

Weakly consistent regularisation methods for ill-posed problems

Erik Burman and Lauri Oksanen

Abstract These notes were written as support for a 9 hours course at the Institut Henri Poincaré in September 2016. The course was divided in three parts. In the first part, which is not included herein, the aim was to first recall some basic aspects of stabilised finite element methods for convection–diffusion problems. We focus entirely on the second and third parts which were dedicated to ill-posed problems and their approximation using stabilised finite element methods. First we introduce the concept of conditional well-posedness and conditional stability. Then we consider the elliptic Cauchy-problem and a data assimilation problem in a unified setting and show how stabilised finite element methods may be used to derive error estimates that are consistent with the stability properties of the problem and the approximation properties of the finite element space. In the third part we extend the result to a data assimilation problem subject to the heat equation.

1 Introduction

In these notes we will give an overview of some recent work on finite element methods for ill-posed problems. For well-posed problems it is known that in the presence of non-symmetric operators approximation using Galerkin finite element methods may have poor accuracy, due to the lack of H^1 -coercivity. A popular remedy is then to add some stabilising terms that should be balanced in such a way that they cure the stability issue, but vanish quickly enough under mesh-refinement so that optimal error estimates can be obtained. For ill-posed problems on the other hand the state of the art is to add some regularising terms on the continuous level to obtain a well-posed continuous problem that can then typically be discretised using standard finite element methods. Here our aim is to attempt to make the ideas from the former class

Erik Burman, Lauri Oksanen
Department of Mathematics, University College London, London, UK–WC1E 6BT, United Kingdom, e-mail: e.burman@ucl.ac.uk, l.oksanen@ucl.ac.uk

of problems carry over to the ill-posed case, using weakly consistent regularisation that is defined on the discrete level. Indeed prior to discretisation no regularisation is applied, instead the ill-posed problem and associated data are discretised in the form of a minimisation problem, where some suitable distance between the discrete solution and the measured data is minimised under the constraint of the discrete form of the partial differential equation. Regularisation terms may then be devised that are in some sense the minimal choice necessary to achieve a well-posed discrete system. To analyse the resulting approximation we rely on conditional stability estimates for the continuous problem typically obtained through Carleman estimates.

Compared to the state of the art methods such as the quasireversibility method by Lions' and Lattes (and the recent improvements on this technique by Bourgeois and Darde [3, 4]) or the penalty method by Kohn and Vogelius the present framework has some interesting features. Since no regularised continuous problem is involved the only (nontrivial) regularisation parameter present is the mesh-size. This is not the case for more traditional methods where the discretisation parameter and the regularisation parameter must be matched carefully, or as is usually assumed, the mesh size is chosen substantially smaller than the regularisation parameter. Maybe more importantly in the present framework the regularisation is independent of the stability of the underlying physical problem. The method can have optimal convergence order with respect to the computational mesh-size and the stability of the physical problem. If regularisation and discretisation errors are to be balanced in the framework of conventional Tikhonov regularisation this appears to inevitably lead to a nontrivial relation between the regularisation, the mesh size and the specific form of the stability of the quantity of which one wishes to control the errors.

With the recent increased understanding of the stability properties of ill-posed problems we believe that these considerations are important. For instance problems with Hölder, or even Lipschitz stability will have that precise order reproduced for the convergence order of the approximation error. This means in particular that ill-posed problem allowing for Lipschitz stability estimates for some quantity may allow for the same error estimates, for this quantity, as in the case where the solution of a well-posed problem is approximated. This is a very pleasing result and although it may seem expected, to the best of our knowledge there exists no results in the literature reporting on such estimates.

The paper consists of two main chapters. In the first we consider stationary ill-posed elliptic problems, such as the elliptic Cauchy problem and the so-called data assimilation problem, where measured data is available in some subdomain of the bulk, but not on the boundary. For these problems interior estimates with Hölder stability are known to hold and we show how to make these estimates translate into error estimates for the computational method. In the second chapter we consider the extension of these ideas to a data assimilation problem subject to the heat equation. In this case a Lipschitz-continuous stability estimate holds for the reconstruction of the solution away from the (unknown) initial datum. Also in this case we show, in a space semi-discretised framework, error estimates that reflect the stability of the physical problem. In this second case the estimates obtained are optimal compared to the approximation order of the finite element space.

2 Preliminary technical results

In this section we will introduce the geometrical setting of the problems that we will consider, the associated finite element spaces and some technical results, including useful inequalities and approximation results. We will stay in the simplest of settings, considering only piecewise affine finite element spaces.

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be a convex polygonal (polyhedral) domain, with boundary $\partial\Omega$ and outward pointing normal n . By $\mathcal{T} = \{T\}$ we denote a quasi-uniform decomposition of Ω in simplices T such that the intersection of two simplices in \mathcal{T} is either the empty set, a shared vertex, a shared face or a shared edge. We also introduce the mesh parameter associated to \mathcal{T} , $h_T = \text{diam}(T)$ where the diameter of T is defined as the diameter of the smallest ball circumscribing T . Setting $h = \max_{T \in \mathcal{T}} h_T$ we consider the family of tessellations $\{\mathcal{T}\}_h$ indexed by h . The simplices are shape regular in the sense that the ratio between the smallest circumscribed ball and the largest inscribed ball of any $T \in \mathcal{T}$ is bounded uniformly, with a constant independent of h . The boundary of T will be denoted ∂T with outward pointing normal n_T . We denote the set of element boundary faces by \mathcal{F} and let \mathcal{F}_i and \mathcal{F}_b denote the set of interior or boundary faces respectively and to each interior face we associate a normal n_F that is fixed, but with arbitrary orientation. The normal on faces on the boundary will be chosen pointing outwards.

We will consider different subsets of $H^1(\Omega)$ for the weak formulations. First let $V := H^1(\Omega)$ and $V^0 := H_0^1(\Omega)$ then if $\Gamma \subset \partial\Omega$ is a connected subset of $\partial\Omega$ we denote

$$V^\Gamma := \{v \in H^1(\Omega) : v|_\Gamma = 0\}.$$

To introduce the discrete spaces we assume that Γ consists of a subset of boundary faces $\mathcal{F}_\Gamma := \{F \in \mathcal{F}_b : \overline{F} \cap \Gamma = F\}$. We define the finite dimensional spaces

$$V_h = \{v_h \in V : v|_T \in \mathbb{P}_1(T); \forall T \in \mathcal{T}\},$$

with $\mathbb{P}_1(T)$ the set of polynomials of degree less than or equal to 1,

$$V_h^0 := V_h \cap V^0$$

and

$$V_h^\Gamma := V_h \cap V^\Gamma.$$

We will also need the space of element wise affine or constant functions, that are not continuous over element boundaries

$$X_h^k := \{v_h \in L^2(\Omega) : v_h|_T \in \mathbb{P}_k(T)\}, \quad k = 0, 1.$$

We will denote the L^2 scalar product over a set Ξ by

$$(v, w)_\Xi := \int_\Xi xy \, d\Xi, \quad \forall v, w \in L^2(\Omega)$$

and the associated norm by

$$\|x\|_{\mathcal{E}} := (x, x)_{\mathcal{E}}^{\frac{1}{2}},$$

the subscript will be dropped whenever $\mathcal{E} \equiv \Omega$. The $H^1(\Omega)$ -norm will be denoted by

$$\|v\|_V := \|v\| + \|\nabla v\|.$$

For the Sobolev spaces consisting of functions with square integrable derivatives of order less than or equal to $s = 1, 2$, over $X \subset \Omega$ we use the standard notation $H^s(X)$ with associated norm $\|\cdot\|_{H^s(X)}$, and for the set of functions that are in L^∞ with gradient in $[L^\infty]^d$ we use the notation $W^{1,\infty}(\Omega)$ with associated norm $\|\cdot\|_{1,\infty}$.

2.1 Inequalities

We will need a few auxiliary results on how different norms or semi norms are related. In particular we will need the following so-called inverse inequality and trace inequalities (see for instance [16])

$$\|\nabla v_h\|_T \leq C_i h^{-1} \|v_h\|_T \quad \forall v_h \in \mathbb{P}_k(T), \quad k \geq 0 \quad (1)$$

$$\|v\|_{\partial T} \leq C_i h_T^{-1/2} (\|v_h\|_T + h_T \|\nabla v\|_T), \quad \forall v \in H^1(T) \quad (2)$$

$$\|v_h\|_{\partial T} \leq C_i h_T^{-1/2} \|v_h\|_T, \quad \forall v_h \in \mathbb{P}_k(T), \quad k \geq 0. \quad (3)$$

We also define the broken norm

$$\|v\|_h := \left(\sum_{K \in \mathcal{T}} \|v\|_K^2 \right)^{\frac{1}{2}}.$$

2.2 Interpolants and approximation

We will use an interpolant $i_h : H^1(\Omega) \mapsto V_h$, that preserves homogeneous boundary conditions and satisfies the following estimates [24]

$$\|u - i_h u\| + h \|\nabla(u - i_h u)\| \leq h^s \|u\|_{H^s(\Omega)}, \quad s = 1, 2. \quad (4)$$

Combining (4) and (2) allows us to prove the estimates

$$\|h^{-\frac{1}{2}}(u - i_h u_h)\|_{\mathcal{F}} + \|h^{\frac{1}{2}}\nabla(u - u_h)\|_{\mathcal{F}} \leq h^{s-1} \|u\|_{H^s(\Omega)}, \quad s = 1, 2. \quad (5)$$

We will also make use of the L^2 -projection, $\pi_h : L^2(\Omega) \mapsto V_h^0$ defined by

$$(\pi_h u, v_h) = (u, v_h) \quad \forall v_h \in V_h^0 \quad (6)$$

and the H^1 -projection $\Pi_h : H_0^1(\Omega) \mapsto V_h^0$ defined by

$$(\nabla \Pi_h u, \nabla v_h) = (\nabla u, \nabla v_h), \quad \forall v_h \in V_h^0. \quad (7)$$

We note that under the assumption of quasi uniformity (for the L^2 -projection) and convexity of the domain (for the H^1 -projection) both these approximations satisfy (4) and (5).

3 Ill-posed problems

It is well known that instabilities may cause suboptimality for approximations of convection–diffusion equations when the standard Galerkin method is applied. Examples of how stabilised methods can improve on the situation include the Galerkin Least Squares method [20, 6], subgrid viscosity [19] or the continuous interior penalty method [11]. This is an example of a problem that is well-posed on the continuous level, but where the discrete system may be ill-conditioned and produce poor quality approximations, unless all the scales of the problem have been resolved, something which may be difficult to achieve in practice. The arguments to analyse such methods use the positivity of the bilinear operator $a(\cdot, \cdot)$ defining the problem. In many cases in practice however the problem is indefinite, such as for Helmholtz equation or for non-coercive convection–diffusion. Then the bilinear form does not satisfy such a positivity property and the inf-sup condition that underpins well-posedness on the continuous level can be difficult to reproduce on the discrete level. This led us to develop a method which does not rely on coercivity or inf-sup stability for its analysis [7]. In that work we observed that this method, since it does not rely on the well-posedness structure for its design it can also be applied to ill-posed problems and this case was then analysed in [8] and applied to a series of different ill-posed problems in [9, 12, 14, 13].

In this section we will discuss how to apply stabilised finite elements to the approximation of ill-posed problems. Of course the class of ill-posed problems is very large and most of these problems are not tractable to the type of high resolution methods that we wish to apply here, so first we will discuss what type of ill-posed problems we are interested in and give some examples. The introduction to the field given below is by necessity rudimentary. For readers interested in delving deeper into the theory of inverse and ill-posed problems and their regularisation, we refer to [22, 25, 2, 17, 21].

To define ill-posed problems it is customary to start by the definition of well-posedness due to Hadamard. To this end we introduce the abstract problem

$$\mathcal{L}u = f \quad (8)$$

where $\mathcal{L} : V \mapsto V', f \in V'$.

Definition 1. (Well-posed problem)

The problem (8) is well-posed if

1. for every $f \in V'$ there exists $u \in V$ satisfying (8). This means that V' is in the range of \mathcal{L} .
2. The solution u is unique in V . That is, \mathcal{L}^{-1} exists.
3. u depends continuously on data.

$$\|u\|_V \leq C\|f\|_{V'}.$$

Definition 2. (Ill-posed problem)

The problem (8) is said to be ill-posed if at least one of the three points defining a well-posed problem in definition (8) fails.

Of course it does not make sense to attempt to approximate any ill-posed problem. However it was recognised by Tikhonov that some ill-posed problems were better behaved than others and we will therefore characterise what we will call conditionally well-posed problems. Here we consider the extended operator

$$\mathcal{K} : V \mapsto (V', M)$$

where M is some dataset. The ill-posed problem that we want to solve can then be cast as, find u , such that

$$\mathcal{L}u = f, \quad |u - u_M|_{\mathcal{M}} = 0$$

where $|\cdot|_{\mathcal{M}}$ is a semi-norm to be specified measuring the data.

This definition differs slightly from the definition given in [21]. First we will introduce subsets of V, V'

$$U := \{v \in V : \exists l \in V', u_M \in M \text{ such that } \mathcal{L}v = l, |v - u_M|_{\mathcal{M}} = 0\}.$$

$$\mathcal{K}(U) := \{(l, u_M) \in (V', M) : \exists v \in U \text{ such that } \mathcal{L}v = l, |v - u_M|_{\mathcal{M}} = 0\}.$$

Definition 3. (Conditionally well-posed problem)

The problem (8) is said to be conditionally well-posed on U if

1. $(f, u_M) \in \mathcal{K}(U)$;
2. a solution u to (8) with $(f, u_M) \in \mathcal{K}(U)$ is unique;
3. Assume that $\|f\|_{V'} < \varepsilon$ and $|u_M|_{\mathcal{M}} \leq \eta$. For some (semi)-norms $\|\cdot\|_S, \|\cdot\|_C$, there holds

$$\|u\|_S \leq C_E \Phi(\varepsilon + \eta),$$

where C_E depends on $\|u\|_C$, $\Phi(x)$ is a continuous monotone growing function that vanishes for $x = 0$.

Observe that this definition says that if the solution exists it must have some partial stability properties. The stability is also conditional on an upper bound $\|u\|_C \leq E$. The problems that we will consider in this course are on the form (8)

and such that for a neighbourhood of (f, u_M) , with diameter δ , $\mathcal{N}_\delta(f, u_M)$ there holds $\mathcal{N}_\delta(f, u_M) \cap \mathcal{K}(U) \neq \emptyset$. Observe that this means that we do not assume that the data of our problem admits a unique solution. We only assume that in a neighbourhood of the data the solution exists. This allows for perturbed data to be used. Below we will let (f, u_M) denote the unperturbed data that is in $\mathcal{K}(U)$, but unavailable and denote the perturbed available data by (\tilde{f}, \tilde{u}_M) .

We will now proceed to give to different examples of problems that are conditionally well-posed in the above sense.

Example 1. (The elliptic Cauchy problem, continuous dependence)

Assume that $\mathcal{L} := -\Delta + \sigma$ where $\sigma \in \mathbb{R}$ and assume that the boundary of Ω is consists of two simply connected parts Γ and Γ' . Consider the problem of finding $u \in H^1(\Omega)$ such that

$$\mathcal{L}u = f \text{ in } \Omega \quad (9)$$

$$u = g \text{ on } \Gamma \quad (10)$$

$$\nabla u \cdot n = \psi \text{ on } \Gamma. \quad (11)$$

For this problem

$$|u_M|_{\mathcal{M}} := \|g\|_{H^{\frac{1}{2}}(\Gamma)} + \|\psi\|_{H^{-\frac{1}{2}}(\Gamma)}. \quad (12)$$

Let us now exemplify the failure of continuous dependence for this problem. Let $\Omega := \{(x, y) \in \mathbb{R}^2 : x > 0\}$ and $\Gamma = \{(x, y) \in \mathbb{R}^2 : x = 0\}$, $\sigma = 0$, $f = 0$, $g = 0$ and

$$\psi(y) = \frac{1}{n} \sin(ny).$$

It is easy to verify that the solution in that case is

$$u(x, y) = \frac{1}{2n^2} \sin(ny)(e^{nx} - e^{-nx}).$$

Clearly as n becomes large $\|\psi\|_{L^\infty(\Gamma)}$ goes to zero, but $\sup_{y \in \mathbb{R}} |u(x, y)|$ blows up for any $x > 0$, showing the failure of continuous dependence.

Example 2. (The elliptic data assimilation problem, uniqueness)

Assume that $\mathcal{L} := -\Delta$ and that measurements u_M of u are available in some open connected subset of Ω , $\omega \subset \Omega$, then we can formulate the data assimilation problem as

$$\mathcal{L}u = f \text{ in } \Omega \quad (13)$$

$$u = u_M \text{ in } \omega. \quad (14)$$

Here

$$|u_M|_{\mathcal{M}} := \|u_M\|_{\omega}. \quad (15)$$

Assume that u_M, f are such that there exists a solution $u \in H^1(\Omega)$ to (13) and (14). Then this solution is unique. The proof is an elementary consequence of the

properties of harmonic functions. Assume that there exists two solutions and let $v = u_1 - u_2$. Then

$$\mathcal{L}v = 0 \text{ in } \Omega \quad (16)$$

$$v = 0 \text{ in } \omega. \quad (17)$$

This means that v is a harmonic function in Ω and hence analytic. By equation (14) v vanishes in ω with non-zero measure and hence since it is analytic, $v \equiv 0$ in Ω .

Remark 1. Using the compatibility condition $\mathcal{L}u_M = f$ in ω one may show that (13)-(14) is equivalent to the Cauchy problem

$$\mathcal{L}u = f \text{ in } \Omega \setminus \omega \quad (18)$$

$$u = u_M \text{ on } \partial\omega \quad (19)$$

$$\nabla u \cdot n = \nabla u_M \cdot n \text{ on } \partial\omega. \quad (20)$$

The conditional stability for these problems can be derived from the analysis in [1]. This uses the weak formulation of the problem and we will therefore propose weak formulations for the model problems proposed in the examples. This will then also serve in the formulation of the associated finite element methods.

3.1 Weak formulations of the model problems

Let us first consider the Cauchy-problem and introduce the spaces

$$V^\Gamma := \{v \in H^1(\Omega) : v|_\Gamma = 0\}$$

and

$$W^\Gamma := \{v \in H^1(\Omega) : v|_{\Gamma'} = 0\} (= V^{\Gamma'}).$$

Now observe that the solution of (9) can be sought in V^Γ and multiply (9) by some $v \in W^\Gamma$ and integrate by parts to obtain

$$(\mathcal{L}u, v) = (\nabla u, \nabla v) + (\sigma u, v) - \int_\Gamma \underbrace{\nabla u \cdot n}_= -\psi v \, ds - \int_{\Gamma'} \nabla u \cdot n \underbrace{v}_= 0 \, ds$$

By defining

$$a(u, v) := (\nabla u, \nabla v) + (\sigma u, v)$$

we arrive at the weak formulation: find $u \in V^\Gamma$ such that

$$a(u, v) = (f, v) + (\psi, v)_\Gamma, \quad \forall v \in W^\Gamma. \quad (21)$$

This weak formulation looks deceptively much as the weak formulation for the Poisson problem, but observe that the choice $v = u$ is not allowed since $u \notin W^\Gamma$.

Turning now to the data assimilation problem we obtain recalling the spaces V^0 and V from the introduction and observe that for problem (13) we may multiply with $v \in V^0$ to obtain

$$(\mathcal{L}u, v) = (\nabla u, \nabla v) - \int_{\partial\Omega} \nabla u \cdot n \underbrace{v}_{=0} ds.$$

This time we define

$$a(u, v) := (\nabla u, \nabla v).$$

It follows that we formally write problem (13)-(14), find $u \in V$ such that $u|_{\omega} = u_M$ and

$$a(u, v) = (f, v) \quad \forall v \in V^0. \quad (22)$$

For this solution to exist it is of course required that the compatibility $\mathcal{L}u_M|_{\omega} = f|_{\omega}$ is satisfied. Once again it is not allowed to take $v = u$ due to the different choices of spaces.

To simplify the discussion below we will use V for the primal space and W for the test space.

3.2 Conditional stability

Conditional stability for the two model problems can be deduced from [1]. Typically these estimates depend on a data term and the size of the right hand side. We introduce a semi norm measuring the measured data

$$|\cdot|_{\mathcal{M}}$$

and the dual norm measuring the data in the pde (observe that they are not necessarily independent, c.f. the Cauchy problem),

$$\|l\|_{W'} = \sup_{w \in W} \frac{|l(w)|}{\|w\|_W}.$$

Example 3. In the case of the Cauchy problem the measurements are given by ψ and we measure the $H^{-1/2}(\Gamma)$ norm of the normal gradient

$$|u|_{\mathcal{M}} := \|\nabla u \cdot n\|_{H^{-1/2}(\Gamma)}.$$

In the case of the data assimilation problem the \mathcal{M} -norm is given by the L^2 -norm over the subset ω .

$$|u|_{\mathcal{M}} := \|u\|_{\omega}.$$

We can then state the following result that is valid for both cases. Here we state the conditions on ω' in slightly simplified form, for a precise definition of ω' see [1, Theorem 1.7]) (9)

Theorem 1. Let $u \in V$ be such that, with $l \in W'$,

$$a(u, v) = l(v).$$

let $|\cdot|_{\mathcal{M}}$ be defined by (12) for the Cauchy problem (9) and by (15) for the data assimilation problem (13) and $|u - u_M|_{\mathcal{M}} = 0$. Then if $|u_M|_{\mathcal{M}} \leq \eta$ and $\|l\|_{W'} \leq \varepsilon$ then for every open simply connected $\omega' \subset \Omega$ such that $\text{dist}(\partial\omega', \partial\Omega) > 0$ there holds

$$\|u\|_{\omega'} \leq C\|u\|^{1-\tau}(\eta + \varepsilon)^\tau \text{ for some } \tau \in (0, 1)$$

Proof. For a proof of this result with full detail on involved constants see [1, Theorem 1.7] for the Cauchy problem and [1, Theorem 4.4] for the data assimilation case.

Remark 2. A similar result for global stability of u on the form

$$\|u\|_{\Omega} \leq C(\|u\|_V) |\log(\eta + \varepsilon)|^{-\tau}, \quad \tau \in (0, 1)$$

is also derived in [1] and may be used to derive global error estimates using the techniques below.

Remark 3. Conditional stability has been used before to tune the regularisation parameters for Tikhonov regularisation methods see for instance [15]. What is new in the approach that we advocate is that the conditional stability is used to drive the design of the regularisation terms on the discrete level in order to obtain the best possible accuracy with respect to approximation and stability.

4 Finite element approximation of ill-posed problems

The aim of the present section is to draw on our experience of stabilised FEM from the previous section to design methods that perturb the original problem as little as possible, while still remaining stable. The ideas that are presented below are mainly taken from [9, 13].

Assume that we wish to attempt to discretise ill-posed problem on the form, find $u \in V$ such that

$$a(u, v) = l(w), \quad \forall w \in W \tag{23}$$

and

$$|u - u_M|_{\mathcal{M}} = 0. \tag{24}$$

We assume that the data u_M is such that a solution $u \in V$, to (23) exists, satisfying (24).

Consider the case of the Cauchy problem, where the measurements are built into the weak formulation and we only need to consider (23). Define the finite element spaces (assuming here that the mesh is fitted to the subsets of the boundary Γ and Γ' ,

$$V_h^\Gamma := V_h \cap V^\Gamma$$

and

$$W_h^\Gamma := V_h \cap W^\Gamma.$$

We then have the discrete formulation: find $u_h \in V_h^\Gamma$ such that

$$a(u_h, w_h) = (f, w_h) + (\Psi, w_h)_\Gamma, \quad \forall w_h \in W_h^\Gamma \quad (25)$$

observe that the corresponding linear system can not be invertible in the general case, because there is no reason that the system matrix is square. Indeed this only holds in the special case when the number of vertices in Γ is the same as the number of vertices in Γ' . Similarly the matrix corresponding to a naive finite element discretisation of (22) (using V_h and $W_h := V_h^0$ and assuming that the mesh is fitted to the domain ω) is not square and in general the system is singular even if we impose $u_h|_\omega = 0$.

The idea is then to combine the satisfaction of (24) and (23) on the discrete level, by *minimising* (24) under the *constraint* (23). In practice $|\cdot|_{\mathcal{M}}$ may not be practical to use for minimisation purposes and may then often be replaced by another form that is equivalent on discrete spaces $|\cdot|_{\mathcal{M}_h}$. One must then in an additional step show that the minimisation of the error in $|\cdot|_{\mathcal{M}_h}$ indeed leads to a bound of the error in $|\cdot|_{\mathcal{M}}$. Below we will assume that this is true and that the discrete semi-norm $|\cdot|_{\mathcal{M}}$ is defined by a scalar product

$$|v|_{\mathcal{M}} := (v, v)_{\mathcal{M}_h}^{\frac{1}{2}}$$

and that for $u \in H^2(\Omega)$,

$$|u - i_h u|_{\mathcal{M}_h} \leq Ch|u|_{H^2(\Omega)}.$$

To handle both model problems in the same abstract framework we will denote the trial spaces V_h and the test spaces W_h in both cases.

We form the tentative Lagrangian

$$\mathfrak{L}(u_h, z_h) := \frac{1}{2} \gamma_M |u_h - \tilde{u}_M|_{\mathcal{M}_h}^2 + a(u_h, z_h) - \tilde{l}(z_h),$$

where $\tilde{u}_M = u_M + \delta u$ is the perturbed data available for the reconstruction and $\tilde{l}(z_h) = l(z_h) + \delta l(z_h)$ is a perturbed right hand side.

Example 4. For the case of the Cauchy problem we take

$$|u_h - \tilde{u}_M|_{\mathcal{M}_h}^2 := \int_\Gamma h (\partial_n u_h - \tilde{\Psi})^2 ds. \quad (26)$$

For the data assimilation problem we immediately have

$$|u_h - \tilde{u}_M|_{\mathcal{M}_h}^2 := |u_h - \tilde{u}_M|_{\mathcal{M}}^2 = \int_\omega (u_h - \tilde{u}_M)^2 dx. \quad (27)$$

Observe that if u is a solution to (23) and (24) then it will minimise the Lagrangian (if $\delta u = \delta l = 0$) with the associated multiplier $z = 0$. Unfortunately the associated minimisation problem may not be well-posed on the discrete level due to the ill-posedness of $a(\cdot, \cdot)$, even if the data of the continuous problem is in $\mathcal{H}(U)$ and there is a unique solution in U . It follows that we need some regularisation.

4.1 Regularisation – stabilisation

The classical way of obtaining a well-posed optimisation problem is through Tikhonov regularisation. In this case the natural choice would be to add regularising terms in the H^1 -semi norm for both the primal and the dual variable to obtain

$$\mathcal{L}(u_h, z_h) := \frac{1}{2} \gamma_M |u_h - \tilde{u}_M|_{\mathcal{M}_h}^2 + \gamma_1 \|\nabla u_h\|^2 - \gamma_2 \|\nabla z_h\|^2 + a(u_h, z_h) - \tilde{I}(z_h).$$

If we compute the Euler-Lagrange equations for this Lagrangian we obtain the system: find $(u_h, z_h) \in V_h \times W_h$ such that

$$a(u_h, w_h) - \gamma_2 (\nabla z_h, \nabla w_h) = \tilde{I}(w_h) \quad \forall w_h \in W_h \quad (28)$$

$$a(v_h, z_h) + \gamma_1 (\nabla u_h, \nabla v_h) + \gamma_M (u_h, v_h)_{\mathcal{M}_h} = \gamma_M (\tilde{u}_M, v_h)_{\mathcal{M}_h} \quad (29)$$

$$\forall v_h \in V_h.$$

Here the form $(u_h - \tilde{u}_M, v_h)_{\mathcal{M}_h}$ is obtained by deriving the forms suggested in example 4.

Remark 4. This system bears a strong resemblance to the quasi reversibility method [22] on mixed form for the Cauchy problem proposed for the continuous problem in [3]. Therein it was proven that if the exact solution exists, and the data are unperturbed, then if $\gamma_1 \rightarrow 0$ for bounded γ_2 (that may tend to zero, but at a lower rate than γ_1) then the regularised solution converges to the exact solution.

Drawing on our experience from stabilised finite element methods we would like to modify the regularisation terms, so that they vanish at an optimal rate when if $u_h \rightarrow u \in H^2(\Omega)$, $z_h \rightarrow 0$, while keeping the regularisation parameters fixed. We therefore introduce the abstract regularisation operators $s : V_h \times V_h \mapsto \mathbb{R}$ and $s^* : W_h \times W_h \mapsto \mathbb{R}$ in the Lagrangian

$$\mathcal{L}(u_h, z_h) := \frac{1}{2} \gamma_M |u_h - \tilde{u}_M|_{\mathcal{M}_h}^2 + \frac{1}{2} s(u_h, u_h) - \frac{1}{2} s^*(z_h, z_h) + a(u_h, z_h) - \tilde{I}(z_h). \quad (30)$$

The corresponding Euler-Lagrange equations then reads

$$a(u_h, w_h) - s^*(z_h, w_h) = \tilde{I}(w_h) \quad (31)$$

$$a(v_h, z_h) + s(u_h, v_h) + \gamma_M (u_h, v_h)_{\mathcal{M}_h} = \gamma_M (\tilde{u}_M, v_h)_{\mathcal{M}_h}. \quad (32)$$

The primal stabilisation operator should be weakly consistent,

$$s(i_h u, i_h u)^{\frac{1}{2}} \leq Ch|u|_{H^2(\Omega)}, \quad (33)$$

and be bounded, $s(v_h, v_h) \leq C\|v_h\|_V^2$. The dual stabilisation on the other hand must be equivalent with the W norm $c_1\|w_h\|_W^2 \leq s^*(w_h, w_h) \leq c_2\|w_h\|_W^2$, but the lower bound is not necessarily uniform in h .

Anticipating the results of the next section we propose the following examples of stabilisation operators,

$$s(v_h, v_h) := \gamma_1 \|h\sigma u_h\|^2 + \gamma_1 \sum_{F \in \mathcal{F}_i} (h_F \llbracket \nabla v_h \rrbracket, \llbracket \nabla v_h \rrbracket)_F =: \gamma_1 |v_h|_V^2 \quad (34)$$

and

$$s^*(v_h, v_h) = \gamma_2 (\nabla v_h, \nabla v_h)_\Omega =: \gamma_2 \|v_h\|_W^2. \quad (35)$$

Observe that for $u \in H^2(\Omega)$ in this case there holds $s(u, v_h) = \gamma_1 (h^2 \sigma^2 u, v_h)_\Omega$ for all $v_h \in V_h$, since the jump term vanishes when applied to sufficiently smooth functions. The remaining L^2 -term, is weakly consistent to the right order for piecewise affine elements. For higher order polynomial approximation of order k , the primal stabilisation operator in the Lagrangian (30) must be replaced by a strongly consistent residual based stabilisation of the form

$$s(v_h, v_h) := \|h^k \nabla u_h\|_\Omega^2 + \gamma_1 \|h(f + \Delta u_h - \sigma u_h)\|_h^2 + \gamma_1 \sum_{F \in \mathcal{F}_i} (h_F \llbracket \nabla v_h \rrbracket, \llbracket \nabla v_h \rrbracket)_F, \quad (36)$$

for details see the discussion in [9]. Recall that $z = 0$ is the solution to the unperturbed problem where data are such that a unique $u \in V$ exists and it follows that the stabilization (35) is also consistent. Observe that s defines a semi-norm on $V_h + H^2(\Omega)$ but that s^* defines a norm on W .

We then introduce the mesh dependent norm

$$\| |(u_h, z_h) | \|^2 := \gamma_M |u_h|_{\mathcal{M}_h}^2 + \gamma_1 |u_h|_V^2 + \gamma_2 \|z_h\|_W^2 + \min(\gamma_1, \gamma_M) h^2 \|u_h\|_{H^1(\Omega)}^2 \quad (37)$$

and observe that using (4) and (5) it is straightforward to prove the interpolation inequality

$$\| |(u - i_h u, 0) | \|^2 \leq Ch|u|_{H^2(\Omega)}. \quad (38)$$

To include the last term in the definition (37) we can apply a discrete Poincaré inequality.

Lemma 1. (*discrete Poincaré inequality*) *There exists $c_p > 0$ such that for all $v_h \in V_h$ there holds*

$$c_p h \|u_h\|_{H^1(\Omega)} \leq |u_h|_{\mathcal{M}_h} + |u_h|_V.$$

Proof. In the case of the Cauchy problem where $|\cdot|_{\mathcal{M}_h}$ is defined by (26) and $u_h|_\Gamma = 0$ this is a consequence of the Poincaré inequalities of [5] and for the data assimilation case where $|\cdot|_{\mathcal{M}_h}$ is defined by (50) the result was proved in [13].

The stabilisation parameters γ_1, γ_2 are real numbers and will not change during computation.

The system (31)-(32) can be cast on the compact form, find $(u_h, z_h) \in V_h \times W_h$ such that

$$A_h[(u_h, z_h), (v_h, w_h)] = \tilde{l}(w_h) + \gamma_M(\tilde{u}_M, v_h)_{\mathcal{M}_h}, \quad \forall (v_h, w_h) \in V_h \times W_h, \quad (39)$$

where

$$A_h[(u_h, z_h), (v_h, w_h)] := a(u_h, w_h) - s^*(z_h, w_h) + a(v_h, z_h) + s(u_h, v_h) + \gamma_M(u_h, v_h)_{\mathcal{M}_h}$$

Proposition 1. *The system (39) admits a unique solution $(u_h, z_h) \in V_h \times W_h$.*

Proof. By construction, for all (v_h, w_h)

$$\gamma_M |v_h|_{\mathcal{M}_h}^2 + \gamma_1 |v_h|_V^2 + \gamma_2 \|w_h\|_W^2 = A_h[(v_h, w_h), (v_h, -w_h)]$$

and therefore by Lemma 1 there exists $\alpha > 0$ such that

$$\alpha \| \|(v_h, w_h)\| \|^2 \leq A_h[(v_h, w_h), (v_h, -w_h)] \quad (40)$$

since the linear system is square it follows and by the above positivity there can exist no zero eigenvalues we conclude that the matrix is invertible.

Comparing with the exact problem (23)-(24) and assuming that $u \in H^2(\Omega)$, we see that the formulation (39) satisfies the following consistency relation

$$A_h[(u_h - u, z_h), (v_h, w_h)] = \delta l(w_h) + \gamma_M(\delta u, v_h)_{\mathcal{M}_h}, \quad \forall (v_h, w_h) \in V_h \times W_h. \quad (41)$$

Remark 5. Observe that the analysis may also be carried out using $s^* = s$, that is the form (34), but will lead to higher sensibility with respect to perturbations in data.

4.2 Error analysis using conditional stability

First we will introduce some continuity properties of the bilinear form using the stabilisations. Assume that $u \in H^2(\Omega)$, then there holds

$$a(u - i_h u, v_h) \leq Ch |u|_{H^2(\Omega)} \|v_h\|_W \quad (42)$$

and for all $u_h \in V_h$ and all $w \in W$, $i_h w \in W_h$

$$a(u_h, w - i_h w) \leq (C(u)h + \| \|(u - u_h, 0)\| \| \|w\|_W, \quad (43)$$

where the constants may depend on the parameters γ_1 , γ_2 and γ_M .

Example 5. For the data assimilation problem the equation (42) follows by an application of the Cauchy-Schwarz inequality and (4) and (43) follows by the integration by parts followed by (4) and (5) leading to

$$\begin{aligned} a(u_h, w - i_h w) &\leq |(\sigma u_h, w - i_h w)| + \sum_{F \in \mathcal{F}_i} \int_F |h^{\frac{1}{2}} [\nabla u_h]| h^{-\frac{1}{2}} |w - i_h w| \, ds \\ &\leq C \gamma_1^{-\frac{1}{2}} (|u - u_h|_V + \|\sigma h u\|) \|w\|_W \end{aligned}$$

The results for the Cauchy problem are obtained in a similar fashion and left as an exercise.

We are now ready to prove a first error estimate that holds independently of the stability properties of the continuous model.

Proposition 2. *If (u_h, z_h) is the solution of (39) and $u \in H^2(\Omega)$ is the solution of (23)-(24) then there holds*

$$\| (u - u_h, z_h) \| \leq C_\gamma h |u|_{H^2(\Omega)} + \delta_\gamma \quad (44)$$

where the perturbation is defined by $\delta_\gamma := \gamma_2^{-1/2} \|\delta l\|_{W'} + \gamma_M^{1/2} |\delta u|_{\mathcal{M}_h}$ and $C_\gamma := C(\gamma_1^{\frac{1}{2}} + \gamma_2^{-\frac{1}{2}})$.

Proof. To prove (44) we observe that by (38) and the triangle inequality it is enough to consider the discrete error $\xi_h = i_h u - u_h$. By (40) we have

$$\alpha \| (\xi_h, z_h) \| \leq A_h [(\xi_h, z_h), (\xi_h, -z_h)].$$

Using the Galerkin orthogonality (41) we may write

$$A_h [(\xi_h, z_h), (\xi_h, -z_h)] = A_h [(i_h u - u, 0), (\xi_h, -z_h)] - \delta l(z_h) + \gamma_M (\delta u, \xi_h)_{\mathcal{M}_h}.$$

Recalling that, by the continuity (42) there holds

$$\begin{aligned} A_h [(i_h u - u, 0), (\xi_h, -z_h)] &= a(u - i_h u, z_h) + s(i_h u - u, \xi_h) + \gamma_M (i_h u - u, \xi_h)_{\mathcal{M}_h} \\ &\leq Ch \gamma_2^{-\frac{1}{2}} |u|_{H^2(\Omega)} \gamma_2^{\frac{1}{2}} \|z_h\|_W + \underbrace{|i_h u - u|_V}_{\leq Ch \gamma_1^{\frac{1}{2}} |u|_{H^2(\Omega)}} \gamma_1^{\frac{1}{2}} |\xi_h|_V + \gamma_M |i_h u - u|_{\mathcal{M}_h} |\xi_h|_{\mathcal{M}_h}. \end{aligned}$$

Bounding also the perturbation terms

$$\delta l(w_h) \leq \gamma_2^{-\frac{1}{2}} \|\delta l\|_{W'} \gamma_2^{\frac{1}{2}} \|z_h\|_W$$

and

$$(\delta u, \xi_h)_{\mathcal{M}_h} \leq |\delta u|_{\mathcal{M}_h} |\xi_h|_{\mathcal{M}_h}$$

we arrive at

$$A_h [(\xi_h, z_h), (-\xi_h, z_h)] \leq Ch |u|_{H^2(\Omega)} \| (\xi_h, z_h) \| + \delta_\gamma \| (\xi_h, z_h) \|.$$

We conclude by dividing by $\| (\xi_h, z_h) \|$.

This proof is insufficient to show error estimates. However for unperturbed data and $u \in H^2(\Omega)$, it may be used to show that $u_h \rightarrow u$ as $h \rightarrow 0$, by a compactness argument.

To prove error estimates we must rely on the stability of Theorem 1. The idea behind the argument is to consider the error $e = u - u_h$ and observe that this error satisfies

$$a(e, w) = l(w) - a(u_h, w) =: r(w), \quad \forall w \in W, \quad (45)$$

for the data value $|e|_{\mathcal{M}}$. We must then use the result of Proposition 2 to show that $\|r\|_{W'} \leq \varepsilon$ and $|e|_{\mathcal{M}} \leq \eta$ in order to be able to apply Theorem 1.

In the data assimilation case we have $|e|_{\mathcal{M}} = |e|_{\mathcal{M}_h} = \|e\|_{\omega}$ so this quantity is immediately bounded by (44). For the Cauchy problem the continuous and discrete data matching terms are not the same, but one can prove that a suitable bound can be obtained for a perturbed error \tilde{e} by adding a small perturbation to u_h in the interface zone such that

$$|\tilde{e}|_{\mathcal{M}} \leq \|e, 0\|. \quad (46)$$

The error analysis then uses the arguments below together with a perturbation argument for \tilde{e} , for details see [10]. We will not consider that case here, instead focussing on the data assimilation case.

Theorem 2. *Let u be the exact solution to (23)-(24), with $l(w) := (f, w)$, $f \in L^2(\Omega)$, and where $|\cdot|_{\mathcal{M}} = \|\cdot\|_{\omega}$. Let u_h be the solution of (31)-(32) with the stabilisation operators (34)-(35). Then, for all $\omega' \subset \Omega$ satisfying the assumptions (1) there holds*

$$\|u - u_h\|_{\omega'} \leq C(\|f\| + |u|_{H^2(\Omega)} + h^{-1}\delta)^{(1-\tau)}(h(\|f\| + |u|_{H^2(\Omega)}) + \delta_{\gamma} + \|\delta l\|_{W'})^{\tau}$$

where the constant depends on the mesh geometry and negative and positive powers of γ_1 , γ_2 and γ_M .

Proof. The estimate is shown by applying Theorem (1) to the problem satisfied by the error.

We now that e is a solution to (23) with $l(w) = r(w)$ as per equation (45). By Proposition 2 the following bounds hold

$$|e|_{\mathcal{M}_h} = \|e\|_{\omega} \leq C_{\gamma} h |u|_{H^2(\Omega)} + \delta_{\gamma} \quad (47)$$

and

$$\|e\|_V \leq C_{\gamma} |u|_{H^2(\Omega)} + h^{-1} \delta_{\gamma}. \quad (48)$$

Now observe that using the equation (31) we have

$$r(w) = r(w) - r(i_h w) + r(i_h w) = l(w - i_h w) - a(u_h, w - i_h w) - s(z_h, i_h w) - \delta l(i_h w).$$

We estimate the terms of the right hand side, assuming that $\|w\|_W = 1$

$$l(w - i_h w) = (f, w - i_h w) \leq \|f\| \|w - i_h w\| \leq Ch \|f\|,$$

and using the relation (43)

$$a(u_h, w - i_h w) \leq (Ch + \|(u - u_h, 0)\|) \|w\|_W.$$

Then applying Proposition 2 we obtain the bound

$$a(u_h, w - i_h w) \leq \gamma_1^{-\frac{1}{2}} (C_\gamma h \|u\|_{H^2(\Omega)} + \delta_\gamma)$$

The two remaining terms are handled using the Cauchy-Schwarz inequality in the first case and the duality pairing H^{-1}/H^1 in the second, followed by the stability of the interpolant i_h in the W -norm,

$$s(z_h, i_h w) \leq \gamma_2 \|z_h\|_W \|w\|_W \leq \gamma_2^{\frac{1}{2}} (C_\gamma h \|u\|_{H^2(\Omega)} + \delta_\gamma)$$

$$\delta l(i_h w) \leq C \|\delta l\|_{W'}$$

Collecting the terms above we have for all $w \in W$ with $\|w\|_W = 1$,

$$r(w) \leq Ch \|f\| + (\gamma_1^{-\frac{1}{2}} + \gamma^{\frac{1}{2}}) (C_\gamma h \|u\|_{H^2(\Omega)} + \delta_\gamma) + C \|\delta l\|_{W'}.$$

But then

$$\begin{aligned} \|r\|_{W'} &= \sup_{w \in W: \|w\|_W=1} \langle r, w \rangle_{(W', W)} = \sup_{w \in W: \|w\|_W=1} r(w) \\ &\leq Ch \|f\| + (\gamma_1^{-\frac{1}{2}} + \gamma^{\frac{1}{2}}) (C_\gamma h \|u\|_{H^2(\Omega)} + \delta_\gamma) + \|\delta l\|_{W'}. \end{aligned} \quad (49)$$

We conclude that e satisfies the assumptions of Theorem 1 by (48) with $\eta = Ch \|u\|_{H^2(\Omega)} + \delta_\gamma$, c.f. equation (47) and $\varepsilon = C(h \|f\| + h \|u\|_{H^2(\Omega)} + \delta_\gamma + \|\delta l\|_{W'})$, c.f. (49). In the last step we dropped the dependence on constants of γ , but it can be traced in the proof.

Remark 6. We detailed Theorem 2 only in the case of the data assimilation problem, but the same arguments also leads to an analysis for the Cauchy problem, under the assumption (46)

Remark 7. One may prove Theorem 2 for the data assimilation problem if s^* is defined by (34). In this case an additional factor h^{-1} multiplies the term measuring perturbations in data.

4.3 A numerical example

We consider the problem in Example 1 on the unit square Ω . The exact solution is $u = 30.0 * x * (1 - x) * y * (1 - y)$, with $f = \mathcal{L}u$, and the data domain ω is defined by

$$\omega := \{(x, y) \in \Omega : |x - 0.5| < 0.25; |y - 0.5| < 0.25\}.$$

We use the formulation (31)-(32) with $s(\cdot, \cdot)$ given by (34) for piecewise affine approximation and (36) for piecewise quadratic approximation. The adjoint stabiliser $s^*(\cdot, \cdot)$ was defined by (35), and the norm $|\cdot|_{\mathcal{M}_h}$ by (50) with $\alpha = 0$ or -2 . (Observe that if $\alpha = 0$ then γ_M must have the unit of the square of an inverse length for the equations to be dimensionally correct.)

We chose $\gamma_2 = \gamma_M = 1$ and $\gamma_1 = 10^{-3}$ for all computations. The latter value is similar to that used for computations in the well-posed case. We meshed the domain using structured meshes that were made to fit the boundary of ω . We performed computations on a sequence of meshes with $n_{ele} = 40, 80, 160, 320$, elements on each side of the square, using piecewise affine and piecewise quadratic elements. In Figure 1, left graphic, we show a computational mesh and on the right graphic we illustrate the domains ω (the inner square) and ω' (the middle square). In Figure 2, left plot, we show the contourlines of an approximate solution and in the right plot the contour lines of the computational error. Observe that the error has a form that is similar to Hadamard's counter-example discussed in Example 1, but growing exponentially in the radial direction and oscillating in the direction tangential to the boundary of ω .

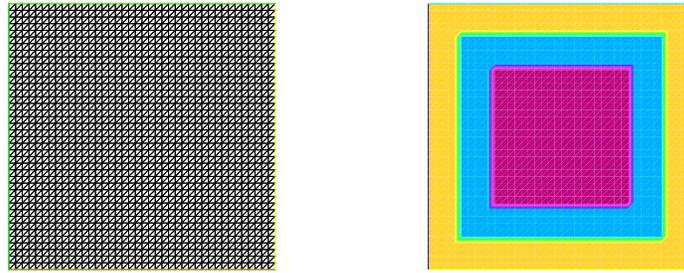


Fig. 1 Left: computational mesh, $n_{ele}=40$. Right: the different subdomains ω and ω' .

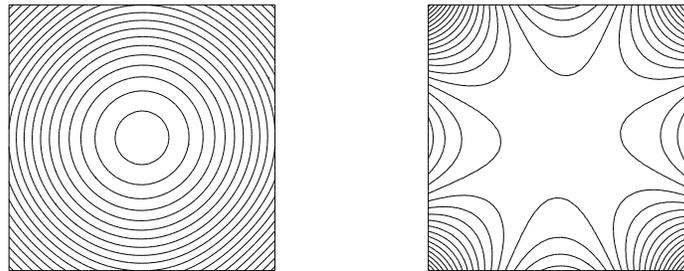


Fig. 2 Left: contour lines of approximate solution, $n_{ele}=40$. Right: contour lines of the computational error.

In the tables below we report the error in the normalised global L^2 -error, the normalised local error for the subset

$$\omega' := \{(x, y) \in \mathbb{R}^2 : |x - 0.5| < 0.375; |y - 0.5| < 0.375\},$$

the data assimilation term, $|u - u_h|_\omega$, and the size of the weakly consistent regularisation

$$|(u - u_h, z)|_s := \sqrt{s(u - u_h, u - u_h) + s^*(z_h, z_h)}. \quad (50)$$

The experimental convergence rates are given in parenthesis, where appropriate. We report the results for unperturbed data and $\alpha = 0$ in tables 1 and 5 and for $\alpha = -2$ in tables 2 and 6. In all cases we observe the expected $O(h^k)$ convergence of the stabilising terms (??), with $k = 1$ for piecewise affine approximation and $k = 2$ in the quadratic case. We also observe that consistently with theory we have $\|u - u_h\|_\omega = O(h^{k-\alpha/2})$. The convergence of the data term is more even for $\alpha = -2$. For the global and local L^2 -norms we see that the error is a factor 5 – 10 larger when $\alpha = 0$ compared with the case where $\alpha = -2$. Most likely this is due to the fact that the missing length-scale present for $\alpha = 0$ is not well represented when $\gamma_M = 1.0$. Indeed the weak penalty does not impose the data sufficiently well compared to the other terms, when $\alpha = -2$ on the other hand the data penalty term is so strong that the data error is very small already on coarse meshes leading to improved local and global errors. We observe convergence compatible with Hölder stability for all quantities, indicating that possibly we are not yet in the asymptotic regime on these scales. Only on the finest meshes in table 6 we see clearly the decreasing orders characteristic for logarithmic convergence in the global error.

We then make the same sequence of computations but adding a perturbation of 2.5% to the data in ω in the piecewise affine case and 1% in the quadratic case. The results are reported for affine approximation in tables 3 ($\alpha = 0$) and 4 ($\alpha = -2$). We observe that although the data assimilation term stagnates, the local and global errors decrease under refinement for $\alpha = 0$. In this case the stabilisation norm also converges to optimal order in spite of the perturbation. When $\alpha = -2$ only the error in the stabilisation semi-norm show any decrease under refinement. On the finest scale we see that both the global error and the error in the stabilisation semi-norm has started to grow. For piecewise affine approximation it appears that the choice $\alpha = -2$ is superior both for perturbed and unperturbed data (at least for the choice $\gamma_M = 1$).

For quadratic approximation the results are reported in tables 7 ($\alpha = 0$) and 8 ($\alpha = -2$). Here the effect of the perturbation is present already on the coarsest mesh and the amplification of the error clearly much stronger for $\alpha = -2$. Indeed whereas for $\alpha = 0$ all error quantities still decrease under mesh refinement, the errors for $\alpha = -2$ all stagnate or increase. For the stabilisation norm we clearly see that the error doubles under mesh refinement on finer meshes, which is consistent with theory. In this case it appears that for resolutions where the mesh-size is of similar order as the perturbation it is advantageous to take $\alpha = 0$, also in accordance with theory.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
40	0.211594 (-)	0.050922 (-)	0.00816074 (-)	0.0289235 (-)
80	0.175512 (0.3)	0.0407488 (0.3)	0.00618422 (0.4)	0.0147585 (1.0)
160	0.113346 (0.6)	0.0235298 (0.8)	0.00337103 (0.9)	0.00791309 (0.9)
320	0.0672893 (0.75)	0.0102456 (1.2)	0.00119201 (1.5)	0.0042852 (0.9)
640	0.0510429 (0.4)	0.00529074 (1.0)	0.000342379 (1.8)	0.00221974 (0.9)

Table 1 Computed quantities for the data assimilation problem using piecewise affine approximation, $\alpha = 0$ and unperturbed data.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
40	0.0476335 (-)	0.00481282 (-)	0.000333429 (-)	0.0352793 (-)
80	0.0403148 (0.2)	0.00312934 (0.6)	8.0272e-05 (2.0)	0.0179655 (1.0)
160	0.0304957 (0.4)	0.00188862 (0.7)	1.998e-05 (2.0)	0.00911884 (1.0)
320	0.0227619 (0.4)	0.0009549 (1.0)	4.71016e-06 (2.1)	0.00464924 (1.0)
640	0.0200062 (0.2)	0.000642748 (0.6)	1.15698e-06 (2.0)	0.00234456 (1.0)

Table 2 Computed quantities for the data assimilation problem using piecewise affine approximation, $\alpha = -2$ and unperturbed data.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
40	0.206909	0.0490942	0.0148287	0.0289287 (-)
80	0.176546	0.0409112	0.013946	0.0146984 (1.0)
160	0.119693	0.0267951	0.0131763	0.0077906 (0.9)
320	0.0793605	0.0180773	0.0125264	0.00416117 (0.9)
640	0.0640708	0.0158747	0.0124993	0.00214582 (1.0)

Table 3 Computed quantities for the data assimilation problem using piecewise affine approximation, $\alpha = 0$ and 2.5% perturbation in data.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
40	0.0520752	0.0145883	0.0124714	0.03529
80	0.0507222	0.014398	0.0125092	0.0186372
160	0.0502568	0.0143645	0.0127194	0.0142032
320	0.0537505	0.0143083	0.0125169	0.0224315
640	0.0427351	0.0138826	0.0125888	0.0434341

Table 4 Computed quantities for the data assimilation problem using piecewise affine approximation, $\alpha = -2$ and 2.5% perturbation in data.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
20	0.0113854 (-)	0.0020353 (-)	0.000272026 (-)	0.00263335 (-)
40	0.00701791 (0.7)	0.000668735 (1.6)	4.36798e-05 (2.6)	0.00067804 (2.0)
80	0.00630128 (0.16)	0.000458704 (0.54)	1.0293e-05 (2.1)	0.000171095 (2.0)
160	0.00457823 (0.5)	0.000278068 (0.72)	5.50828e-06 (1.0)	4.33632e-05 (2.0)
320	0.00275223 (0.7)	9.14176e-05 (1.6)	7.11806e-07 (2.8)	1.10465e-05 (2.0)

Table 5 Computed quantities for the data assimilation problem using piecewise quadratic approximation, $\alpha = 0$ and unperturbed data.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
20	0.00594613 (-)	0.000454428 (-)	1.92029e-05 (-)	0.00269387 (-)
40	0.00364274 (0.7)	0.000194766 (1.2)	3.21386e-06 (-2.6)	0.00069238 (-)
80	0.0023773 (0.6)	6.52831e-05 (1.6)	2.95005e-07 (3.4)	0.000176426 (2.0)
160	0.00159176 (0.6)	2.93421e-05 (1.2)	3.91486e-08 (2.9)	4.45628e-05 (2.0)
320	0.00118008 (0.4)	1.27615e-05 (1.2)	4.3179e-09 (3.2)	1.12277e-05 (2.0)

Table 6 Computed quantities for the data assimilation problem using piecewise quadratic approximation, $\alpha = -2$ and unperturbed data.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
20	0.0146381	0.00619699	0.00510402	0.00260206
40	0.0137215	0.00593519	0.00492976	0.00066236 (2.0)
80	0.0135235	0.00594218	0.00498009	0.000167333 (2.0)
160	0.0110434	0.00593666	0.00497521	4.82896e-05 (1.8)
320	0.00982659	0.0058722	0.00497389	1.23888e-05 (2.0)

Table 7 Computed quantities for the data assimilation problem using piecewise quadratic approximation, $\alpha = 0$ and 1% perturbation in data.

nele	$\ u - u_h\ $	$\ u - u_h\ _{\omega'}$	$\ u - u_h\ _{\omega}$	$ (u - u_h, z) _s$
20	0.0177247	0.00638777	0.00513258	0.00275637
40	0.026475	0.00628408	0.00495361	0.00164336
80	0.0503314	0.00644259	0.00500485	0.002676516
160	0.159728	0.0079909	0.0050097	0.00510579
320	0.335852	0.00962178	0.0050035	0.0101055

Table 8 Computed quantities for the data assimilation problem using piecewise quadratic approximation, $\alpha = -2$ and 1% perturbation in data.

5 Time dependent problems – data assimilation

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 $L^2((0, T) \times \omega)$. In this section we consider the extension of the methods in the previous section to the time dependent case, where several interesting new features appear. In particular we can here consider a problem which has Lipschitz stability and prove that our method can exploit this in the form of error estimates that are optimal compared to approximation. We consider a data assimilation problem for the heat equation

$$\partial_t u - \Delta u = f, \quad \text{in } (0, T) \times \Omega, \quad (51)$$

with homogeneous Dirichlet conditions and where $T > 0$ and $\Omega \subset \mathbb{R}^n$ is an open convex polyhedral domain. Let $\omega \subset \Omega$ be open and let $0 < T_1 < T_2 \leq T$. The data assimilation problems is of the following form: determine the restriction $u|_{(T_1, T_2) \times \Omega}$

of a solution to the heat equation (51) given f and the restriction $u|_{(0,T)\times\omega}$. For this case the following stability estimate holds [23, 26, 14]

Theorem 3. *Let $\omega \subset \Omega$ be open and non-empty, and let $0 < T_1 < T$. Then there is $C > 0$ such that for all u in the space*

$$H^1(0, T; H^{-1}(\Omega)) \cap L^2(0, T; H_0^1(\Omega)), \quad (52)$$

it holds that

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

where $L = \partial_t - \Delta$ and $\|\cdot\|_{(0, -1)} = \|\cdot\|_{L^2(0, T; H^{-1}(\Omega))}$. The data of the data assimilation problem is given by

$$q = u|_{(0, T)\times\omega}, \quad f = Lu. \quad (53)$$

We will 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

$$G_f(u, z) = 0, \quad z \in L^2(0, T; H_0^1(\Omega)), \quad (54)$$

give the weak formulation of $\partial_t u - \Delta u = f$.

5.1 Optimisation based finite element space discretisation

We consider only the problem semi-discretised in space, and show that the time continuous dynamical system is well-posed for every fixed h . This section summarizes part of the analysis from [14], where also a problem with weaker stability, similar to that of the data assimilation problem in the previous section was considered. The analysis carries over to the fully discrete case, but the stabilisation operators are not the same. In particular in the fully discrete case, the adjoint stabilisation can be omitted.

Since the problem is time dependent we introduce the space-time spaces \mathcal{V}_h and \mathcal{W}_h ,

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

Observe that contrary to the developments in the previous section both spaces are equipped with Dirichlet conditions in space. The difference between the two spaces

here is the regularity in time. Following the development in the previous sections our approach to solve the data assimilation problem is based on minimizing the Lagrangian functional

$$\mathfrak{L}_{q,f}(u, z) = \frac{1}{2} \|u - q\|_{\omega}^2 + \frac{1}{2} s(u, u) - \frac{1}{2} s^*(z, z) + G_f(u, z), \quad (55)$$

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 the primal and dual stabilizers, respectively. Note that minimizing $\mathfrak{L}_{q,f}$ can be seen as fitting $u|_{(0,T) \times \omega}$ to the data q under the constraint (54), z can be interpreted as a Lagrange multiplier, and $s/2$ and $s^*/2$ as regularizing penalty terms.

Let $q \in L^2((0, T) \times \omega)$ and $f \in H^{(0,-1)}$. The Lagrangian $\mathfrak{L}_{q,f}$, defined by (55), satisfies

$$\begin{aligned} D_u \mathfrak{L}_{q,f} v &= (u - q, v)_{\omega} + s(u, v) + G(v, z), \\ D_z \mathfrak{L}_{q,f} w &= -s^*(z, w) + G(u, w) - \langle f, w \rangle, \end{aligned}$$

and therefore the critical points $(u, z) \in \mathcal{V}_h \times \mathcal{W}_h$ of $\mathfrak{L}_{q,f}$ satisfy

$$A[(u, z), (v, w)] = (q, v)_{\omega} + \langle f, w \rangle, \quad (v, w) \in \mathcal{V}_h \times \mathcal{W}_h, \quad (56)$$

where A is the symmetric bilinear form

$$A[(u, z), (v, w)] = (u, v)_{\omega} + s(u, v) + G(v, z) - s^*(z, w) + G(u, w). \quad (57)$$

Note that

$$A[(u, z), (u, -z)] = s(u, u) + \|u\|_{\omega}^2 + s^*(z, z),$$

Herein we consider only semi-discretisations, that is, we minimize $\mathfrak{L}_{q,f}$ on a scale of spaces that are discrete in the spatial variable but not in the time variable. As before the spatial mesh size $h > 0$ will be the only parameter controlling the convergence of the approximation, and we use piecewise affine finite elements. For simplicity we set all regularisation parameters to one and we consider only the case of unperturbed data.

5.2 A framework for stabilisation

Before proceeding to the analysis of the data assimilation problem, we introduce an abstract stabilisation framework.

Let s and s^* be bilinear forms on the spaces \mathcal{V}_h and \mathcal{W}_h , respectively. Let $|\cdot|_{\mathcal{V}}$ be a semi-norm on \mathcal{V}_h and let $\|\cdot\|_{\mathcal{W}}$ be a norm on \mathcal{W}_h . We relax (34) and (35) by requiring only that s and s^* are continuous with respect to $|\cdot|_{\mathcal{V}}$ and $\|\cdot\|_{\mathcal{W}}$, that is,

$$s(u, u) \leq C|u|_{\mathcal{V}}^2, \quad s^*(z, z) \leq C\|z\|_{\mathcal{W}}^2, \quad u \in \mathcal{V}_h, z \in \mathcal{W}_h, h > 0. \quad (58)$$

Let $\|\cdot\|_*$ be the norm of a continuously embedded subspace H^* of the energy space (52). The space H^* encodes the a priori smoothness. We assume that the stabilizations and norms introduced are such that the following continuities hold

$$G(u, z - \pi_h z) \leq C|u|_{\mathcal{V}} \|z\|_{(0,1)}, \quad u \in \mathcal{V}_h, z \in H_0^{(0,1)}, \quad (59)$$

$$G(u - \pi_h u, z) \leq Ch \|z\|_{\mathcal{W}} \|u\|_*, \quad u \in H^*, z \in \mathcal{W}_h, \quad (60)$$

where π_h is an interpolator satisfying the properties bounds

$$\pi_h : H_0^1(\Omega) \rightarrow W_h, \quad h > 0. \quad (61)$$

$$\|\pi_h u\|_{H^1(\Omega)} \leq C \|u\|_{H^1(\Omega)}, \quad u \in H^1(\Omega), \quad h > 0, \quad (62)$$

$$\|u - \pi_h u\|_{H^m(\Omega)} \leq Ch^{k-m} \|u\|_{H^k(\Omega)}, \quad u \in H^k(\Omega), \quad h > 0, \quad (63)$$

where $k = 1, 2$ and $m = 0, k - 1$. We assume that the following upper bound holds

$$|\pi_h u|_{\mathcal{V}} \leq Ch \|u\|_*, \quad u \in H^*, \quad (64)$$

and require that analogously to the stationary case

$$\|\pi_h z\|_{\mathcal{W}} \leq C \|z\|_{(0,1)}, \quad z \in H_0^{(0,1)}, \quad (65)$$

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

$$\|u\| \leq Ch^{-1} (|u|_{\mathcal{V}} + \|u\|_{\omega}), \quad u \in \mathcal{V}_h. \quad (66)$$

Observe that, by the lower bound (66),

$$\| \! \| (u, z) \! \| \! \| = |u|_{\mathcal{V}} + \|u\|_{\omega} + \|z\|_{\mathcal{W}},$$

is a norm on $\mathcal{V}_h \times \mathcal{W}_h$. Finally, in the abstract setting, we assume that the s and s^* are sufficiently strong so that the following weak coercivity holds where A and $\| \! \| \cdot \! \| \! \|$ are defined as above.

$$\| \! \| (u, z) \! \| \! \| \leq C \sup_{(v,w) \in \mathcal{V}_h \times \mathcal{W}_h} \frac{A[(u,z), (v,w)]}{\| \! \| (v,w) \! \| \! \|}, \quad (u,z) \in \mathcal{V}_h \times \mathcal{W}_h \quad (67)$$

and for all $(v,w) \in \mathcal{V}_h \times \mathcal{W}_h$,

$$\sup_{(x,y) \in \mathcal{V}_h \times \mathcal{W}_h} |A[(x,y), (v,w)]| > 0. \quad (68)$$

The Babuska-Lax-Milgram theorem implies that the equation (56) has a unique solution in $\mathcal{V}_h \times \mathcal{W}_h$. As we shall see below, these design criteria are sufficient to derive optimal error estimates in the transient case, provided the problem has a conditional stability property.

5.3 The data assimilation problem

We will now proceed to a specific case.

We choose the stabilizers and semi-norms as follows,

$$s(u, u) = \|h\nabla u(0, \cdot)\|_{L^2(\Omega)}^2, \quad s^* = a, \quad (69)$$

$$\|u\|_{\mathcal{V}} = s(u, u)^{1/2} + \|h\partial_t u\|, \quad \|z\|_{\mathcal{W}} = s^*(z, z)^{1/2}, \quad (70)$$

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

$$(\nabla \pi_h u, \nabla v)_{L^2(\Omega)} = (\nabla u, \nabla v)_{L^2(\Omega)}, \quad u \in H_0^1(\Omega), \quad v \in W_h. \quad (71)$$

As Ω is a convex polyhedron, it is well known that this choice satisfies (61)-(63), see e.g. [18, Th. 3.12-18].

Lemma 2. *The choices (69)-(71) satisfy the properties (58)-(64), (65) and (67). Moreover, $\|\cdot\|$ is a norm on $\mathcal{V}_h \times \mathcal{W}_h$.*

Proof. It is clear that the continuities (58) hold. We begin with the lower bound (59). By the orthogonality (71),

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

Towards the lower bound (60), we use the orthogonality (71) as above,

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

The bound (60) follows from the Poincaré inequality of Lemma 1.

The bound (64) follows from the continuity of the trace

$$\|\nabla u(0, \cdot)\|_{L^2(\Omega)} \leq \|u\|_{(1,1)}, \quad (72)$$

and the continuity of the projection π_h . The bound (65) follows immediately from the continuity of π_h .

We turn to the weak coercivity (67). The essential difference between the time dependent case and the stationary case is that in the latter case, the choice $w = u$ is prohibited. In this case it is allowed, but due to the time-derivative and the lack of initial condition it does not lead to stability. Instead we observe that $\partial_t u \in \mathcal{W}_h$ when $u \in \mathcal{V}_h$ so that this can be used as a test function $w = \partial_t u$ to obtain

$$A[(u, z), (0, \partial_t u)] = -s^*(z, \partial_t u) + G(u, \partial_t u) = \|\partial_t u\|^2 + a(u, \partial_t u) - a(z, \partial_t u),$$

and thus using bilinearity of A ,

$$A[(u, z), (u, \alpha h^2 \partial_t u - z)] = s(u, u) + \alpha \|h \partial_t u\|^2 + \|u\|_\omega^2 + s^*(z, z) \quad (73)$$

$$+ \alpha h^2 a(u, \partial_t u) - \alpha h^2 a(z, \partial_t u),$$

where $\alpha > 0$. We will establish (67) by showing that there is $\alpha \in (0, 1)$ such that

$$\| \| (u, w - z) \| \| \leq C \| \| (u, z) \| \|, \quad (74)$$

$$\| \| (u, z) \| \| ^2 \leq CA[(u, z), (u, w - z)], \quad (75)$$

where $w = \alpha h^2 \partial_t u$.

Towards (74) we observe that

$$\| \| (u, w - z) \| \| ^2 = \| \| (u, z) \| \| ^2 - 2s^*(z, w) + s^*(w, w) \leq 2\| \| (u, z) \| \| ^2 + 2s^*(w, w).$$

We use the discrete inverse inequality (1) to bound the second term

$$s^*(w, w) = \alpha^2 h^4 \|\partial_t \nabla u\|^2 \leq C \alpha^2 h^2 \|\partial_t u\|^2 \leq C \alpha^2 \| \| (u, z) \| \| ^2, \quad \alpha > 0.$$

It remains to show (75). Towards bounding the first cross term in (73) we observe that

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

Hence $\alpha h^2 a(u, \partial_t u) \geq -\alpha s(u, u)/2$. We use the arithmetic-geometric inequality,

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

and the discrete inverse inequality (1) to bound the second cross term in (73),

$$\alpha h^2 a(z, \partial_t u) \leq \alpha (4\varepsilon)^{-1} a(z, z) + \alpha \varepsilon h^4 \|\partial_t \nabla u\|^2 \leq \alpha (4\varepsilon)^{-1} a(z, z) + C \alpha \varepsilon \|h \partial_t u\|^2.$$

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

$$A[(u, z), (u, w - z)] \geq (1 - \alpha/2)s(u, u) + \alpha \|h \partial_t u\|^2/2 + \|u\|_\omega^2 + (1 - C\alpha/2)s^*(z, z),$$

and therefore (75) holds with small enough $\alpha > 0$.

The second condition (68) follows similarly using the symmetry of A . Fix $(v, w) \in \mathcal{V} \times \mathcal{W}$ and take $x = v$, $y = -w + \alpha h^2 \partial_t v$ for α sufficiently small and proceed as in the proof of (75). Finally, using the Poincaré inequality of Lemma 1, we see that $\| \| (u, z) \| \| = 0$ implies $z = 0$ and $u(0, \cdot) = 0$. As also $\partial_t u = 0$, we have $u = 0$, and therefore $\| \| \cdot \| \|$ is a norm.

5.4 Error estimates

We are now in a situation to prove an error estimate using the abstract theory.

Theorem 4. *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 (58)-(64), (65) and (67). Then (56) 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 $h \in (0, 1)$*

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

Proof. We begin again by showing the estimate

$$\| (u_h - \pi_h u, z_h) \| \leq Ch \|u\|_*. \quad (76)$$

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

$$\begin{aligned} G(u, w) &= \langle f, w \rangle, \quad w \in L^2(0, T; H_0^1(\Omega)), \\ (q - u, v)_\omega &= 0, \quad v \in L^2((0, T) \times \omega), \end{aligned} \quad (77)$$

and the equations (56) and (77) imply for all $v \in \mathcal{V}_h$ and $w \in \mathcal{W}_h$ that

$$A[(u_h - \pi_h u, z_h), (v, w)] = (u - \pi_h u, v)_\omega + G(u - \pi_h u, w) - s(\pi_h u, v). \quad (78)$$

The weak coercivity (67) implies that in order to show (76) it is enough bound the three terms on the right hand side of (78). For the first of them, that is, $(u - \pi_h u, v)_\omega$, we use (63). The lower bound (60) applies to the second term $G(u - \pi_h u, w)$, and for the third one we use the continuity (58) and the upper bound (64),

$$s(\pi_h u, v) \leq C |\pi_h u|_{\mathcal{V}} |v|_{\mathcal{V}} \leq Ch \|u\|_* |v|_{\mathcal{V}}.$$

We define the residual r as follows. By taking $v = 0$ in (56) we get $G(u_h, w) = \langle f, w \rangle + s^*(z_h, w)$, $w \in \mathcal{W}_h$, and therefore

$$\begin{aligned} \langle r, w \rangle &= G(u_h, w) - \langle f, w \rangle - G(u_h, \pi_h w) + G(u_h, \pi_h w) \\ &= G(u_h, w - \pi_h w) - \langle f, w - \pi_h w \rangle + s^*(z_h, \pi_h w), \quad w \in H_0^{(0,1)}. \end{aligned} \quad (79)$$

We now wish to arrive to the estimate

$$\|r\|_{(0,-1)} \leq C(|u_h|_{\mathcal{V}} + \|z_h\|_{\mathcal{W}} + h\|f\|). \quad (80)$$

To show that (80) holds, it is enough to bound the three terms on the right hand side of (79). The lower bound (59) applies to the first term $G(u_h, w - \pi_h w)$, for the second term $\langle f, w - \pi_h w \rangle$ we use (63), for the third term we use the continuity (58) and the upper bound (65)

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

The inequalities (80), (76) and (64) imply

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

Finally using the above bound on r , Theorem 3 implies that

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

The claim follows by using (76) and (63),

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

Here we used also the assumption that H^* is a continuously embedded subspace of the energy space (52), namely, this implies that the embedding $H^* \subset H^{(0,1)}$ is continuous.

Remark 8. If the data q, f is perturbed in this time-dependent case, the data assimilation problem behaves like a typical well posed problem, that is, a term

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

needs to be added on the right-hand side of the estimate in Theorem 6, but this time without any negative power of h . The proof is similar as in the stationary case and left as an exercise.

6 Conclusion

We have shown on some model problems how weakly consistent regularisation may be applied in the context of finite element approximation of ill-posed problems as a means to obtain approximations to the exact solution that are optimal with respect to the approximation order of the finite element space and the (conditional) stability of the physical problem. We have only considered piecewise affine approximation here but the extension to high order polynomial approximation (and with associated enhanced accuracy for smooth solutions) is possible using the ideas from [9]. Ongoing work focuses on problems where the stability depends on the parameters of the physical problem in a more intricate way such as for the convection–diffusion equation or the Helmholtz equation. Further work will also address the extension to systems such as the linearised Navier-Stokes’ equations.

References

1. G. Alessandrini, L. Rondi, E. Rosset, and S. Vessella. The stability for the Cauchy problem for elliptic equations. *Inverse Problems*, 25(12):123004, 47, 2009.

2. J. Baumeister. *Stable solution of inverse problems*. Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig, 1987.
3. L. Bourgeois. A mixed formulation of quasi-reversibility to solve the Cauchy problem for Laplace's equation. *Inverse Problems*, 21(3):1087–1104, 2005.
4. L. Bourgeois and J. Dardé. The “exterior approach” to solve the inverse obstacle problem for the Stokes system. *Inverse Probl. Imaging*, 8(1):23–51, 2014.
5. S. C. Brenner. Poincaré-Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Anal.*, 41(1):306–324, 2003.
6. A. N. Brooks and T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 32(1-3):199–259, 1982. FENOMECH '81, Part I (Stuttgart, 1981).
7. E. Burman. Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: Elliptic equations. *SIAM J. Sci. Comput.*, 35(6):A2752–A2780, 2013.
8. E. Burman. Error estimates for stabilized finite element methods applied to ill-posed problems. *C. R. Math. Acad. Sci. Paris*, 352(7-8):655–659, 2014.
9. E. Burman. Stabilised finite element methods for ill-posed problems with conditional stability. *ArXiv e-prints*, Dec. 2015.
10. E. Burman. The elliptic Cauchy problem revisited: Control of boundary data in natural norms. *Comptes Rendus Mathématique*, 2017. <http://dx.doi.org/10.1016/j.crma.2017.02.014>.
11. E. Burman and P. Hansbo. Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems. *Comput. Methods Appl. Mech. Engrg.*, 193(15-16):1437–1453, 2004.
12. E. Burman and P. Hansbo. Stabilized nonconforming finite element methods for data assimilation in incompressible flows. *ArXiv e-prints*, Sept. 2016.
13. E. Burman, P. Hansbo, and M. Larson. Solving ill-posed control problems by stabilized finite element methods: an alternative to Tikhonov regularization. *ArXiv e-prints*, Sept. 2016.
14. E. Burman and L. Oksanen. Data assimilation for the heat equation using stabilized finite element methods. *ArXiv e-prints*, Sept. 2016.
15. J. Cheng and M. Yamamoto. One new strategy for a priori choice of regularizing parameters in tikhonov's regularization. *Inverse Problems*, 16(4):L31, 2000.
16. D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2012.
17. H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
18. A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004.
19. J.-L. Guermond. Stabilization of Galerkin approximations of transport equations by subgrid modeling. *M2AN Math. Model. Numer. Anal.*, 33(6):1293–1316, 1999.
20. C. Johnson, U. Nävert, and J. Pitkäranta. Finite element methods for linear hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 45(1-3):285–312, 1984.
21. S. I. Kabanikhin. Definitions and examples of inverse and ill-posed problems. *J. Inverse Ill-Posed Probl.*, 16(4):317–357, 2008.
22. R. Lattès and J.-L. Lions. *Méthode de quasi-réversibilité et applications*. Travaux et Recherches Mathématiques, No. 15. Dunod, Paris, 1967.
23. J.-P. Puel. A nonstandard approach to a data assimilation problem and Tychonov regularization revisited. *SIAM J. Control Optim.*, 48(2):1089–1111, 2009.
24. L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.
25. A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill-posed problems*. V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York-Toronto, Ont.-London, 1977.
26. M. Yamamoto. Carleman estimates for parabolic equations and applications. *Inverse Problems*, 25(12):123013, 75, 2009.