy. Meteorol. Soc.* **134** (2008) 1971–1996.

[2] E. Bécache, L. Bourgeois, L. Franceschini and J. Dardé, Application of mixed formulations of quasi-reversibility to solve ill-posed problems for heat and wave equations: the 1D case. *Inverse Probl. Imaging* **9** (2015) 971–1002.

[3] F. Boyer, F. Hubert and J. Le Rousseau, Uniform controllability properties for space/time-discretized parabolic equations. *Numer. Math.* **118** (2011) 601–661.

[4] E. Burman, Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: Elliptic equations. *SIAM J. Sci. Comput.* **35** (2013) A2752–A2780.

[5] E. Burman, Error estimates for stabilized finite element methods applied to ill-posed problems. *C. R. Math. Acad. Sci. Paris* **352** (2014) 655–659.

[6] E. Burman, Stabilised finite element methods for ill-posed problems with conditional stability, in Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Vol. 114 of *Lect. Notes Comput. Sci. Eng.* Springer, Cham (2016) 93–127.

[7] E. Burman and L. Oksanen, Data Assimilation for the Heat Equation Using Stabilized Finite Element Methods. *Numer. Math.* **139** (2016) 505–528.

[8] E. Burman, M.G. Larson and P. Hansbo, *Solving ill-Posed Control Problems by Stabilized Finite Element Methods: An Alternative to Tikhonov Regularization.* Technical report (2016)

[9] C. Castro, N. Cîndea and A. Münch, Controllability of the linear one-dimensional wave equation with inner moving forces. *SIAM J. Control Optim.* **52** (2014) 4027–4056.

[10] N. Cîndea and A. Münch, Inverse problems for linear hyperbolic equations using mixed formulations. *Inverse Probl.* **31** (2015) 075001.

[11] N. Cîndea and A. Münch, A mixed formulation for the direct approximation of the control of minimal L^2-norm for linear type wave equations. *Calcolo* **52** (2015) 245–288.

[12] N. Cîndea and A. Münch, Simultaneous reconstruction of the solution and the source of hyperbolic equations from boundary measurements: a robust numerical approach. *Inverse Probl.* **32** (2016) 115020.

[13] N. Cîndea, E. Fernández-Cara and A. Münch, Numerical controllability of the wave equation through primal methods and Carleman estimates. *ESAIM: COCV.* **19** (2013) 1076–1108.

[14] P. Courtier *et al.*, A strategy for operational implementation of 4D-Var, using an incremental approach. *Quaterly J. Roy. Meteorol. Soc.* **120** (1994) 1367–1387.

[15] O.Y. Èmanuilov, Controllability of parabolic equations. *Mat. Sb.* **186** (1995) 109–132.

[16] A. Ern and J.-L. Guermond, Theory and Practice of Finite Elements. Vol. 159 of *Applied Mathematical Sciences.* New York (2004).

[17] V. Isakovm, Inverse Problems For Partial Differential Equations, 2nd edn. Vol. 127 of *Applied Mathematical Sciences.* Springer, New York (2006).

[18] S.E. Jenkins, C.J. Budd, M.A. Freitag and N.D. Smith, The effect of numerical model error on data assimilation. *J. Comput. Appl. Math.* **290** (2015) 567–588.

[19] E. Jones, T. Oliphant, P. Peterson *et al.*, SciPy: open source scientific tools for python (2001). Online; version 0.19.1 [accessed 19/07/17].

[20] M.V. Klibanov, Estimates of initial conditions of parabolic equations and inequalities via lateral Cauchy data. *Inverse Probl.* **22** (2006) 495–514.

[21] M.V. Klibanov, Carleman estimates for global uniqueness, stability and numerical methods for coefficient inverse problems. *J. Inverse Ill-Posed Probl.* **21** (2013) 477–560.

[22] M.V. Klibanov, Carleman estimates for the regularization of ill-posed Cauchy problems. *Appl. Numer. Math.* **94** (2015) 46–74.

[23] M.V. Klibanov and P.G. Danilaev, Solution of coefficient inverse problems by the method of quasi-reversibility. *Dokl. Akad. Nauk SSSR* **310** (1990) 528–532.

[24] R. Lattès and J.-L. Lions, Méthode de quasi-réversibilité et applications. Travaux et Recherches Mathématiques, No. 15. Dunod, Paris (1967).

[25] F.X. Le Dimet and O. Talagrand, Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A* **38** (1986) 97–110.

[26] J.-L. Lions, Équations différentielles opérationnelles et problèmes aux limites. Die Grundlehren der mathematischen Wissenschaften, Bd.111, Springer-Verlag, Berlin-Göttingen-Heidelberg (1961).

[27] A. Münch and D.A. Souza, A mixed formulation for the direct approximation of L^2-weighted controls for the linear heat equation. *Adv. Comput. Math.* **42** (2016) 85–125.

[28] C.C. Paige and M.A. Saunders, Solutions of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.* **12** (1975) 617–629.

[29] Y. Sasaki, Some basic formalisms in numerical variational analysis. *Mon. Weather Rev.* **98** (1970).

[30] M. Tadi, M.V. Klibanov and W. Cai, An inversion method for parabolic equations based on quasireversibility. *Comput. Math. Appl.* **43** (2002) 927–941.

[31] O. Talagrand, Initialisation d'un modèle numérique d'atmosphère à partir de données distribuées dans le temps, in Computing Methods in Applied Sciences and Engineering, in *Proc. Third Internat. Sympos., Versailles, 1977, II.* Vol. 91 of *Lecture Notes in Physics.* Springer, Berlin-New York (1979) 217–231.

[32] O. Talagrand, On the mathematics of data assimilation. *Tellus* **33** (1981) 321–339.

[33] O. Talagrand and P. Courtier, Variational assimilation of meteorological observations with the adjoint vorticity equation. i: theory. *Q. J. Royal Meteorol. Soc.* **113** (1987) 1311–1328.

[34] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems. Vol. 25 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin (1997).

[35] Y.B. Wang, J. Cheng, J. Nakagawa and M. Yamamoto. A numerical method for solving the inverse heat conduction problem without initial value. *Inverse Probl. Sci. Eng.* **18** (2010) 655–671.

[36] M. Yamamoto, Carleman estimates for parabolic equations and applications. *Inverse Probl.* **25** 123013 (2009).