INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus
Mathématique

Erik Burman, Guillaume Delay, Alexandre Ern and Lauri Oksanen

A stability estimate for data assimilation subject to the heat equation with initial datum

Volume 361 (2023), p. 1521-1530

https://doi.org/10.5802/crmath.506

CENTRE MERSENNE

Les Comptes Rendus. Mathématique *sont membres du Centre Mersenne pour l'édition scientifique ouverte*
www.centre-mersenne.org
e-ISSN : 1778-3569

Comptes Rendus Mathématique
2023, Vol. 361, p. 1521-1530
https://doi.org/10.5802/crmath.506

Partial differential equations, Control theory / *Équations aux dérivées partielles, Théorie du contrôle*

A stability estimate for data assimilation subject to the heat equation with initial datum

Une estimée de stabilité pour l'assimilation de données pour l'équation de la chaleur avec une donnée initiale

Erik Burman [a], Guillaume Delay [*, b], Alexandre Ern [c] and Lauri Oksanen [d]

[a] Department of Mathematics, University College London, London, UK–WC1E 6BT, UK

[b] Sorbonne Université, CNRS, Université Paris Cité, LJLL, F-75005 Paris, France

[c] CERMICS, École des Ponts, 77455 Marne-la-Vallée cedex 2, and INRIA, Paris, France

[d] University of Helsinki, Department of Mathematics and Statistics, P.O 68, 00014 University of Helsinki, Finland

E-mails: e.burman@ucl.ac.uk (E. Burman), guillaume.delay@sorbonne-universite.fr (G. Delay), alexandre.ern@enpc.fr (A. Ern), lauri.oksanen@helsinki.fi (L. Oksanen)

Abstract. This paper studies the unique continuation problem for the heat equation. We prove a so-called conditional stability estimate for the solution. We are interested in local estimates that are Hölder stable with the weakest possible norms of data on the right-hand side. Such an estimate is useful for the convergence analysis of computational methods dealing with data assimilation. We focus on the case of a known solution at initial time and in some subdomain but that is unknown on the boundary. To the best of our knowledge, this situation has not yet been studied in the literature.

Résumé. Cette contribution traite du problème de continuation unique pour l'équation de la chaleur. Nous prouvons une estimée conditionnelle de stabilité pour la solution de ce problème. Nous sommes intéressés par une estimée locale qui est Hölder-stable avec les normes les plus faibles possibles pour le terme de droite. Une telle estimée est utile pour l'analyse de convergence des méthodes de calcul traitant du problème d'assimilation de données. Nous nous intéressons en particulier au cas où la solution est connue à l'instant initial et dans un certain sous-domaine mais est inconnue sur la frontière du domaine. Cette situation ne semble pas avoir été déjà traitée dans la littérature.

2020 Mathematics Subject Classification. 35A23, 35B30, 35K05.

Funding. GD was supported by the Emergence grant of Sorbonne University. EB was supported by EPSRC grants EP/T033126/1 and EP/V050400/1.

Manuscript received 26 October 2022, revised 13 April 2023, accepted 14 April 2023.

* Corresponding author.

1. Introduction

The goal of the present work is to derive a conditional stability estimate for the data assimilation problem subject to the heat equation. This problem consists in finding the solution to the heat equation in a target subdomain with the knowledge of its values in another subdomain and of its initial datum. The main difficulty is that the boundary conditions of the problem are not known. This situation frequently arises in variational data assimilation, when a background state obtained from a previous assimilation cycle is available as (approximate) initial condition. Integration of such a background state is a requirement in weather forecasting, but also for optimization algorithms that divide the assimilation window into several shorter time intervals and perform assimilation on these intervals sequentially. Stability estimates for this situation do not appear to be available in the literature, and existing techniques cannot be adapted "off the shelf". Therefore, we give a self-contained proof with special care taken to design an estimate on a form that is readily applicable to the error analysis of numerical schemes in the spirit of [2,3,6].

More precisely, let $\Omega \subset \mathbb{R}^n$ ($n \in \{1,2,3\}$) be an open bounded set, let $\omega \subset\subset \Omega$ (i.e., $\overline{\omega} \subset \Omega$) be the open and non-empty subset where the solution is known, and let $T > 0$. We use the shorthand notation $L := \partial_t - \Delta$ for the space-time differential operator associated with the heat equation, $\dot{u} := \partial_t u$ for the time derivative, and $M := (0,T) \times \Omega$ for the space-time cylinder. We consider the following data assimilation problem: Find $u : M \to \mathbb{R}$ such that

$$L(u) = f \qquad \text{in } M, \tag{1}$$

$$u(0,\cdot) = u_0(\cdot) \qquad \text{in } \Omega, \tag{2}$$

$$u = g \qquad \text{in } (0,T) \times \omega, \tag{3}$$

where $f \in L^2(0,T;L^2(\Omega))$, $u_0 \in H^1(\Omega)$, and $g \in H^1(0,T;(H^1(\omega))') \cap L^2(0,T;H^1(\omega))$ are given. Notice that no information is given on the boundary $\partial\Omega$. We assume that f, u_0 and g are chosen so that there exists a solution to the data assimilation problem (1)-(3). Since this problem is ill-posed, however, one cannot hope for a stability estimate in the usual form. Nevertheless, one can derive a so-called conditional stability estimate which bounds the energy norm of the solution u in a target subdomain $B \subset\subset \Omega$ using (i) the measurements in $(0,T) \times \omega$; (ii) the initial datum u_0; (iii) the source term f; and (iv) an a priori bound on the solution under the form of its L^2-norm over the whole domain Ω. Our main result establishes Hölder stability of the solution to the data assimilation problem in the interior of the target space-time subdomain.

Theorem 1 (Three-cylinders inequality). *Let $\omega \subset\subset \Omega$ be open and non-empty, and let $0 < T_1 < T$. Let $B \subset\subset \Omega$ be open and connected. Then there are $C > 0$ and $\kappa \in (0,1)$ such that any space-time function u in the space*

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

satisfies

$$\|u\|_{L^2(0,T_1;H^1(B))} \le C(\|u\|_{L^2((0,T)\times\omega)} + F(u))^{\kappa}(\|u\|_{L^2((0,T)\times\Omega)} + F(u))^{1-\kappa}, \tag{5}$$

where

$$F(u) := \|u|_{t=0}\|_{L^2(\Omega)} + \|L(u)\|_{L^2(0,T;H^{-1}(\Omega))}. \tag{6}$$

The proof of Theorem 1, which hinges on a suitable pointwise Carleman estimate, is developed through the following two sections.

Carleman estimates for parabolic problems can be found in [1, 8, 10, 12, 15]. However, most works in the literature are concerned with the initialization problem [3, 6, 12], where boundary conditions are known, but not the initial condition. Here, we are instead interested in the opposite case, where the initialization problem has been solved and therefore the initial datum is known, but the boundary conditions are unknown. The estimate derived in Theorem 1 uses the initial datum in the upper bound and bounds the solution up to the initial time. Instead, in the usual setting in which the initial datum is unknown, the solution is estimated only in a space-time subdomain that is kept away from the initial time; see for instance [6, Theorems 1&2]. Furthermore, a similar control problem is considered in [7] where well-posedness is proven using Carleman estimates. The main difference with the present work is that the function that is estimated therein vanishes on the lateral boundary instead of the initial time; consequently, the weight function that is used in the proof is singular at the initial time.

The rest of this contribution is organized as follows. In Section 2, we prove a pointwise Carleman estimate. In Section 3, this estimate is used to prove a preliminary three-cylinders inequality. This result is then adapted to complete the proof of Theorem 1.

2. A pointwise Carleman estimate

The main result of this section is the pointwise Carleman estimate established in Lemma 2. We consider two functions $\rho \in C^3(\overline{M})$ and $w \in C^2(\overline{M})$ and a real number $\tau > 0$. Notice that the functions ρ and w both depend on time $t \in [0,T]$ and space $x \in \overline{\Omega}$, i.e., we have $\rho = \rho(t,x)$ and $w = w(t,x)$. In Section 3, the function w will be chosen in a specific way in terms of the function u in Theorem 1, and the real number τ will be chosen large enough. A specific choice for the function ρ will be made as well.

In what follows, we suppose that $\nabla\rho \neq 0$ in $\overline{M}$. We denote by $\lambda > 0$ a bound on $D^2\rho$ (uniformly in $\overline{M}$) and by $\theta > 0$ a real number such that $\theta \le |\nabla\rho|^2 \le \theta^{-1}$ (uniformly in $\overline{M}$). We fix a real number α such that $\alpha > 3\theta^{-3}\lambda$. Here, the differential operators ∇ and D^2 act only on the space variables. Moreover, we use the shorthand notation $Y \lesssim Z$ with positive real numbers Y, Z for the inequality $Y \le CZ$ where the value of the generic constant C can change at each occurrence provided it is independent of w and τ. The value of C can depend on ρ since this function will be fixed once and for all in Section 3.

We introduce the following functions that are defined using ρ, w, and τ:

$$\phi := e^{\alpha\rho}, \qquad \ell := \tau\phi, \qquad v := e^{\ell} w. \tag{7}$$

We also define the following auxiliary quantities:

$$a := 3\lambda\alpha\tau\phi, \quad \sigma := a + \Delta\ell, \quad q := a + |\nabla\ell|^2, \quad b := -\sigma v - 2(\nabla v, \nabla\ell), \quad c := (|\nabla v|^2 - q v^2)\nabla\ell, \tag{8a}$$

$$r := (\nabla\sigma, \nabla v)v + (\operatorname{div}(a\nabla\ell) - a\sigma)\, v^2, \quad Q := q + \dot{\ell}, \quad B := b - \dot{v}, \tag{8b}$$

$$R := r + \frac{1}{2}\dot{Q}v^2 + \operatorname{div}(\dot{\ell}\nabla\ell)v^2 - \sigma\dot{\ell}v^2. \tag{8c}$$

The main result of this section is that $|\nabla w|^2 + |w|^2$ (with weights depending on τ and ρ) can be upper bounded by $|L(w)|^2$ and additional terms subject to a divergence or a time derivative. This result already contains the structure for Theorem 1. Indeed, the divergence and time-derivative terms will disappear when an integration will be performed over M.

Lemma 2 (Pointwise Carleman estimate). *There is $\tau_0 > 0$ such that for all $\tau > \tau_0$ and all $w \in C^2(\overline{M})$, we have*

$$e^{2\tau\phi}(\tau|\nabla w|^2 + \tau^3|w|^2) \lesssim e^{2\tau\phi}|L(w)|^2 - \operatorname{div}(b\nabla v + c) + \operatorname{div}(\dot{v}\nabla v + v^2\dot{\ell}\nabla\ell) - \frac{1}{2}\partial_t(|\nabla v|^2 - Qv^2). \tag{9}$$

The rest of this section is devoted to proving Lemma 2.

2.1. *Preliminary results*

The first step towards proving Lemma 2 is to upper bound $e^{2\tau\phi}(\tau|\nabla w|^2+\tau^3|w|^2)$ as follows.

Lemma 3. *For all $\tau>0$ and all $w\in C^2(\overline{M})$, we have*

$$e^{2\tau\phi}(\tau|\nabla w|^2+\tau^3|w|^2)\lesssim a|\nabla v|^2+2D^2\ell(\nabla v,\nabla v)+(-a|\nabla\ell|^2+2D^2\ell(\nabla\ell,\nabla\ell))v^2. \tag{10}$$

Proof. We argue as in the proof of [5, Proposition 1]. For a vector $X\in\mathbb{R}^n$, we have

$$D^2\phi(X,X)=\alpha\phi(\alpha(\nabla\rho,X)^2+D^2\rho(X,X)). \tag{11}$$

Since the first term is positive, this implies

$$D^2\phi(X,X)\geq\alpha\phi D^2\rho(X,X)\geq-\lambda\alpha\phi|X|^2, \tag{12}$$

where we recall that $\lambda>0$ is an upper bound on $D^2\rho$ (uniformly in $\overline{M}$). Taking $X:=\nabla\ell=\alpha\tau\phi\nabla\rho$ in (11), we obtain

$$D^2\ell(\nabla\ell,\nabla\ell)=\tau\alpha\phi(\alpha(\nabla\rho,\nabla\ell)^2+D^2\rho(\nabla\ell,\nabla\ell))\geq(\alpha\tau\phi)^3(\alpha|\nabla\rho|^4-\lambda|\nabla\rho|^2).$$

Recalling the choice for the real number θ, this gives

$$D^2\ell(\nabla\ell,\nabla\ell)\geq(\alpha\tau\phi)^3(\alpha\theta^2-\lambda\theta^{-1}). \tag{13}$$

We also have

$$|\nabla\ell|^2=\tau^2|\nabla\phi|^2=(\alpha\tau\phi)^2|\nabla\rho|^2\leq(\alpha\tau\phi)^2\theta^{-1}. \tag{14}$$

Using (12) with $X:=\nabla v$, we get

$$2D^2\ell(\nabla v,\nabla v)\geq-2\lambda\alpha\tau\phi|\nabla v|^2.$$

Recalling that $a:=3\lambda\alpha\tau\phi$, we obtain

$$a|\nabla v|^2+2D^2\ell(\nabla v,\nabla v)\geq\lambda\alpha\tau\phi|\nabla v|^2. \tag{15}$$

Combining (13) and (14), we get

$$(-a|\nabla\ell|^2+2D^2\ell(\nabla\ell,\nabla\ell))v^2\geq\underbrace{(2\alpha\theta^2-5\lambda\theta^{-1})}_{=:c(\alpha)}(\alpha\tau\phi)^3v^2. \tag{16}$$

Summing (15) and (16), we see that

$$c(\alpha)(\alpha\tau\phi)^3v^2+\lambda\alpha\tau\phi|\nabla v|^2\leq a|\nabla v|^2+2D^2\ell(\nabla v,\nabla v)+(-a|\nabla\ell|^2+2D^2\ell(\nabla\ell,\nabla\ell))v^2. \tag{17}$$

Notice that the right-hand side of (17) is the one we have in the statement of Lemma 3. To bound the left-hand side of (17) from below in terms of $|\nabla w|^2$ and w^2 we notice that

$$|\nabla v|^2=e^{2\tau\phi}|\tau w\nabla\phi+\nabla w|^2\geq e^{2\tau\phi}\frac{1}{2}|\nabla w|^2-e^{2\tau\phi}|\nabla\phi|^2\tau^2w^2, \tag{18}$$

where we used the Cauchy–Schwarz and Young inequalities. Owing to the choice $\alpha>3\theta^{-3}\lambda$, we infer that

$$c(\alpha)(\alpha\phi)^3-\lambda\alpha\phi|\nabla\phi|^2\geq(\alpha\phi)^3(2\alpha\theta^2-6\lambda\theta^{-1})=:c_\alpha>0, \tag{19}$$

where the first inequality comes from the definition of $c(\alpha)$ and $-\lambda\alpha\phi|\nabla\phi|^2\geq-\lambda\theta^{-1}\alpha^3\phi^3$ (as a consequence of (14)) and the second inequality comes from $\alpha>3\theta^{-3}\lambda$. Multiplying (18) by $\lambda\alpha\tau\phi$, multiplying (19) by $\tau^3v^2=e^{2\tau\phi}\tau^3w^2$ and summing both equations gives

$$\lambda\alpha\tau\phi e^{2\tau\phi}\frac{1}{2}|\nabla w|^2+c_\alpha e^{2\tau\phi}\tau^3w^2\leq c(\alpha)(\alpha\tau\phi)^3v^2+\lambda\alpha\tau\phi|\nabla v|^2.$$

Invoking (17), we get

$$c_\rho e^{2\tau\phi}(\tau|\nabla w|^2+\tau^3|w|^2)\leq a|\nabla v|^2+2D^2\ell(\nabla v,\nabla v)+(-a|\nabla\ell|^2+2D^2\ell(\nabla\ell,\nabla\ell))v^2,$$

where $c_\rho=\min(c_\alpha,\lambda\alpha\inf_{(x,t)\in\overline{M}}|\phi(x,t)|/2)>0$. This ends the proof. □

The next step is to observe that the terms on the right-hand side of (10) are equal to a weighted square norm of the heat operator plus some more terms.

Lemma 4. *For all $\tau > 0$ and all $w \in C^2(\overline{M})$, we have*

$$\begin{aligned} e^{2\ell}|L(w)|^2/2 = {} & (\Delta v + Qv)^2/2 + B^2/2 \\ & + a|\nabla v|^2 + 2D^2\ell(\nabla v, \nabla v) + \left(-a|\nabla\ell|^2 + 2D^2\ell(\nabla\ell,\nabla\ell)\right) v^2 \\ & + \operatorname{div}(b\nabla v + c) - \operatorname{div}(\dot v\nabla v + v^2\dot\ell\nabla\ell) + \partial_t(|\nabla v|^2 - Qv^2)/2 + R. \end{aligned}$$

Proof. We have

$$e^{2\ell}|L(w)|^2/2 = e^{2\ell}|\Delta w|^2/2 + e^{2\ell}|\dot w|^2/2 - e^{2\ell}\dot w\Delta w.$$

Moreover, $\dot v = \partial_t(e^\ell w) = e^\ell \dot w + \dot\ell v$. Hence, $-e^\ell \dot w = -\dot v + \dot\ell v$. Straightforward computations give

$$e^\ell\Delta w = \Delta v - \sigma v - 2(\nabla v, \nabla\ell) + \sigma v - (\Delta\ell)v + |\nabla\ell|^2 v = \Delta v + b + qv.$$

Combining previous relations, we get

$$e^{2\ell}|L(w)|^2/2 = e^{2\ell}|\Delta w|^2/2 + |\dot v - \dot\ell v|^2/2 + (-\dot v + \dot\ell v)(\Delta v + b + qv).$$

We now invoke [4, Lemma 1] (setting $k := 0$ therein) to obtain

$$\begin{aligned} e^{2\ell}|\Delta w|^2/2 = {} & (\Delta v + qv)^2/2 + b^2/2 \\ & + a|\nabla v|^2 + 2D^2\ell(\nabla v, \nabla v) + \left(-a|\nabla\ell|^2 + 2D^2\ell(\nabla\ell,\nabla\ell)\right) v^2 \\ & + \operatorname{div}(b\nabla v + c) + r. \end{aligned}$$

Using this identity and recalling that $Q := q + \dot\ell$ and $B := b - \dot v$, we get

$$\begin{aligned} e^{2\ell}|L(w)|^2/2 & = (\Delta v + qv)^2/2 + b^2/2 + (\dot v^2 + (\dot\ell v)^2)/2 - \dot v\dot\ell v + (\dot\ell v - \dot v)(\Delta v + b + qv) \\ & \quad + a|\nabla v|^2 + 2D^2\ell(\nabla v, \nabla v) + \left(-a|\nabla\ell|^2 + 2D^2\ell(\nabla\ell,\nabla\ell)\right) v^2 + \operatorname{div}(b\nabla v + c) + r \\ & = (\Delta v + Qv)^2/2 + B^2/2 - \dot\ell v(\Delta v + qv) + b\dot v - \dot v\dot\ell v + (\dot\ell v - \dot v)(\Delta v + b + qv) \\ & \quad + a|\nabla v|^2 + 2D^2\ell(\nabla v, \nabla v) + \left(-a|\nabla\ell|^2 + 2D^2\ell(\nabla\ell,\nabla\ell)\right) v^2 + \operatorname{div}(b\nabla v + c) + r \\ & = (\Delta v + Qv)^2/2 + B^2/2 - \dot v\dot\ell v + \dot\ell v b - \dot v(\Delta v + qv) \\ & \quad + a|\nabla v|^2 + 2D^2\ell(\nabla v, \nabla v) + \left(-a|\nabla\ell|^2 + 2D^2\ell(\nabla\ell,\nabla\ell)\right) v^2 + \operatorname{div}(b\nabla v + c) + r. \end{aligned} \tag{20}$$

It remains to rewrite $-\dot v\dot\ell v + \dot\ell v b - \dot v(\Delta v + qv)$. We have

$$-\dot v\Delta v = -\operatorname{div}(\dot v\nabla v) + (\nabla v, \nabla\dot v) = -\operatorname{div}(\dot v\nabla v) + \partial_t|\nabla v|^2/2, \tag{21}$$

$$-\dot v\dot\ell v - qv\dot v = -Qv\dot v = -Q\partial_t(v^2/2) = -\partial_t(Qv^2)/2 + \dot Q v^2/2. \tag{22}$$

Moreover, we have $-2(\nabla v, \nabla\ell)\dot\ell v = -(\nabla(v^2), \dot\ell\nabla\ell) = v^2\operatorname{div}(\dot\ell\nabla\ell) - \operatorname{div}(v^2\dot\ell\nabla\ell)$ and recalling that $b := -\sigma v - 2(\nabla v, \nabla\ell)$, we obtain

$$\dot\ell v b = \dot\ell v(-\sigma v - 2(\nabla v, \nabla\ell)) = -\sigma\dot\ell v^2 + v^2\operatorname{div}(\dot\ell\nabla\ell) - \operatorname{div}(v^2\dot\ell\nabla\ell). \tag{23}$$

The claim follows by injecting (21), (22) and (23) into (20). □

2.2. *Proof of Lemma 2*

The proof of Lemma 2 combines the results of Lemmas 3 and 4. We still argue as the proof of [5, Prop. 1]. Rewriting the result of Lemma 4 under the form

$$\begin{aligned} & a|\nabla v|^2 + 2D^2\ell(\nabla v, \nabla v) + \left(-a|\nabla\ell|^2 + 2D^2\ell(\nabla\ell,\nabla\ell)\right) v^2 \\ & \qquad \le e^{2\ell}|L(w)|^2/2 - \operatorname{div}(b\nabla v + c) + \operatorname{div}(\dot v\nabla v + v^2\dot\ell\nabla\ell) - \partial_t(|\nabla v|^2 - Qv^2)/2 - R, \end{aligned}$$

and applying this bound to the right-hand side of (10), we infer that

$$e^{2\tau\phi}(\tau|\nabla w|^2 + \tau^3|w|^2) \\ \lesssim e^{2\ell}(Lw)^2/2 - \operatorname{div}(b\nabla v + c) + \operatorname{div}(\dot{v}\nabla v + v^2\dot{\ell}\nabla\ell) - \partial_t(|\nabla v|^2 - Qv^2)/2 - R. \quad (24)$$

It only remains to bound R. We recall that

$$R = (\nabla\sigma, \nabla v)v + (\operatorname{div}(a\nabla\ell) - a\sigma)\,v^2 + \dot{Q}v^2/2 + \operatorname{div}(\dot{\ell}\nabla\ell)\,v^2 - \sigma\dot{\ell}v^2.$$

For $\tau > 0$ large enough, we have (recall that the value of the hidden constant C in any inequality of the form $X \lesssim Y$ can depend on ρ but is independent of τ and w)

$$\begin{aligned} \operatorname{div}(a\nabla\ell) - a\sigma &= (\nabla a, \nabla\ell) - a^2 = 3\lambda\alpha\tau^2|\nabla\phi|^2 - (3\lambda\alpha\tau)^2\phi^2 \lesssim \tau^2, \\ \dot{Q} = \dot{a} + 2(\nabla\dot{\ell}, \nabla\ell) + \ddot{\ell} &= 3\lambda\alpha\tau\dot{\phi} + 2\tau^2(\nabla\phi, \nabla\dot{\phi}) + \tau\ddot{\phi} \lesssim \tau^2, \\ \operatorname{div}(\dot{\ell}\nabla\ell) &= \tau^2\operatorname{div}(\dot{\phi}\nabla\phi) \lesssim \tau^2, \\ \sigma\dot{\ell} &= \tau^2(3\lambda\alpha\phi + \Delta\phi)\dot{\phi} \lesssim \tau^2, \end{aligned}$$

and

$$(\nabla\sigma, \nabla v)v = \tau(3\alpha\lambda\nabla\phi + \nabla\Delta\phi, \nabla w e^{\ell} + \nabla\ell e^{\ell} w)e^{\ell} w \\ \lesssim \tau(|\nabla w|e^{\ell} + |\nabla\ell|e^{\ell}|w|)e^{\ell}|w| \lesssim e^{2\tau\phi}(\tau|\nabla w||w| + \tau^2|w|^2) \lesssim e^{2\tau\phi}(|\nabla w|^2 + \tau^2|w|^2).$$

Thus, for $\tau > 0$ large enough, we have

$$|R| \lesssim e^{2\tau\phi}(|\nabla w|^2 + \tau^2|w|^2).$$

This shows that for $\tau > 0$ large enough, R can be absorbed in the left-hand side of (24). The proof is complete.

3. Proof of Theorem 1

The goal of this section is to prove Theorem 1. First, using the pointwise Carleman estimate from Section 2, we establish a preliminary three-cylinders inequality (Proposition 5). Then we conclude the proof of Theorem 1 by improving on the norms used on the right-hand side of the preliminary three-cylinders inequality.

3.1. *Preliminary three-cylinders inequality*

In the earlier work [6], a three-cylinders inequality was proved via a reduction to Isakov's Carleman estimate [9]. The change in the present three-cylinders inequality is that we prove stability up to $t = 0$, to the price of requiring an estimate on the initial datum in the right-hand side. The proof by Isakov does not keep track of the boundary term at $t = 0$, and for this reason we are forced to give a full proof of an analogous Carleman estimate that handles that term (Lemma 2 and its integrated version (26)). Notice also that the time-dependent part of our weight function is different from that in [6], since we want to provide stability up to $t = 0$. The same observation also holds for the cutoff function. In addition, contrary to Isakov, we derive our bound starting from a pointwise Carleman estimate (Lemma 2) that is not yet available in the context of the heat equation to the best of our knowledge.

Carleman estimates with boundary terms are typically somewhat more complicated to prove than those for compactly supported functions. Often the former estimates are called global and the latter local. In this sense, the estimate in the present paper is local in space but global in time. To our knowledge, this combination has not been studied previously. Yet, it is of interest in applications, for example those in weather forecasting, where repeated data assimilation tasks require a background state to initialize the simulation.

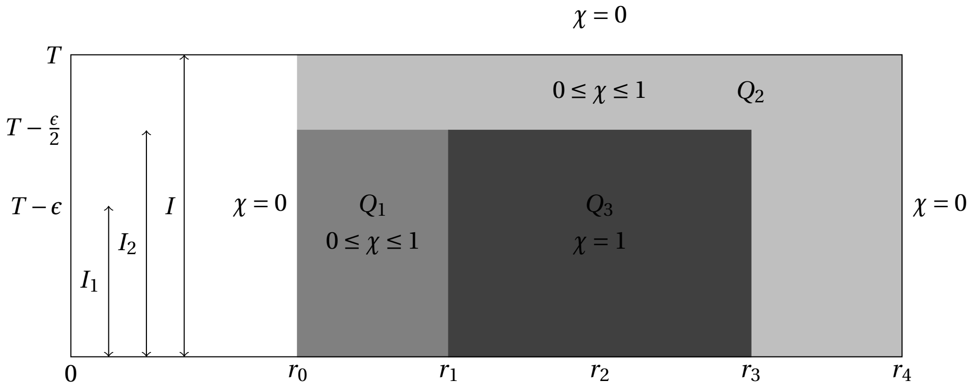

Figure 1. Choice of the cutoff function χ.

Proposition 5 (Preliminary three-cylinders inequality). *Let* $x_0 \in \Omega$ *and* $0 < r_1 < r_2 < d(x_0, \partial\Omega)$. *Write* $B_j = B(x_0, r_j)$, $j \in \{1,2\}$. *Let* $0 < \epsilon < T$. *Then there are* $C > 0$ *and* $\kappa \in (0,1)$ *such that for all* $u \in C^2(\mathbb{R} \times \Omega)$,

$$\|u\|_{L^2(0,T-\epsilon;H^1(B_2))} \le C(\|u\|_{L^2(0,T;H^1(B_1))} + \|L(u)\|_{L^2((0,T)\times\Omega)} + \|u|_{t=0}\|_{H^1(\Omega)})^\kappa \|u\|^{1-\kappa}_{L^2(0,T;H^1(\Omega))}. \quad (25)$$

Proof. The idea is to integrate the pointwise Carleman estimate (9) using adequate functions ρ and w.

Step 1. Choice of the function ρ**.** Let $0 < r_0 < r_1$ and $r_2 < r_3 < r_4 < d(x_0, \partial\Omega)$. Define $B_j = B(x_0, r_j)$, $j \in \{0,3,4\}$. We choose a function $\rho_1 \in C^\infty(\Omega)$ such that $-r_0/2 \ge \rho_1 > -r_0$ in B_0 and that $\rho_1(x) = -d(x, x_0)$ outside B_0 (notice that $\rho_1 < 0$). Setting $I := (0,T)$, $I_1 := (0, T-\epsilon)$, and $I_2 := (0, T-\epsilon/2)$, we choose a function $\rho_2 \in C^\infty(\mathbb{R})$ such that $\rho_2(s) \le -r_3$ for $s \ge T - \epsilon/2$ and $\rho_2 = 0$ in I_1 (notice that $\rho_2 \le 0$). We define $\rho(t,x) := \rho_1(x) + \rho_2(t)$. Notice that $|\nabla\rho| = 1$ outside B_0. We use the notation (see Figure 1)

$$Q_1 := I_2 \times (B_1 \setminus B_0), \quad Q_2 := ((I \setminus I_2) \times (B_4 \setminus B_0)) \cup (I \times (B_4 \setminus B_3)), \quad Q_3 := I_2 \times (B_3 \setminus B_1).$$

We also define $\Phi(r) := e^{-\alpha r}$. Recalling that $\phi = e^{\alpha\rho}$, we observe that the following bounds hold true:

$$\begin{aligned} \phi &\le \Phi(r_3) \quad \text{in } Q_2, \\ \phi &\ge \Phi(r_2) \quad \text{in } I_1 \times (B_2 \setminus B_1). \end{aligned}$$

Indeed, the first bound is a consequence of the fact that $\rho_2(s) \le -r_3$ for $s \ge T - \epsilon/2$ and $\rho_1(x) = -d(x, x_0)$ outside B_0. The second bound comes from $\rho_2 = 0$ in I_1 and $\rho_1(x) \ge -r_2$ in $B_2 \setminus B_1$.

Step 2. Choice of the function w**.** We define a cutoff function $\chi \in C_0^\infty((-1,T) \times (B_4 \setminus B_0))$ that satisfies $\chi = 1$ in Q_3 and $0 \le \chi \le 1$ in $(-1,T) \times B_4$; see Figure 1. We then set $w := \chi u$.

Step 3. Integrating (9) over $I \times \Omega$ and observing that $\nabla\rho$ does not vanish on $I \times (B_4 \setminus B_0)$, we get for all $\tau > 0$ large enough,

$$\begin{aligned} \int_0^T \int_\Omega (\tau|\nabla w|^2 + \tau^3|w|^2) e^{2\tau\phi} \,\mathrm{d}x\mathrm{d}t &= \int_0^T \int_{B_4 \setminus B_0} (\tau|\nabla w|^2 + \tau^3|w|^2) e^{2\tau\phi} \,\mathrm{d}x\mathrm{d}t \\ &\lesssim \int_0^T \int_{B_4 \setminus B_0} |L(w)|^2 e^{2\tau\phi} \,\mathrm{d}x\mathrm{d}t + \int_{B_4 \setminus B_0} (|\nabla w|^2 + \tau^2|w|^2) e^{2\tau\phi} \mathrm{d}x|_{t=0}. \quad (26) \end{aligned}$$

Indeed, the divergence terms on the right-hand side of (9) disappear since $\int_{B_4\setminus B_0}\operatorname{div}(A) = \int_{\partial(B_4\setminus B_0)} An = 0$ for any A that fulfills $A=0$ on $\partial(B_4\setminus B_0)$. Moreover, concerning the time-derivative on the right-hand side of (9), we used that for all $\tau>0$ large enough, $Q = a+|\nabla\ell|^2+\dot{\ell} \lesssim \tau^2$ so that

$$-\int_0^T\int_{B_4\setminus B_0}\partial_t(|\nabla v|^2-Qv^2)\mathrm{d}x\mathrm{d}t = \int_{B_4\setminus B_0}(|\nabla v|^2-Qv^2)\mathrm{d}x|_{t=0} \lesssim \int_{B_4\setminus B_0}(|\nabla w|^2+\tau^2|w|^2)e^{2\tau\phi}\mathrm{d}x|_{t=0},$$

where the hidden constant depends in particular on the first-order derivatives of ρ in time and in space.

Step 4. The next step is to replace w by u in (26). To this end, we define the commutator

$$[L,\chi](u) := L(\chi u)-\chi L(u) = (\partial_t\chi-\Delta\chi)u-2(\nabla u,\nabla\chi). \tag{27}$$

Since $L(w) = L(\chi u) = \chi L(u)+[L,\chi](u)$, we have $|L(w)|^2 \lesssim |L(u)|^2+|[L,\chi](u)|^2$. Moreover, the commutator $[L,\chi]$ vanishes on Q_3 since $\chi=1$ on Q_3. Therefore, the first term on the right-hand side of (26) satisfies

$$\begin{aligned}\int_0^T\int_{B_4\setminus B_0}|L(w)|^2e^{2\tau\phi}\,\mathrm{d}x\mathrm{d}t &\lesssim \int_{(0,T)\times B_4}|L(u)|^2e^{2\tau\phi}\mathrm{d}x\mathrm{d}t+\int_{Q_1\cup Q_2}|[L,\chi](u)|^2e^{2\tau\phi}\mathrm{d}x\mathrm{d}t\\ &\lesssim e^{2\tau}\|L(u)\|^2_{L^2((0,T)\times B_4)}+e^{2\tau}\|u\|^2_{L^2(0,T;H^1(B_1))}+e^{2\tau\Phi(r_3)}\|u\|^2_{L^2(0,T;H^1(B_4))},\end{aligned} \tag{28}$$

where we used $\phi\le 1$, $\phi\le\Phi(r_3)$ in Q_2, and the fact that $[L,\chi](u)$ can be bounded by L^2- and H^1-norms of u owing to (27). (Notice that the hidden constant above depends on the first-order derivatives in time and first- and second-order derivatives in space of χ. Therefore, the constant in (28) blows up as $\epsilon\to 0$, or if $r_4\to d(x_0,\partial\Omega)$.) Consider now the second term on the right-hand side of (26). First, we notice that $\|\nabla w\|_{L^2(B_4\setminus B_0)} \lesssim \|u\|_{H^1(B_4\setminus B_0)}$. For the low-order term, we observe that $\rho\le -r_0/2$ implies $\phi<1$ and, therefore, $\tau^2e^{2\tau\phi}\lesssim e^{2\tau}$ for $\tau>0$ large enough. It follows that for $\tau>0$ large enough,

$$\int_{B_4\setminus B_0}(|\nabla w|^2+\tau^2|w|^2)e^{2\tau\phi}\mathrm{d}x|_{t=0} \lesssim e^{2\tau}\|u|_{t=0}\|^2_{H^1(\Omega)}. \tag{29}$$

Furthermore, using $\chi=1$, i.e. $w=u$, as well as $\phi\ge\Phi(r_2)$ in $I_1\times(B_2\setminus B_1)$, we infer that, for $\tau\ge1$, the left-hand side of (26) can be bounded from below by

$$\int_{I_1\times(B_2\setminus B_1)}\left(\tau|\nabla u|^2+\tau^3|u|^2\right)e^{2\tau\phi}\,\mathrm{d}x\mathrm{d}t \ge e^{2\tau\Phi(r_2)}\|u\|^2_{L^2(0,T-\epsilon;H^1(B_2\setminus B_1))}. \tag{30}$$

Altogether, the inequalities (26) and (28)-(30) imply that, for $\tau>0$ large enough,

$$\begin{aligned}&\|u\|_{L^2(0,T-\epsilon;H^1(B_2))}\\ &\qquad\lesssim e^{\tau}\left(\|L(u)\|_{L^2((0,T)\times B_4)}+\|u\|_{L^2(0,T;H^1(B_1))}+\|u|_{t=0}\|_{H^1(\Omega)}\right)+e^{-p\tau}\|u\|_{L^2(0,T;H^1(B_4))},\end{aligned}$$

with $p:=\Phi(r_2)-\Phi(r_3)>0$. Here, we used that $e^{2\tau(1-\Phi(r_2))}<e^{2\tau}$ since $\Phi(r_2)>0$.

Step 5. Finally, the claim follows by a direct application of [11, Lemma 5.2], the idea being to optimize in τ, under the constraint $\tau\ge\tau_0$ for some $\tau_0>0$ large enough (see also [14]). □

3.2. *End of the proof*

In this last step of the proof of Theorem 1, we improve the norms on the right-hand side of (25) to $\|u|_{t=0}\|_{L^2(\Omega)}$, $\|L(u)\|_{L^2(0,T;H^{-1}(\Omega))}$, $\|u\|_{L^2((0,T)\times\omega)}$, and $\|u\|_{L^2((0,T)\times\Omega)}$.

Step 1. We set $T_1 := T - \epsilon$ for some $0 < \epsilon < T$. Upon replacing ω by a smaller set, we may assume without loss of generality that it is a ball of the form $B_1 := B(x_0, r_1)$ for some $x_0 \in \Omega$. We will show the following local version of the conditional stability estimate (5) where B is replaced by a ball of the form $B_2 := B(x_0, r_2)$ with $0 < r_1 < r_2 < d(x_0, \partial\Omega)$: For all $u \in C^\infty(\mathbb{R} \times \Omega)$,

$$\|u\|_{L^2(0,T_1;H^1(B_2))} \lesssim (\|u\|_{L^2((0,T)\times B_1)} + F(u))^\kappa (\|u\|_{L^2((0,T)\times\Omega)} + F(u))^{1-\kappa}, \tag{31}$$

where $F(u) := \|u|_{t=0}\|_{L^2(\Omega)} + \|L(u)\|_{L^2(0,T;H^{-1}(\Omega))}$. The general case for B follows by covering B by a finite chain of balls starting from ω, and by iterating the local result (see [13]), up to some rescaling in ϵ. Since smooth functions are dense in the space defined in (4), it is sufficient to consider the case where $u \in C^\infty(\mathbb{R} \times \Omega)$. In the rest of this proof, we consider $B_0 := B(x_0, r_0)$ and $B_3 := B(x_0, r_3)$ with $0 < r_0 < r_1$ and $r_2 < r_3 < d(x_0, \partial\Omega)$.

Step 2. Let us first weaken the norm of $L(u)$ and u_0. To this end, let $w \in L^2(0,T;H^1(B_3)) \cap H^1(0,T;H^{-1}(B_3))$ solve

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

and set $v := u|_{(0,T)\times B_3} - w \in L^2(0,T;H^1(B_3)) \cap H^1(0,T;H^{-1}(B_3))$. Since $L(v) = 0$ and $v|_{t=0} = 0$, Proposition 5 (with B_0 instead of B_1 and with B_3 instead of Ω) implies that

$$\|v\|_{L^2(0,T_1;H^1(B_2))} \le C\|v\|^\kappa_{L^2(0,T;H^1(B_0))}\|v\|^{1-\kappa}_{L^2(0,T;H^1(B_3))}. \tag{32}$$

Using the standard energy estimate for the heat equation on w, $\|w\|_{L^2(0,T;H^1(B_3))} \lesssim F$ with $F := \|u|_{t=0}\|_{L^2(\Omega)} + \|L(u)\|_{L^2(0,T;H^{-1}(\Omega))}$ together with the triangle inequality, we infer that

$$\|v\|_{L^2(0,T_1;H^1(B_i))} \lesssim \|u\|_{L^2(0,T_1;H^1(B_i))} + F, \quad i \in \{0,3\}, \tag{33}$$

and

$$\|u\|_{L^2(0,T_1;H^1(B_2))} \lesssim \|v\|_{L^2(0,T_1;H^1(B_2))} + F. \tag{34}$$

Applying (32) to the right-hand side of (34) and (33) to the right-hand side of (32) we obtain

$$\|u\|_{L^2(0,T_1;H^1(B_2))} \lesssim (\|u\|_{L^2(0,T;H^1(B_0))} + F)^\kappa (\|u\|_{L^2(0,T;H^1(B_3))} + F)^{1-\kappa}.$$

Step 3. Let us finally weaken the norms of $u|_{(0,T)\times B_0}$ and $u|_{(0,T)\times B_3}$. Choosing $\chi \in C_0^\infty(B_1)$ such that $\chi = 1$ in B_0, we see that χu satisfies

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

Since $[L,\chi](u)$ is of first-order in space and zeroth-order in time (with respect to u), standard energy estimates yield the following bounds:

$$\begin{aligned} \|u\|_{L^2(0,T;H^1(B_0))} = \|\chi u\|_{L^2(0,T;H^1(B_0))} &\le \|\chi u\|_{L^2(0,T;H^1(B_1))} \\ &\lesssim \|\chi u|_{t=0}\|_{L^2(B_1)} + \|L(\chi u)\|_{L^2(0,T;H^{-1}(B_1))} \\ &\lesssim F + \|[L,\chi](u)\|_{L^2(0,T;H^{-1}(B_1))} \lesssim F + \|u\|_{L^2((0,T)\times B_1)}. \end{aligned}$$

Reasoning analogously, we obtain

$$\|u\|_{L^2(0,T;H^1(B_3))} \lesssim F + \|u\|_{L^2((0,T)\times\Omega)}.$$

Putting everything together establishes (31). This ends the proof. □

References

[1] L. Bourgeois, "Quantification of the unique continuation property for the heat equation", *Math. Control Relat. Fields* **7** (2017), no. 3, p. 347-367.

[2] E. Burman, G. Delay, A. Ern, "The unique continuation problem for the heat equation discretized with a high-order space-time nonconforming method", *SIAM J. Numer. Anal.* **61** (2023), no. 5, p. 2534-2557.

[3] E. Burman, J. Ish-Horowicz, L. Oksanen, "Fully discrete finite element data assimilation method for the heat equation", *ESAIM, Math. Model. Numer. Anal.* **52** (2018), no. 5, p. 2065-2082.

[4] E. Burman, M. Nechita, L. Oksanen, "Unique continuation for the Helmholtz equation using stabilized finite element methods", *J. Math. Pures Appl.* **129** (2019), p. 1-22.

[5] ———, "A stabilized finite element method for inverse problems subject to the convection-diffusion equation. I: diffusion-dominated regime", *Numer. Math.* **144** (2020), p. 451-477.

[6] E. Burman, L. Oksanen, "Data assimilation for the heat equation using stabilized finite element methods", *Numer. Math.* **139** (2018), p. 505-528.

[7] O. Y. Èmanuilov, "Controllability of parabolic equations", *Mat. Sb.* **186** (1995), no. 6, p. 109-132.

[8] O. Y. Imanuvilov, M. Yamamoto, "Carleman estimate for a parabolic equation in a Sobolev space of negative order and its applications", in *Control of nonlinear distributed parameter systems (College Station, TX, 1999)*, Lecture Notes in Pure and Applied Mathematics, vol. 218, Marcel Dekker, 2001, p. 113-137.

[9] V. Isakov, "Carleman type estimates in an anisotropic case and applications", *J. Differ. Equations* **105** (1993), p. 217-238.

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

[11] J. Le Rousseau, G. Lebeau, "On Carleman estimates for elliptic and parabolic operators. Applications to unique continuation and control of parabolic equations", *ESAIM, Control Optim. Calc. Var.* **18** (2012), no. 3, p. 712-747.

[12] J.-P. Puel, "A nonstandard approach to a data assimilation problem and Tychonov regularization revisited", *SIAM J. Control Optim.* **48** (2009), no. 2, p. 1089-1111.

[13] L. Robbiano, "Théorème d'unicité adapté au contrôle des solutions des problèmes hyperboliques", *Commun. Partial Differ. Equations* **16** (1991), no. 4-5, p. 789-800.

[14] ———, "Fonction de coût et contrôle des solutions des équations hyperboliques", *Asymptotic Anal.* **10** (1995), no. 2, p. 95-115.

[15] M. Yamamoto, "Carleman estimates for parabolic equations and applications", *Inverse Probl.* **25** (2009), no. 12, article no. 123013 (75 pages).