Systems & Control Letters 158 (2021) 105055

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Averaged dynamics and control for heat equations with random diffusion

Jon Asier Bárcena-Petisco [a,b,*,1], Enrique Zuazua [c,b,a,1]

[a] Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid, Spain
[b] Chair of Computational Mathematics, Fundación Deusto, Avenida de las Universidades 24, 48007 Bilbao, Basque Country, Spain
[c] Chair of Dynamics, Control and Numerics, Alexander von Humboldt-Professorship, Department of Data Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany

ARTICLE INFO

Article history:
Received 6 October 2020
Received in revised form 20 July 2021
Accepted 27 September 2021
Available online 30 October 2021

Keywords:
Averaged controllability
Averaged observability
Observability
Random heat equation

ABSTRACT

This paper deals with the averaged dynamics for heat equations in the degenerate case where the diffusivity coefficient, assumed to be constant, is allowed to take the null value. First we prove that the averaged dynamics is analytic. This allows to show that, most often, the averaged dynamics enjoys the property of unique continuation and is approximately controllable. We then determine if the averaged dynamics is actually null controllable or not depending on how the density of averaging behaves when the diffusivity vanishes. In the critical density threshold the dynamics of the average is similar to the $\frac{1}{2}$-fractional Laplacian, which is well-known to be critical in the context of the controllability of fractional diffusion processes. Null controllability then fails (resp. holds) when the density weights more (resp. less) in the null diffusivity regime than in this critical regime.

1. Introduction

We analyse the problem of controlling the averaged value of the heat equation with random diffusion. This problem is relevant in applications in which the control has to be chosen independently of the random value, in a robust way. This problem has been studied in the literature in bounded domains and with diffusivities independent of the space and time variables and with a strictly positive minimum common to almost every realization. Notably, in [1,2] the authors consider diffusivities which follow the uniform and exponential probability distributions respectively, whereas a more general study is done in [3]. In those papers it is shown that, under those assumptions, the averaged dynamics inherits many properties from the dynamics of the heat equation (regularity, controllability, observability, etc.), with the only notable exception of the semi-group property. This is done by considering the Fourier representation of the averaged solutions.

In this paper we pursue the study to diffusivities which are allowed to take any positive value. In this scenario the averaged dynamics is still analytic (see Proposition 4.1), and we prove that the averaged dynamics is approximately controllable provided that we have a hierarchic decay in the time variable of the different frequencies. However, the averaged dynamics may acquire a fractional nature, or an even less diffusive one, so it may not be null controllable. What determines if we can control it is how fast the density of averaging decays when the diffusivity α vanishes. In the critical threshold, which is given by all the random variables whose density functions $\rho(\alpha)$ decay like $e^{-C_\rho\alpha^{-1}}$ for some $C_\rho > 0$ when $\alpha \to 0$, the dynamics of the average is similar to the $\frac{1}{2}$-fractional Laplacian, which is well-known to be critical in the context of controllability of fractional diffusion processes.

The mathematical model and main results

In this paper we treat the random heat equation described by the following system:

$$\begin{cases} y_t - \alpha\Delta y = g, & \text{in } (0,T)\times G,\\ y = h, & \text{on } (0,T)\times\partial G,\\ y(0,\cdot) = y^0, & \text{on } G, \end{cases} \tag{1.1}$$

for G a domain, g a source term, h the Dirichlet boundary conditions, y^0 the initial configuration and α the diffusivity coefficient, which is a positive random variable with density function ρ (the regime in which α is allowed to take negative values is studied

* Corresponding author at: Departamento de Matemáticas, Universidad Autónoma de Madrid, 28049 Madrid, Spain.
E-mail addresses: jonasier.barcena@ehu.eus (J.A. Bárcena-Petisco), enrique.zuazua@fau.de (E. Zuazua).

[1] The two authors have contributed equally on proving the main and the auxiliary results, on performing the simulations and on the composition of the paper.

https://doi.org/10.1016/j.sysconle.2021.105055
0167-6911/

in Section 8). The averaged solution of (1.1) is given by:

$$\tilde{y}(t,x;y^0,g,h) := \int_0^{+\infty} y(t,x;\alpha,y^0,g,h)\rho(\alpha)d\alpha.$$

Moreover, we can model a control f located in $G_0 \subset G$ or on $\Gamma \subset \partial G$ by posing $g = f1_{G_0}$ or $h = f1_{\Gamma}$ respectively.

In order to study (1.1) we rely on being the eigenfunctions of $-\alpha\Delta$ independent of α, as this allows us to work with the Fourier representation of the averaged solution. Thus, we may not use the same techniques in more general heat equations, like:

$$y_t - \mathrm{div}(\sigma(x,\alpha)\nabla y) + A(x,\alpha)\cdot\nabla y + a(x,\alpha)y = 0. \tag{1.2}$$

Indeed, even if we assume that $w \mapsto -\mathrm{div}(\sigma(x,\alpha)\nabla w) + A(x,\alpha)\cdot\nabla w + a(x,\alpha)w$ can be diagonalized, its eigenfunctions depend on α and, in particular, the averages of the eigenfunctions with respect to α may not form an orthogonal set. However, by studying the dynamics and controllability of (1.1) we highlight some of the most fundamental phenomena involving (1.2). The techniques presented in this paper also work for heat equations of the type:

$$y_t - \alpha\mathcal{L}y = 0,$$

for $\mathcal{L}$ any self-adjoint elliptic operator of order 2 with compact resolvent.

Since studying controllability with internal or boundary controls is almost equivalent, this paper is mainly devoted to controllability with an internal control and the few differences are explained in Section 8. In addition, to study the controllability properties of (1.1) we follow the classical duality approach (see Section 7.1 for further details) and focus on the observability properties of its adjoint system, which is given by:

$$\begin{cases} -\varphi_t - \alpha\Delta\varphi = 0, & \text{in } (0,T)\times G, \\ \varphi = 0, & \text{on } (0,T)\times\partial G, \\ \varphi(T,\cdot) = \phi, & \text{on } G. \end{cases} \tag{1.3}$$

To lighten the notation we work, as usual, in its time-reversed system, which is given by:

$$\begin{cases} u_t - \alpha\Delta u = 0, & \text{in } (0,T)\times G, \\ u = 0, & \text{on } (0,T)\times\partial G, \\ u(0,\cdot) = \phi, & \text{on } G. \end{cases} \tag{1.4}$$

We define the average of (1.4) as:

$$\tilde{u}(t,x;\phi) := \int_0^{+\infty} u(t,x;\alpha,\phi)\rho(\alpha)d\alpha.$$

The first property of $\tilde{u}$ that we prove is its analyticity in the time variable from $(0,+\infty)$ to $L^2(G)$. Next, using this together with a hierarchic decay in the time variable of the different frequencies, we obtain some unique continuation results for (1.4). Finally we determine when the averaged dynamics of (1.4) is null observable by combining the Fourier representation of the solutions of (1.4) and the monotonicity of the solutions of (1.1) with respect to the boundary conditions.

In order to illustrate the effect of averaging in the dynamics, let us study the dynamics of (1.4) when $G = \mathbb{R}^d$. As averaging and the Fourier transform commute, we work on the Fourier transform of the fundamental solution of the heat equation, which is given by:

$$\exp(-\alpha|\xi|^2 t).$$

Consequently, the Fourier transform of the average of the fundamental solutions is given by:

$$\int_0^{+\infty} \exp(-\alpha|\xi|^2 t)\rho(\alpha)d\alpha;$$

i.e. the Laplace transform of ρ evaluated in $|\xi|^2 t$. In particular, for $r \in (0,1)$ if $\rho(\alpha) \sim_{0^+} e^{-C\alpha^{-\frac{r}{1-r}}}$ we have that:

$$\int_0^{+\infty} \exp(-\alpha|\xi|^2 t)\rho(\alpha)d\alpha \sim \exp(-C|\xi|^{2r}t^r), \tag{1.5}$$

when $|\xi|^2 t \to +\infty$ as shown in (2.8). Thus, for those density functions the averaged dynamics in $\mathbb{R}^d$ has a fractional nature. As we are going to prove, for G bounded this is also true and we have the usual controllability and observability results of fractional dynamics (see, for example, [4–8]); that is, (1.5) implies that the averaged unique continuation is preserved, but (1.5) preserves the null averaged observability if and only if $r > 1/2$, being the threshold density functions those which satisfy:

$$\rho(\alpha) \sim_{0^+} e^{-C\alpha^{-1}}. \tag{1.6}$$

2. Quantification of the main results

In this section we introduce the precise definition of the previously introduced observability notions, quantify the main results and give some specific examples.

To start with, we recall the definitions of the introduced observability notions:

Definition 2.1. Let $G \subset \mathbb{R}^d$ be a domain and $G_0 \subset G$ be a subdomain. System (1.4) is *null averaged observable* or *null observable in average* in G_0 if for all $T > 0$ there is a constant $C > 0$ such that for any $\phi \in L^2(G)$:

$$\left\|\tilde{u}(T,\cdot;\phi)\right\|_{L^2(G)} \le C\|\tilde{u}(\cdot;\phi)\|_{L^2((0,T)\times G_0)}. \tag{2.1}$$

If (1.4) is null averaged observable, we define the *cost* of the null averaged observability as:

$$K(G,G_0,\rho,T) = \sup_{\phi\in L^2(G)\setminus\{0\}} \frac{\left\|\tilde{u}(T,\cdot;\phi)\right\|_{L^2(G)}}{\|\tilde{u}(\cdot;\phi)\|_{L^2((0,T)\times G_0)}}. \tag{2.2}$$

Definition 2.2. Let $G \subset \mathbb{R}^d$ be a domain and $G_0 \subset G$ be a subdomain. System (1.4) satisfies the *averaged unique continuation property* in G_0 if for all $T > 0$ the equality $\tilde{u} = 0$ in $(0,T)\times G_0$ implies that $\phi = 0$.

To continue with, we state the precise hypotheses on ρ. For that, we focus on the Laplace transform of ρ, which also appears naturally when G is a bounded domain (see (3.1)). We use the asymptotic notation $f(s) \gtrsim g(s)$, which means that there is $C > 0$ such that $f(s) \ge Cg(s)$ for s large enough.

- To have the unique continuation we need for some $r > 0$ that:

$$-\frac{d}{ds}\ln\left(\int_0^{+\infty} e^{-s\alpha}\rho(\alpha)d\alpha\right) = \frac{\int_0^{+\infty} e^{-s\alpha}\alpha\rho(\alpha)d\alpha}{\int_0^{+\infty} e^{-s\alpha}\rho(\alpha)d\alpha} \gtrsim s^{r-1}. \tag{2.3}$$

- To have the null observability we need (2.3) for some $r > \frac{1}{2}$.
- To prove the lack of null observability we need for some $C > 0$ and $r \in \left[0,\frac{1}{2}\right)$ that:

$$\int_0^{+\infty} e^{-s\alpha}\rho(\alpha)d\alpha \gtrsim e^{-Cs^r}. \tag{2.4}$$

Let us now state the main results of this paper:

- The first main result of this paper is that in many cases we have the unique continuation property:

Theorem 2.3. *Let $G \subset \mathbb{R}^d$ be a Lipschitz domain, $G_0 \subset G$ be a subdomain, and $\rho = 1_{(0,1)}$ or ρ be a density function which satisfies (2.3) for some $r > 0$. Then, system (1.4) satisfies the averaged unique continuation property in G_0.*

The proof of Theorem 2.3 is given in Section 4. For the uniform distribution it relies on explicit computations of the averaged solutions, whereas for the more diffusive case it relies on the analyticity of the averaged dynamics from $t \in (0, +\infty)$ to $L^2(G)$ (see Proposition 4.1) and on the fact that there is some hierarchy in how the frequencies decay, a technique dating back to [9].

- The second main result of this paper concerns some cases in which we do not have averaged observability:

Theorem 2.4. *Let $G \subset \mathbb{R}^d$ be a Lipschitz domain, $G_0 \subset G$ be a subdomain such that $G_0 \neq G$ and ρ be a density function which satisfies (2.4) for some $C > 0$ and $r \in \left[0, \frac{1}{2}\right)$. Then, system (1.4) is not null observable in average in G_0.*

We know from Theorem 2.3 that the lack of observability is not caused by a lack of unique continuation. In fact, we prove Theorem 2.4 in Section 5 by giving a sequence $\phi_N \in L^2(G)$ such that:

$$\lim_{N\to\infty} \frac{\|\tilde{u}(T, \cdot; \phi_N)\|_{L^2(G)}}{\left(\int_0^T \int_{G_0} |\tilde{u}(t, x; \phi_N)|^2 dxdt\right)^{1/2}} = +\infty. \tag{2.5}$$

This sequence is constructed with functions supported in $G \setminus G_0$, orthogonal to some low frequencies and, at the same time, not too concentrated on high frequencies. Estimate (2.4) ensures us that the mid frequencies do not decay too fast. The fact that the proof works for all $d \in \mathbb{N}$ and $r \in [0, 1/2)$ is a step forwards with respect to the literature, as in analogous situations with fractional dynamics the lack of controllability for $d \geq 2$ and $r \in [0, 1/2)$ is still unproved.

Remark 2.5. If $G = G_0$ system (1.4) has the averaged unique continuation property and is null observable in average. This is an immediate consequence of the fact that $t \mapsto \|\tilde{u}(t, \cdot; \phi)\|_{L^2(G)}$ is a decreasing function (see Remark 3.8).

- The last main result of the paper concerns some cases in which we have averaged observability:

Theorem 2.6. *Let $G \subset \mathbb{R}^d$ be a Lipschitz locally star-shaped domain, $G_0 \subset G$ be a subdomain, $T > 0$ and ρ be a density function which satisfies (2.3) for some $r > \frac{1}{2}$. Then, system (1.4) is null observable in average. In addition, there are $T_0, C > 0$ such that for all $T \in (0, T_0]$ we have that:*

$$K(G, G_0, \rho, T) \leq Ce^{CT^{-(2r-1)^{-1}}}. \tag{2.6}$$

We recall that the locally star-shaped domains are defined in [10, Section 3] and include all the C^2 domains. We prove Theorem 2.6 in Section 6 by adapting the ideas of [11]; that is, we use the Fourier representation and the decay properties of the averaged dynamics.

Remark 2.7. The estimate (2.6) is an upper estimate for short-time horizons. Ideally, it would also be good to have a lower bound and to precise the constant of the exponential by some geometric terms as in the heat equation (see, for instance, [12–16]), though this problem goes beyond the objective of this work.

Example 2.8. The density functions which satisfy the hypotheses of Theorems 2.3 and 2.6 include those which decay sufficiently fast when the diffusivity vanishes. Similarly, the density functions which satisfy the hypothesis of Theorem 2.4 are those which do not decay fast enough (including those which do not decay at all) when the diffusivity vanishes. Meaningful examples include:

1. Any density function ρ such that $\rho(\alpha) = 0$ for all $\alpha \in (0, \delta_\rho)$ for some $\delta_\rho > 0$ satisfies (2.3) for $r = 1$. Hence, all these functions satisfy the hypotheses of Theorems 2.3 and 2.6.
2. If $k \in (0, +\infty)$, and p and q are two polynomials such that $p(\alpha)+q(\alpha^{-1}) \neq 0$ for some $\alpha \in (0, 1)$, the density function:
$$\rho(\alpha) = \frac{\left(p(\alpha) + q(\alpha^{-1})\right) e^{-\alpha^{-k}} 1_{(0,1)}(\alpha)}{\int_0^1 \left(p(s) + q(s^{-1})\right) e^{-s^{-k}} ds}, \tag{2.7}$$
satisfies (2.3) for $r = \frac{k}{k+1}$. Thus, (2.7) satisfies the hypotheses of Theorem 2.3 if $k > 0$ and of Theorem 2.6 if $k > 1$.
3. The density functions given by (2.7) satisfy (2.4) for $r = \frac{k}{k+1}$. Thus, if $k \in (0, 1)$ they satisfy the hypothesis of Theorem 2.4.
4. The density functions $\rho(\alpha) = 1_{(0,1)}(\alpha)$ (that is, when α is a random variable with uniform distribution in $(0, 1)$) and $\rho(\alpha) = e^{-\alpha} 1_{(0,+\infty)}(\alpha)$ (that is, when α is a random variable with exponential distribution in $(0, +\infty)$) satisfy (2.4) for all $r > 0$. Indeed, any continuous density function ρ such that $\rho(0) > 0$ does so. Thus, all these functions satisfy the hypotheses of Theorem 2.4.

The proofs of items 1 and 4 are straightforward. As for items 2 and 3, we can prove them by considering the asymptotic result:

$$\int_0^1 \alpha^r e^{-s\alpha - \alpha^{-k}} d\alpha \sim s^{-\frac{2+2r+k}{2+2k}} e^{-c_k s^{\frac{k}{k+1}}}, \tag{2.8}$$

for all $r \in \mathbb{R}$ for some $c_k > 0$ (independent of r) when $s \to +\infty$. These asymptotic similarities can be proved with the Laplace method. In fact, we have that:

$$\int_0^1 \alpha^r e^{-s\alpha - \alpha^{-k}} d\alpha \sim \int_0^{+\infty} \alpha^r e^{-s\alpha - \alpha^{-k}} d\alpha$$
$$= s^{-\frac{1+r}{1+k}} \int_0^{+\infty} t^r e^{-s^{\frac{k}{k+1}}\left(t + t^{-k}\right)} dt,$$

where we have used the change of variables $\alpha = ts^{-\frac{1}{k+1}}$. Next, we can show the equivalence by using the classical Laplace method. In fact, if ϕ is any convex function in $(0, +\infty)$ with minimum at $\bar{t}$, and if f is a continuous function in a neighbourhood of $\bar{t}$ with subexponential growth it is well-known the limit:

$$\lim_{\theta\to\infty} \frac{\int_0^\infty f(t) e^{-\theta\phi(t)} dt}{f\left(\bar{t}\right) e^{-\theta\phi(\bar{t})}} \sqrt{\frac{\theta\phi''\left(\bar{t}\right)}{2\pi}} = 1,$$

which is proved by considering that the mass of the integral is concentrated on a neighbourhood of $\bar{t}$, by using a Taylor expansion of order 2 in the exponent and the continuity of f, then extending again the integral to $(0, +\infty)$ and finally explicitly computing the Gaussian function. For a more detailed proof of (2.8) one can consult, for instance, [17, (6.4.35) and Example 6.4.9].

The rest of the paper is organized as follows: in Section 3 we present some basic results, in Section 4 we prove Theorem 2.3, in Section 5 we prove Theorem 2.4, in Section 6 we prove Theorem 2.6, in Section 7 we resume the controllability problem, in Section 8 we comment some possible extensions, and in Appendix we prove a technical result.

3. Preliminaries

In this section we introduce some basic facts and notation that we use later on. In particular, we study the spectral properties of the Dirichlet Laplacian, the size of the solutions of the heat equation and the decay implied by (2.3).

3.1. Some results on the spectral decomposition of the Dirichlet Laplacian

As usual, e_i denotes (starting at $i=0$) the eigenfunctions of the Dirichlet Laplacian, λ_i their respective eigenvalues and $\Lambda_\lambda := \{i : \lambda_i \leq \lambda\}$. In addition, for any $\lambda > 0$, $\mathcal{P}_\lambda$ denotes the orthogonal projection of $L^2(G)$ into $\langle e_i \rangle_{i\in\Lambda_\lambda}$ and $\mathcal{P}_\lambda^\perp$ the orthogonal projection of $L^2(G)$ into $\langle e_i \rangle_{i\in\Lambda_\lambda}^\perp$.

To begin with, we recall that, as shown in [1], the Fourier representation of the averaged solution is:

$$\begin{aligned}\tilde{u}(t,x;\phi) &:= \int_0^{+\infty} u(t,x;\alpha,\phi)\rho(\alpha)d\alpha \\ &= \sum_{i\in\mathbb{N}} \left(\int_0^{+\infty} e^{-\alpha\lambda_i t}\rho(\alpha)d\alpha\right)\langle \phi, e_i\rangle_{L^2(G)} e_i(x). \end{aligned} \tag{3.1}$$

Next, we recall that the eigenvalues have a growth limited by Weyl's law:

Lemma 3.1 (*Weyl's Law*)**.** *Let $G \subset \mathbb{R}^d$ be a Lipschitz domain. We have:*

$$\lim_{\lambda\to\infty} \frac{|\Lambda_\lambda|}{\lambda^{d/2}} = \frac{\mathrm{Vol}(B(0,1))\,\mathrm{Vol}(G)}{(2\pi)^d}.$$

In particular, there is $C > 0$ such that for all $\lambda \geq \lambda_0$:

$$|\Lambda_\lambda| \leq C\lambda^{d/2}. \tag{3.2}$$

Weyl's law is proved for instance in [18].

Finally, we recall the following elliptic result proved in [10, Theorem 3]:

Lemma 3.2 (*[10]*)**.** *Let G be a locally star-shaped domain and $G_0 \subset G$ a subdomain. There exists a constant $C > 0$ such that for all $\lambda > 0$ and $\{c_i\} \subset \mathbb{R}$:*

$$\left(\sum_{i\in\Lambda_\lambda} |c_i|^2\right)^{1/2} \leq Ce^{C\sqrt{\lambda}} \left\| \sum_{i\in\Lambda_\lambda} c_i e_i \right\|_{L^2(G_0)}. \tag{3.3}$$

This result is a refinement of [19, Theorem 1.2], which was a refinement of the results proved in [20].

3.2. Some results on the heat equation

In this subsection we state some properties of the solutions of the heat equation. We first recall that their time derivative can be estimated by using the analyticity and contraction of the semigroup of the heat equation (see [21, Sections 2.5 and 5.6]) and Cauchy's integration formula:

Lemma 3.3. *Let G be a bounded domain. Then, there is $C > 0$ such that for all $k \in \mathbb{N}$, $s \in \mathbb{R}^+$ and $\phi \in L^2(G)$ we have that:*

$$\|\partial_s^k v(s,\cdot;\phi)\|_{L^2(G)} \leq \frac{C^k k!}{s^k}\|\phi\|_{L^2(G)}, \tag{3.4}$$

for v the solution of:

$$\begin{cases} v_t - \Delta v = 0, & \text{in } (0,T)\times G, \\ v = 0, & \text{on } (0,T)\times\partial G, \\ v(0,\cdot) = \phi, & \text{on } G. \end{cases} \tag{3.5}$$

It is interesting to consider the solutions of (3.5) because of the identity:

$$u(t,x;\alpha,\phi) = v(t\alpha, x;\phi), \tag{3.6}$$

for u the solution of (1.4).

Another result that we need is that the propagation of the mass, when the initial value in some subdomain is null, is exponentially slow:

Lemma 3.4. *Let G be a bounded domain and let $\hat{G}, G_0 \subset G$ be Lipschitz domains satisfying $\hat{G} \subset\subset G\setminus G_0$. Then, there are $c, C > 0$ such that for all ϕ satisfying $\mathrm{supp}(\phi) \subset \hat{G}$ and all $T, \alpha > 0$ we have that:*

$$\|u(\cdot;\alpha,\phi)\|_{C^0([0,T];L^2(G_0))} \leq Ce^{-\frac{c}{\alpha T}}\|\phi\|_{L^2(G)}, \tag{3.7}$$

for u the solution of (1.4).

We recall that $A \subset\subset B$ means that A is contained in a compact set contained in B. Lemma 3.4, whose originality we do not claim, is a consequence of the comparison principle. Indeed, following for example the ideas of [22, Lemma 4], we obtain Lemma 3.4 by comparing the solutions of (1.4) with initial value $\pm\phi$ to the solution of the heat equation in $\mathbb{R}^d$ with initial value $|\phi|1_G$, a solution which can be estimated by using its representation with the kernel of the heat equation.

3.3. Decay properties implied by (2.3)

In this subsection we show that if the density function ρ satisfies (2.3), the averaged solutions of (1.4) have a decay similar to that of the solutions of the fractional heat equation. In particular, we prove the following result:

Lemma 3.5. *Let ρ be a density function which satisfies (2.3) for some $r \in (1/2, 1]$. Then, there is $c > 0$ such that for all $\lambda \geq \lambda_0$ and $t_1, t_2 \in [0,1)$ satisfying $t_1 < t_2$ we have that:*

$$\int_0^{+\infty} e^{-t_2\lambda\alpha}\rho(\alpha)d\alpha \leq e^{-c\lambda^r(t_2-t_1)} \int_0^{+\infty} e^{-t_1\lambda\alpha}\rho(\alpha)d\alpha. \tag{3.8}$$

We recall that λ_0 is the first eigenvalue of the Laplacian.

Proof. In order to prove Lemma 3.5 we first remark that for all $s \geq 0$ we have that:

$$\begin{aligned} -\frac{d}{ds}\ln\left(\int_0^{+\infty} e^{-s\alpha}\rho(\alpha)d\alpha\right) &= \frac{\int_0^{+\infty} e^{-s\alpha}\alpha\rho(\alpha)d\alpha}{\int_0^{+\infty} e^{-s\alpha}\rho(\alpha)d\alpha} \\ &\geq c\min\{s^{r-1}, 1\}. \end{aligned} \tag{3.9}$$

Indeed, (3.9) follows from (2.3) and from the continuity of $s \mapsto \frac{\int_0^{+\infty} e^{-s\alpha}\alpha\rho(\alpha)d\alpha}{\int_0^{+\infty} e^{-s\alpha}\rho(\alpha)d\alpha}$ in $[0,+\infty)$. Thus, from (3.9) we obtain for all $s_1, s_2 \geq 0$ with $s_1 < s_2$ the estimate:

$$\begin{aligned} \int_0^{+\infty} e^{-s_2\alpha}\rho(\alpha)d\alpha &\leq \exp\left(-c\int_{s_1}^{s_2} \min\{s^{r-1},1\}ds\right) \\ &\quad\times \int_0^{+\infty} e^{-s_1\alpha}\rho(\alpha)d\alpha. \end{aligned} \tag{3.10}$$

Next, we fix t_1 and t_2 satisfying $t_1, t_2 \in [0,1)$ and $t_1 < t_2$, and use different approaches depending on the value of λ:

- If $\lambda \in [\lambda_0, t_2^{-1}]$, from (3.10) taking $s_1 = \lambda t_1$ and $s_2 = \lambda t_2$, since $[s_1, s_2] \subset (0,1]$, we obtain that:

$$\int_0^{+\infty} e^{-t_2\lambda\alpha}\rho(\alpha)d\alpha \leq e^{-c\lambda(t_2-t_1)} \int_0^{+\infty} e^{-t_1\lambda\alpha}\rho(\alpha)d\alpha. \tag{3.11}$$

In addition, since $\lambda \geq \lambda_0$ we have that:

$$-\lambda \leq -\lambda_0^{1-r}\lambda^r. \tag{3.12}$$

Thus, from (3.11) and (3.12) we obtain (3.8) for some $c > 0$ and all $\lambda \in [\lambda_0, t_2^{-1}]$.

- If $\lambda \in [t_1^{-1}, +\infty)$, from (3.10) taking $s_1 = \lambda t_1$ and $s_2 = \lambda t_2$, since $[s_1, s_2] \subset [1, +\infty)$, we obtain that:

$$\int_0^{+\infty} e^{-t_2\lambda\alpha}\rho(\alpha)d\alpha \leq e^{-c\lambda^r(t_2^r - t_1^r)}\int_0^{+\infty} e^{-t_1\lambda\alpha}\rho(\alpha)d\alpha. \tag{3.13}$$

Moreover, we consider that:

$$\begin{aligned} -(t_2^r - t_1^r) &= \frac{(-t_2^r + t_1^r)(t_2^{1-r} + t_1^{1-r})}{t_2^{1-r} + t_1^{1-r}} \\ &= \frac{t_1 - t_2 - t_2^r t_1^{1-r} + t_2^{1-r} t_1^r}{t_2^{1-r} + t_1^{1-r}} \leq \frac{t_1 - t_2}{t_2^{1-r} + t_1^{1-r}} \\ &\leq -\frac{t_2 - t_1}{2}. \end{aligned} \tag{3.14}$$

We have used in the first inequality of (3.14) that $t_1 < t_2$ and $r \in (1/2, 1]$, and in the second one that $t_1 - t_2 < 0$ and $t_1, t_2 \in (0, 1]$. Thus, from (3.13) and (3.14) we obtain (3.8) for some $c > 0$ and all $\lambda \in [t_1^{-1}, +\infty)$.

- If $\lambda \in (t_2^{-1}, t_1^{-1})$, we have that:

$$(t_1, t_2) = (t_1, \lambda^{-1}] \cup (\lambda^{-1}, t_2).$$

For the time interval (λ^{-1}, t_2) we may use the result of the second case for t_1 replaced by λ and obtain that:

$$\int_0^{+\infty} e^{-t_2\lambda\alpha}\rho(\alpha)d\alpha \leq e^{-c\lambda^r(t_2 - \lambda^{-1})}\int_0^{+\infty} e^{-\alpha}\rho(\alpha)d\alpha. \tag{3.15}$$

In addition, for the time interval $(t_1, \lambda^{-1}]$ we may use the result of the second case for t_2 replaced by λ and obtain that:

$$\int_0^{+\infty} e^{-\alpha}\rho(\alpha)d\alpha \leq e^{-c\lambda^r(\lambda^{-1} - t_1)}\int_0^{+\infty} e^{-t_1\lambda\alpha}\rho(\alpha)d\alpha. \tag{3.16}$$

Thus, from (3.15) and (3.16), by taking smaller constants c if necessary, we get (3.8) for all $\lambda \in (t_2^{-1}, t_1^{-1})$. □

Remark 3.6. The first and third cases in the proof of Lemma 3.5 might be empty depending on the values of λ_0, t_1 and t_2. However, since we use this result for t_1 and t_2 arbitrarily small, we need to prove those cases.

In a similar way, we can also prove the following result:

Lemma 3.7. *Let ρ be a density function which satisfies (2.3) for some $r > 0$. Then, there is $c > 0$ such that for all $\lambda, \tilde{\lambda}$ such that $\tilde{\lambda} > \lambda \geq \lambda_0$ and $t \in [1, +\infty]$ we have that:*

$$\int_0^{+\infty} e^{-t\tilde{\lambda}\alpha}\rho(\alpha)d\alpha \leq e^{-ct^r((\tilde{\lambda})^r - \lambda^r)}\int_0^{+\infty} e^{-t\lambda\alpha}\rho(\alpha)d\alpha. \tag{3.17}$$

Indeed, integrating both sides of (2.3) in $(t\lambda, t\tilde{\lambda})$ we find (3.17).

Finally, we underline the following result:

Remark 3.8. A consequence of (3.1) is that $t \mapsto \|\tilde{u}(t, \cdot; \phi)\|_{L^2(G)}$ is a decreasing function. Indeed, we have that:

$$\|\tilde{u}(t, \cdot; \phi)\|_{L^2(G)}^2 = \sum_{i\in\mathbb{N}} \left(\int_0^{+\infty} e^{-\alpha\lambda_i t}\rho(\alpha)d\alpha\right)^2 |\langle\phi, e_i\rangle|^2,$$

which is a series of decreasing functions.

4. Unique continuation property for averaged solutions

In this section we prove the unique continuation property for averaged solutions (Theorem 2.3). We first study the uniform distribution and then the density functions which satisfy (2.3).

4.1. Proof of Theorem 2.3 for the uniform distribution

Let us compute the averaged solutions of (1.4) when α has the uniform distribution in (0, 1). For that, as in [1, Section 3] and [2, Section 3], we present $\tilde{u}$ as the difference of two terms of known nature:

$$\begin{aligned} \tilde{u}(t, x; \phi) &= \sum_{i\in\mathbb{N}} \int_0^1 e^{-\lambda_i\alpha t}\langle\phi, e_i\rangle e_i(x)d\alpha \\ &= \frac{1}{t}\left(\sum_{i\in\mathbb{N}} \frac{1}{\lambda_i}\langle\phi, e_i\rangle e_i(x) - \sum_{i\in\mathbb{N}} \frac{e^{-\lambda_i t}}{\lambda_i}\langle\phi, e_i\rangle e_i(x)\right) \\ &= \frac{1}{t}\left(-\Delta^{-1}\phi + \sum_{i\in\mathbb{N}} e^{-\lambda_i t}\langle\Delta^{-1}\phi, e_i\rangle e_i(x)\right). \end{aligned} \tag{4.1}$$

Consequently, from $\int_0^T \int_{G_0} |\tilde{u}(t, x)|^2 = 0$ and (4.1) we find that:

$$-\Delta^{-1}\phi + \sum_{i\in\mathbb{N}} e^{-\lambda_i t}\langle\Delta^{-1}\phi, e_i\rangle e_i(x) = 0 \text{ in } (0, T) \times G_0,$$

which differentiating in time implies that:

$$\sum_{i\in\mathbb{N}} e^{-\lambda_i t}\langle\phi, e_i\rangle e_i(x) = 0 \text{ in } (0, T) \times G_0. \tag{4.2}$$

Hence, using the analyticity of the solutions of the heat equation we have that (4.2) implies that $\phi = 0$, and thus we have the averaged unique continuation property.

4.2. Proof of Theorem 2.3 for density functions which satisfy (2.3)

The proof consists on several steps. First, we show that assuming (2.3) the averaged dynamic are real-analytic and then use this to prove the unique continuation. To begin with, we prove the analyticity:

Proposition 4.1. *Let G be a Lipschitz domain, α any random variable and $\phi \in L^2(G)$. Then, the function:*

$$U : t \in (0, +\infty) \to \tilde{u}(t, \cdot; \phi) \in L^2(G),$$

is analytic.

Proof. In order to prove Proposition 4.1, we prove that $U \in C^\infty$ and that:

$$\forall a_1, a_2 \in (0, +\infty)\ \exists C > 0 : \sup_{t\in[a_1,a_2]} \|U^{(i)}(t)\|_{L^2(G)} \leq C^i i! \quad \forall i \in \mathbb{N}, \tag{4.3}$$

which is a characterization of analyticity in $\mathbb{R}^+$ (see, for instance, [23, Proposition 1.2.12]). Since:

$$U(t) = \int_0^{+\infty} v(\alpha t, \cdot; \phi)\rho(\alpha)d\alpha,$$

for v the solution of (3.5), we can easily see that:

$$U^{(i)}(t) = \int_0^{+\infty} \alpha^i \left(\partial_t^i v\right)(\alpha t, \cdot; \phi)\rho(\alpha)d\alpha, \tag{4.4}$$

and thus $U \in C^\infty$. Moreover, (4.3) follows from (4.4), the triangular inequality and (3.4). □

To continue with, we present the following auxiliary result:

Lemma 4.2. *Let G be a domain, $G_0 \subset G$ be a subdomain and v be an analytic function from $(0, +\infty)$ to $L^2(G)$. Then, if there is $T^* > 0$ such that $v = 0$ in $(0, T^*) \times G_0$, $v = 0$ in $(0, +\infty) \times G_0$.*

Proof. Lemma 4.2 follows from the fact that if v is analytic from $(0, +\infty)$ to $L^2(G)$ and $\psi \in L^\infty(G)$, then ψv is analytic from $(0, +\infty)$ to $L^2(G)$. Indeed, since v is analytic, by definition, for all $\tilde{t} \in (0, +\infty)$ there are $v_{i,\tilde{t}} \in L^2(G)$ such that $v = \sum_{i=0}^{\infty} v_{i,\tilde{t}}(t - \tilde{t})^i$ in a neighbourhood of $\tilde{t}$, so $\psi v = \sum_{i=0}^{\infty}(\psi v_{i,\tilde{t}})(t - \tilde{t})^i$ in that neighbourhood, and thus ψv is analytic. This implies that $v 1_{G_0}$ is analytic from $(0, +\infty)$ to $L^2(G)$. Consequently, since $v 1_{G_0} = 0$ in $(0, T^*)$, by analyticity $v 1_{G_0} = 0$ in $(0, +\infty) \times G$, so $v = 0$ in $(0, +\infty) \times G_0$. □

Now we are ready to prove Theorem 2.3:

End of the Proof of Theorem 2.3. Let $\phi \in L^2(G)$ such that $\tilde{u}(t, x; \phi) = 0$ in $(0, T) \times G_0$. By Proposition 4.1 and Lemma 4.2 we have that $\tilde{u}(t, x; \phi) = 0$ in $(0, +\infty) \times G_0$. Let us show that the first frequency of ϕ is null by contradiction. If the first frequency is not null, we obtain from (3.1) and (3.17) that:

$$\begin{aligned}
\tilde{u}(t, \cdot; \phi) &= \int_0^{+\infty} e^{-\alpha\lambda_0 t}\rho(\alpha)d\alpha \langle\phi, e_0\rangle_{L^2(G)} e_0 \\
&\quad + \sum_{i\in\mathbb{N}_*}\left(\int_0^{+\infty} e^{-\alpha\lambda_i t}\rho(\alpha)d\alpha\right)\langle\phi, e_i\rangle_{L^2(G)} e_i \\
&= \left(\int_0^{+\infty} e^{-\alpha\lambda_0 t}\rho(\alpha)d\alpha\right) \\
&\quad \times \left[\langle\phi, e_0\rangle_{L^2(G)} e_0 + O\left(e^{-(\lambda_1^r - \lambda_0^r)t^r}\|\phi\|_{L^2(G)}\right)\right]. \qquad (4.5)
\end{aligned}$$

Thus, by considering (4.5) for large values of t we obtain that $\langle\phi, e_0\rangle_{L^2(G)} e_0 = 0$ in G_0, which by Lemma 3.2 implies that $\langle\phi, e_0\rangle_{L^2(G)} e_0 = 0$, arriving at a contradiction.

To continue with, we can prove in a similar way that if ϕ is null up to the Nth frequency, then $\tilde{u}(t, x; \phi) = 0$ in $(0, T) \times G_0$ implies that the $(N + 1)$-th frequency is also null. Consequently, we obtain by induction that $\tilde{u}(t, x; \phi) = 0$ in $(0, T) \times G_0$ implies $\phi = 0$. □

5. Lack of null averaged observability

In this section we prove Theorem 2.4. As for the notation used in this section, C (resp. c) denotes a sufficiently large (resp. small) positive constant that may be different each time it appears and which just depends on G, G_0, T and ρ. In particular, it does not depend on the index N that we are going to introduce.

In order to prove Theorem 2.4 we construct a sequence ϕ_N satisfying (2.5). For that purpose, we first state and justify the properties which allow to have (2.5):

- The first property is that:

$$\bigcup_{N\in\mathbb{N}} \operatorname{supp}(\phi_N) \subset\subset G \setminus \overline{G_0}. \qquad (5.1)$$

  This requirement is very natural as $G \setminus \overline{G_0}$ is the part of the domain that cannot be observed when $\alpha = 0$. We use it in addition to Lemma 3.4 to obtain that $u(t, x; \alpha, \phi_N)$ decays exponentially in $\{(x, t, \alpha) : x \in G_0,\ \alpha t < N^{-1/2}\}$.
- The second property is that:

$$\phi_N \in \langle e_i\rangle^{\perp}_{i\in\Lambda_N}. \qquad (5.2)$$

  The benefit of (5.2) is that $u(t, x; \alpha, \phi_N)$ decays exponentially in $\{(x, t, \alpha) : x \in G,\ \alpha t \geq N^{-1/2}\}$, which follows from (3.1).
- The third property is that for $C > 0$ large enough we have that:

$$\|\mathcal{P}_{CN}\phi_N\|_{L^2(G)} \geq \sqrt{3}\|\phi_N\|_{L^2(G)}/2. \qquad (5.3)$$

  This estimate is needed to make sure that $\|\tilde{u}_N(T, \cdot; \phi_N)\|_{L^2(G)}/\|\phi_N\|_{L^2(G)}$ does not decay too fast.

Let us construct the sequence ϕ_N. For that, we inspire in [24, Section 6] and consider more or less a linear combination of Dirac masses; that is,

$$\phi_N \approx \sum_{i_1,\ldots,i_N=1}^{C\sqrt{N}} c_{i_1,\ldots,i_N,N}\delta^0_{x_{i_1,\ldots,i_N,N}}. \qquad (5.4)$$

In fact, the Dirac masses are replaced by $C_N^{N^d}\varsigma(C_N(x - x_{i_1,\ldots,i_N,N}))$, for ς a regularizing function and C_N to be defined. The property (5.1) is trivial. As for (5.2), we can obtain it by taking the right linear combination. Indeed, we just have to solve a homogeneous linear system which, for $C > 0$ large enough, by Weyl's law (see Lemma 3.1) has more unknowns than equations. Finally, we can obtain (5.3) by choosing the right approximation with functions whose support has a diameter proportional to $N^{-1/2}$. In particular, we can prove that:

Proposition 5.1. *Let $G \subset \mathbb{R}^d$ be a Lipschitz domain and $G_0 \subset G$. Then, there is a sequence $(\phi_N)_{N\geq N_0}$ satisfying (5.1), (5.2) and (5.3).*

The rigorous proof of Proposition 5.1 is a bit technical, so it is postponed to Appendix.

Remark 5.2. Since (5.1), (5.2) and (5.3) just depend on G and G_0, so does the sequence ϕ_N.

Example 5.3. In Fig. 1 we illustrate the solutions of the heat equation given by the proof of Proposition 5.1 to get an insight on how they look like. For doing these graphs we have taken $G = (0, \pi)$, $G_0 = (0, \pi/2)$, $\rho = 1_{(0,1)}$ and:

$$\varsigma(x) = \exp\left(\frac{-1}{10(x-1)^2(x+1)^2}\right)1_{(-1,1)}(x). \qquad (5.5)$$

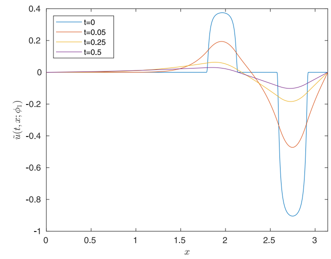
We recall that in $(0, \pi)$ we have that $e_i(x) = \sin(ix)$ and $\lambda_i = i^2$.

Let us now prove rigorously Theorem 2.4:

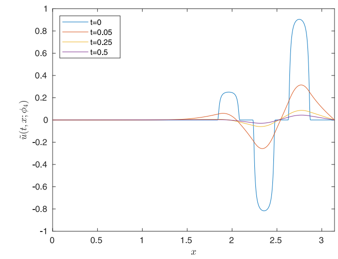
Proof of Theorem 2.4. We consider ϕ_N given by Proposition 5.1 (for N large enough). We easily find that:

$$\begin{aligned}
&\int_0^T\int_{G_0}\left|\int_0^{+\infty} u(t, x; \alpha, \phi_N)\rho(\alpha)d\alpha\right|^2 dxdt \\
&\leq \int_0^T\int_{G_0}\int_0^{+\infty} |u(t, x; \alpha, \phi_N)|^2\rho(\alpha)d\alpha dxdt \\
&= \int_0^T\int_{G_0}\int_0^{+\infty} |u(t, x; \alpha, \phi_N)|^2 1_{\alpha t\leq N^{-1/2}}(t, \alpha)\rho(\alpha)d\alpha dxdt \\
&\quad + \int_0^T\int_{G_0}\int_0^{+\infty} |u(t, x; \alpha, \phi_N)|^2 1_{\alpha t> N^{-1/2}}(t, \alpha)\rho(\alpha)d\alpha dxdt \\
&\leq C\left(e^{-c\sqrt{N}} + e^{-\sqrt{N}}\right)\|\phi_N\|^2_{L^2(G)}. \qquad (5.6)
\end{aligned}$$

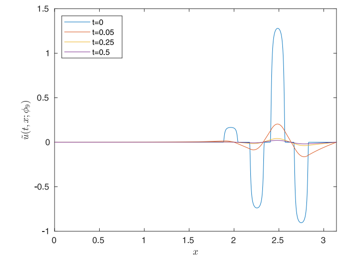
Indeed, for the first inequality of (5.6) we have used that the L^1 norm in a probabilistic space can be estimated by the L^2 norm. As for the second inequality of (5.6), we have used (3.7) and (5.1) for bounding the first integral, whereas we have used (5.2), (3.6) and the Fourier decomposition of the solutions of the heat equation for bounding the second one.



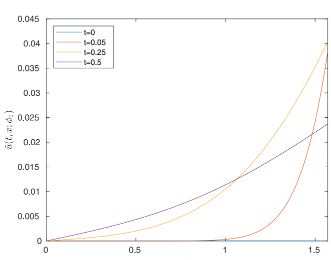
(a) The averaged solution with initial value ϕ_1



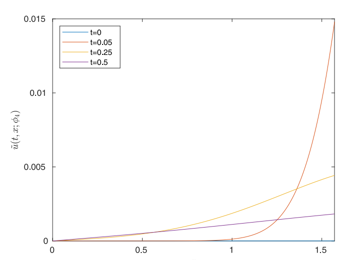
(b) The averaged solution with initial value ϕ_4



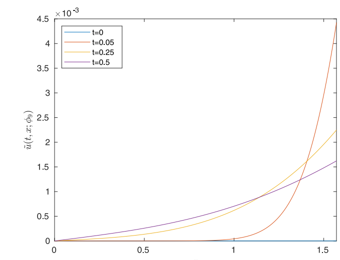
(c) The averaged solution with initial value ϕ_9



(d) The average solution with initial value ϕ_1 zoomed in $(0,\pi/2)$



(e) The average solution with initial value ϕ_4 zoomed in $(0,\pi/2)$



(f) The average solution with initial value ϕ_9 zoomed in $(0,\pi/2)$

Fig. 1. The averaged solutions of the heat equation when $G=(0,\pi)$, $G_0=(0,\pi/2)$, $\rho=1_{(0,1)}$ with initial values of the sequence given in the proof of Proposition (see (A.1)) and with ς given by (5.5).

To continue with, using (3.1), (2.4) and (5.3) we obtain that:

$$\begin{aligned}\|\tilde{u}(T,\cdot;\phi_N)\|^2_{L^2(G)} &= \sum_{i\in\mathbb{N}}\left(\int_0^\infty e^{-\lambda_i\alpha T}\rho(\alpha)d\alpha\right)^2|\langle\phi_N,e_i\rangle|^2\\ &\geq c\sum_{i\in\mathbb{N}}e^{-C(\lambda_i T)^r}|\langle\phi_N,e_i\rangle|^2\\ &\geq ce^{-CN^r}\sum_{i\in\Lambda_{CN}}|\langle\phi_N,e_i\rangle|^2 = ce^{-CN^r}\|\mathcal{P}_{CN}\phi_N\|^2_{L^2(G)}\\ &\geq ce^{-CN^r}\|\phi_N\|^2_{L^2(G)}. \end{aligned}\tag{5.7}$$

Hence, recalling that $r\in[0,1/2)$ we easily obtain (2.5) from (5.6) and (5.7). $\square$

6. Proof of null averaged observability

In this section we prove Theorem 2.6. As for the notation, C (resp. c) denotes a sufficiently large (resp. small) positive constant that may be different each time it appears and which just depends on G, G_0, ρ and r, but which is independent of $T\in(0,T_0)$, for $T_0(G,G_0,\rho)$ small enough.

In order to prove Theorem 2.6 we use the approach introduced in [11, Section 2]. It is not a direct consequence of the results presented in that section because the dynamics of the averaged solution only satisfies a decay property and not a semigroup property.

First, we reformulate [11, Lemma 2.1]:

Lemma 6.1. *Let $G\subset\mathbb{R}^d$ be a domain, G_0 be a subdomain, $T_0>0$, $q\in(0,1)$ and f be a positive function such that $f(t)\to 0$ as $t\to 0^+$. Suppose that we have for all $\phi\in L^2(G)$ and all $t_1,t_2\in(0,T_0]$ satisfying $t_1<t_2$ that:*

$$f(t_2-t_1)\|\tilde{u}(t_2,\cdot;\phi)\|^2_{L^2(G)}-f(q(t_2-t_1))\|\tilde{u}(t_1,\cdot;\phi)\|^2_{L^2(G)} \leq \int_{t_1}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\phi)|^2dxd\tau. \tag{6.1}$$

Then, we have for all $\phi\in L^2(G)$ and $T\in(0,T_0]$ that:

$$\left\|\tilde{u}(T,\cdot;\phi)\right\|_{L^2(G)}\leq\sqrt{f((1-q)T)}\|\tilde{u}(\cdot;\phi)\|_{L^2((0,T)\times G_0)}.$$

The proof of Lemma 6.1 is very similar to that of [11, Lemma 2.1]: a telescopic sum considering $t_{2,i}=Tq^i$ and $t_{1,i}=Tq^{i+1}$ for $i\in\mathbb{N}$.

As in [11], we do not prove (6.1) directly, but we prove a similar version, which is the analogue of [11, Lemma 2.3]:

Lemma 6.2. *Let $G\subset\mathbb{R}^d$ be a domain, G_0 be a subdomain, $T_0,\beta,\gamma_1,\gamma_2,f_0,g_0>0$ satisfying $\gamma_1<\gamma_2$. Suppose that we have for all $\phi\in L^2(G)$ and all $t_1,t_2\in(0,T_0]$ satisfying $t_1<t_2$ the inequality:*

$$f(t_2-t_1)\|\tilde{u}(t_2,\cdot;\phi)\|^2_{L^2(G)}-g(t_2-t_1)\|\tilde{u}(t_1,\cdot;\phi)\|^2_{L^2(G)} \leq\int_{t_1}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\phi)|^2dxd\tau, \tag{6.2}$$

for $f(s)\geq f_0\exp(-2/(\gamma_2 s)^\beta)$ and $g(s)\leq g_0\exp(-2/(\gamma_1 s)^\beta)$. Then, for any $\gamma\in(0,\gamma_2-\gamma_1)$ there is $T'\in(0,T_0]$ such that for all $T\in(0,T']$ and $\phi\in L^2(G)$:

$$\left\|\tilde{u}(T,\cdot;\phi)\right\|_{L^2(G)}\leq\sqrt{f_0^{-1}\exp(1/(\gamma T)^\beta)}\|\tilde{u}(\cdot;\phi)\|_{L^2((0,T)\times G_0)}.$$

Moreover, if $g_0<f_0$, we can take $\gamma=\gamma_2-\gamma_1$ and $T'=T_0$.

The proof of Lemma 6.2 is the same as [11, Lemma 2.3]: bounding superiorly $\frac{g(s)}{f(qs)}$ and using Lemma 6.1.

Now we are ready to prove Theorem 2.6. We do it by following the strategy of [11, Section 2]:

Proof of Theorem 2.6. Let $t_1,t_2\in[0,1)$ such that $t_1<t_2$ and $\phi\in L^2(G)$. We are going to prove:

$$ce^{-C((t_2-t_1)^{-1/4}+\sqrt{\lambda})}\|\tilde{u}(t_2,\cdot;\phi)\|^2_{L^2(G)}-Ce^{-c\lambda^r(t_2-t_1)}\|\tilde{u}(t_1,\cdot;\phi)\|^2_{L^2(G)} \leq\int_{t_1}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\phi)|^2dxd\tau, \tag{6.3}$$

for all $\lambda \geq \lambda_0$ and then use Lemma 6.2 with the appropriate value of λ (depending on t_1 and t_2). First, considering Remark 3.8 and that $\mathcal{P}_\lambda\phi \perp \mathcal{P}_\lambda^\perp\phi$ we have that:

$$\|\tilde{u}(t_2,\cdot;\phi)\|^2_{L^2(G)} \leq \frac{2}{t_2-t_1}\int_{(t_1+t_2)/2}^{t_2}\int_G (|\tilde{u}(\tau,x;\mathcal{P}_\lambda\phi)|^2 + |\tilde{u}(\tau,x;\mathcal{P}_\lambda^\perp\phi)|^2)dxd\tau. \tag{6.4}$$

From Lemma 3.2 and that $\mathcal{P}_\lambda\phi = \phi - \mathcal{P}_\lambda^\perp\phi$ we obtain that:

$$\begin{aligned}
&\frac{2}{t_2-t_1}\int_{(t_1+t_2)/2}^{t_2}\int_G |\tilde{u}(\tau,x;\mathcal{P}_\lambda\phi)|^2 dxd\tau \\
&\leq C\frac{e^{C\sqrt{\lambda}}}{t_2-t_1}\int_{(t_1+t_2)/2}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\mathcal{P}_\lambda\phi)|^2dxd\tau \\
&\leq \frac{Ce^{C\sqrt{\lambda}}}{t_2-t_1}\int_{(t_1+t_2)/2}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\phi)|^2dxd\tau \\
&\quad + \frac{Ce^{C\sqrt{\lambda}}}{t_2-t_1}\int_{(t_1+t_2)/2}^{t_2}\int_G|\tilde{u}(\tau,x;\mathcal{P}_\lambda^\perp\phi)|^2.
\end{aligned} \tag{6.5}$$

Moreover, from the decay property of Remark 3.8 and (3.8) we have that:

$$\begin{aligned}
&\frac{Ce^{C\sqrt{\lambda}}}{t_2-t_1}\int_{(t_1+t_2)/2}^{t_2}\int_G|\tilde{u}(\tau,x;\mathcal{P}_\lambda^\perp\phi)|^2dxd\tau \\
&\leq Ce^{C\sqrt{\lambda}}\left\|\tilde{u}\left((t_2+t_1)/2,\cdot;\mathcal{P}_\lambda^\perp\phi\right)\right\|^2_{L^2(G)} \\
&\leq Ce^{C\sqrt{\lambda}-c\lambda^r(t_2-t_1)}\|\tilde{u}(t_1,\cdot;\mathcal{P}_\lambda^\perp\phi)\|^2_{L^2(G)}.
\end{aligned} \tag{6.6}$$

Thus, from (6.4)–(6.6), $2 \leq Ce^{C\sqrt{\lambda}}$ and $(t_2-t_1)^{-1} \leq Ce^{C(t_2-t_1)^{-(2r-1)^{-1}}}$ (recall that $C > 0$ is a sufficiently large constant) we obtain:

$$\begin{aligned}
\|\tilde{u}(t_2,\cdot;\phi)\|^2_{L^2(G)} &\leq \frac{Ce^{C\sqrt{\lambda}}}{t_2-t_1}\int_{(t_1+t_2)/2}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\phi)|^2dxd\tau \\
&\quad + Ce^{C\sqrt{\lambda}-c\lambda^r(t_2-t_1)}\|\tilde{u}(t_1,\cdot;\mathcal{P}_\lambda^\perp\phi)\|^2_{L^2(G)} \\
&\leq Ce^{C((t_2-t_1)^{-(2r-1)^{-1}}+\sqrt{\lambda})} \\
&\quad \times \int_{(t_1+t_2)/2}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\phi)|^2dxd\tau \\
&\quad + Ce^{C\sqrt{\lambda}-c\lambda^r(t_2-t_1)}\|\tilde{u}(t_1,\cdot;\mathcal{P}_\lambda^\perp\phi)\|^2_{L^2(G)}.
\end{aligned}$$

This inequality implies that, for a sufficiently small $c > 0$, since $e^{-C(t_2-t_1)^{-(2r-1)^{-1}}} < 1$:

$$\begin{aligned}
&ce^{-C((t_2-t_1)^{-(2r-1)^{-1}}+\sqrt{\lambda})}\|\tilde{u}(t_2,\cdot;\phi)\|^2_{L^2(G)} \\
&\leq \int_{t_1}^{t_2}\int_{G_0}|\tilde{u}(\tau,x;\phi)|^2dxd\tau + Ce^{-c\lambda^r(t_2-t_1)}\|\tilde{u}(t_1,\cdot;\mathcal{P}_\lambda^\perp\phi)\|^2_{L^2(G)},
\end{aligned} \tag{6.7}$$

which implies (6.3).

We now define:

$$\lambda(t_2,t_1) = \mathfrak{C}(t_2-t_1)^{-(r-1/2)^{-1}}, \tag{6.8}$$

for $\mathfrak{C} \geq \lambda_0$ a positive constant sufficiently large. If we take in (6.3) λ given by (6.8), we obtain (6.2) for the functions:

$$f(s) = c\exp\left(-C\left(s^{-(2r-1)^{-1}} + \mathfrak{C}^{1/2}s^{-(2r-1)^{-1}}\right)\right),$$

$$g(s) = C\exp\left(-c\mathfrak{C}^r s^{-(2r-1)^{-1}}\right).$$

Indeed, we have for all $s \in (0,1)$ that:

$$f(s) \geq c\exp\left(-C\mathfrak{C}^{1/2}s^{-(2r-1)^{-1}}\right).$$

Since $r > \frac{1}{2}$ the functions f and g satisfy the hypothesis of Lemma 6.2 for $\beta = (2r-1)^{-1}$, $\gamma_1 = (c\mathfrak{C}^r)^{-1/\beta}$ and $\gamma_2 = (C\mathfrak{C}^{1/2})^{-1/\beta}$ by taking $\mathfrak{C}$ large enough, so we end the proof by using Lemma 6.2. $\square$

7. The controllability problem

In this section we first resume the theoretical study of the controllability problem and then perform some simulations.

7.1. A theoretical study

As stated in the introduction, the observability results that we have obtained in this paper have some implications on the controllability of (1.1). Let us consider the controllability problem given by:

$$\begin{cases} y_t - \alpha\Delta y = f1_{G_0}, & \text{in } (0,T)\times G, \\ y = 0, & \text{on } (0,T)\times\partial G, \\ y(0,\cdot) = y^0, & \text{on } G. \end{cases} \tag{7.1}$$

In particular, we focus on the following notions of controllability, which are introduced in [25]:

Definition 7.1. System (7.1) is *null averaged controllable* or *null controllable in average* if for all $T > 0$ there is $C > 0$ such that for any initial value $y^0 \in L^2(G)$ there is $f \in L^2((0,T)\times G_0)$ satisfying:

$$\|f\|_{L^2((0,T)\times G_0)} \leq C\|y^0\|_{L^2(G)},$$

and $\tilde{y}(T,\cdot;y^0,f) = 0$. If (7.1) is null averaged controllable, the *cost* of the null averaged controllability is defined by:

$$\tilde{K}(G,G_0,\rho,T) = \sup_{y^0\in L^2(G)\setminus\{0\}}\ \inf_{f:\tilde{y}(T,\cdot;y^0,f)=0}\frac{\|f\|_{L^2((0,T)\times G_0)}}{\|y^0\|_{L^2(G)}}. \tag{7.2}$$

Definition 7.2. System (7.1) is *approximately averaged controllable* or *approximately controllable in average* if for all $T > 0$, $\varepsilon > 0$ and $y^0, y^1 \in L^2(G)$, there exists a control f^ε such that:

$$\left\|\tilde{y}(T,\cdot;y^0,f^\varepsilon) - y^1\right\|_{L^2(G)} < \varepsilon.$$

We now recall the duality result between observability and controllability:

Theorem 7.3 ([2]). *Let $G \subset \mathbb{R}^d$ be a domain and $G_0 \subset G$ be a subdomain. System (7.1) is null controllable in average if and only if system (1.4) is null observable in average in G_0. In that case, $K = \tilde{K}$ (see (2.2)* and *(7.2)); that is, the cost of the control of null averaged observability equals the cost of null averaged controllability. Similarly, system (7.1) is approximately averaged controllable if and only if system (1.4) satisfies the unique continuation property in G_0.*

The proof of Theorem 7.3 can be found in [2, Appendix A]. As an immediate consequence we obtain that Theorems 2.3, 2.4 and 2.6 and Remark 2.5 imply the following controllability results for system (7.1):

Corollary 7.4. *Let $G \subset \mathbb{R}^d$ be a domain, $G_0 \subset G$ be a subdomain and $T > 0$. Then:*

- *Under the hypotheses of Theorem 2.3, system (7.1) is approximately controllable in average.*
- *Under the hypotheses of Theorem 2.4, system (7.1) is not null controllable in average.*
- *If $G_0 = G$, system (7.1) is null and approximately controllable in average for any probability distribution ρ.*

- *Under the hypotheses of Theorem 2.6, system (7.1) is null controllable in average and there are $C, T_0 > 0$ such that for all $T \in (0, T_0]$ we have the bound:*

$$\tilde{K}(G, G_0, \rho, T) \leq Ce^{CT^{(2r-1)^{-1}}}.$$

As shown in [2, Appendix A.2], the controls can be obtained (when the system is controllable) by minimizing a quadratic functional, so they can be obtained by running simulations.

7.2. Simulations

In this section we illustrate experimentally the controllability results obtained in Corollary 7.4. Our objective is not to develop rigorous numerical methods for obtaining the controls, which goes beyond the objective of this work, but to get an insight on the differences between density functions inside and outside the null-controllability regime. For that, we recall that the optimal control for null controllability in average (when (7.1) is null controllable) is given by $\varphi(t, x; \phi)1_{G_0}$, for φ the solution of (1.3) and ϕ the minimizer of:

$$J(\phi) = \frac{1}{2}\int_0^T \int_{G_0} \left| \int_0^{+\infty} \varphi(t, x; \alpha, \phi)\rho(\alpha)d\alpha \right|^2 dxdt + \left\langle y^0, \int_0^{+\infty} \varphi(0; \alpha, \phi)\rho(\alpha)d\alpha \right\rangle.$$

As the numerical simulations are cumbersome in higher dimensions, to get better illustrations we work in $d = 1$, and in particular in $G = (0, \pi)$. We also consider $G_0 = (1, 2)$, $T = 1$ and $y^0 = \frac{1}{2}$. Moreover, to illustrate these differences, we consider $\rho = 1_{(1,2)}$, which is inside, and $\rho = 1_{(0,1)}$, which is outside.

In order to numerically implement this problem, we approximate it by minimizing J in $V_M := \langle e_i \rangle_{i=1}^M$ for $M = 20$, $M = 50$ and $M = 100$. This is motivated by the results presented in [26] for the random wave equation. In fact, we may prove as in [26, Section 5], by using variational and duality arguments, that if the system is null controllable in average the controls constructed with the minimizer of J in V_M converge to the optimal control in $L^2((0, T) \times G_0)$. We say that a control is constructed with the minimizer of J in V_M if it is the restriction to $(0, T) \times G_0$ of the solution of (1.3) with φ^T replaced by the minimizer. Since V_M is a finite dimensional space, computing the minimizer of J is equivalent to solving numerically a linear system, which can be easily done by using any numerical computing environment (in our case MATLAB). We have obtained the following illustrations:

- We illustrate in Fig. 2 (resp. in Fig. 3) the controls constructed with the minimizer of J restricted to the spaces V_M for $\rho = 1_{(1,2)}$ (resp. for $\rho = 1_{(0,1)}$). For $\rho = 1_{(1,2)}$ we obtain a sequence of functions which seems to converge with respect to M to some function, which is something that can be seen in an even more clear way when $t \in [0, 1/2]$. Of course, the closer the time is to 1, the more slowly the punctual values of the control converge with M (and in $t = 1$ it diverges), but this is a well-known behaviour when controlling a parabolic dynamics (see, for instance, [27–29]). However, for $\rho = 1_{(0,1)}$ the sequence of functions does not seem to converge at all, which is something that we can appreciate in a more detailed way when $t \in [0, 1/2]$. What explains the difference is the controllability properties of each density.
- We illustrate in Fig. 4 the state at $t = 1$ of the respective solutions of the averaged heat equation with the previously obtained functions. For $\rho = 1_{(1,2)}$ we see that the state is taken slowly to 0, whereas for $\rho = 1_{(0,1)}$ the picture shows some oscillations. This matches the theoretical results obtained in Corollary 7.4 as for $\rho = 1_{(1,2)}$ the system is known to be averaged null controllable, whereas for $\rho = 1_{(0,1)}$ the system is known to be not null averaged controllable. The computations are obtained by approximating the density functions with Dirac masses, which explains why the solution in the left figure does not converge exactly to 0.

8. Further comments and open problems

In this section we underline some extensions of our results to analogous situations and comment some interesting open problems:

- **Average of the controls.** A naive and incorrect way of computing a control that takes the average to rest is to compute the average of the controls that take each of the instances to rest. However, we do not get the same trajectory if we consider some source terms f_α and then average on the solutions or if we compute the solutions with source term $\int_0^\infty f_\alpha \rho(\alpha)d\alpha$ and then average (when $\alpha \mapsto f_\alpha$ is measurable). In fact, the solutions of the equation:

  $$\dot{u} + \alpha u = f_\alpha,$$

  are given by:

  $$u(t) = u(0)e^{-\alpha t} + \int_0^t f_\alpha(s)e^{-\alpha(t-s)}ds;$$

  thus, most often we do not get the same trajectory, as:

  $$\int_0^t \int_0^{+\infty} f_\alpha(s)e^{-\alpha(t-s)}\rho(\alpha)d\alpha ds \neq \int_0^t \left(\int_0^{+\infty} f_\alpha(s)\rho(\alpha)d\alpha\right) \times \left(\int_0^{+\infty} e^{-\alpha(t-s)}\rho(\alpha)d\alpha\right) ds.$$

  A specific counter-example can be given with scalar ODEs. Let us consider a dynamic that behaves half of the times like:

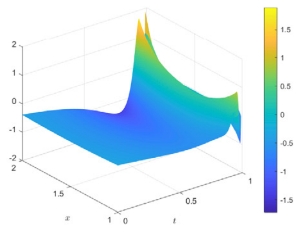
  $$\dot{u} + u = f,$$

  and half of the times like:

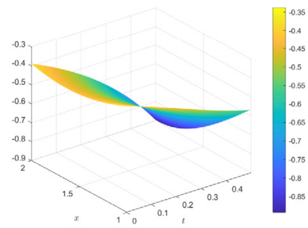
  $$\dot{v} - v = g;$$

  that is, $\mathbb{P}[\alpha = 1] = \mathbb{P}[\alpha = -1] = 1/2$. If the initial value is 1 we can take the solutions to rest at time $T = 1$ with the controls $f(t) = -t$ and $g(t) = t - 2$ (this is done by considering that $1 - t$ is a valid trajectory). The average of the controls is -1. However, for $f(t) = g(t) = \frac{-t+t-2}{2} = -1$, we obtain the trajectories $u(t) = -1 + 2e^{-t}$ and $v(t) = 1$, whose average is e^{-t}, which is not null at $T = 1$.
- **Random initial data.** As in some previous works involving average controllability results (see [2,3,25,30]) we fix the initial value. However, we can prove by linearity that averaged approximate controllability with fixed initial data implies averaged approximate controllability with random initial data y^0_α if $\int_0^{+\infty} \|y^0_\alpha\|^2_{L^2(G)}\rho(\alpha)d\alpha < \infty$. Nonetheless, in the case of null controllability in average to argue with linearity we would need to prove that the range of the average with random initial data is the same as the range with fixed initial value, which is an open problem. Alternatively, applying duality as in Theorem 7.3 could be a possibility. Indeed, we have to prove that

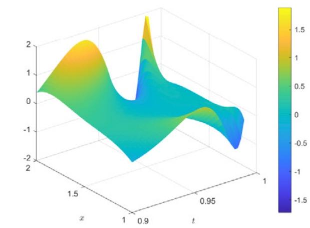
  $$\varphi^T \mapsto \int_0^{+\infty} \langle y^0_\alpha, \varphi(0, \cdot; \alpha)\rangle_{L^2(G)}\rho(\alpha)d\alpha$$



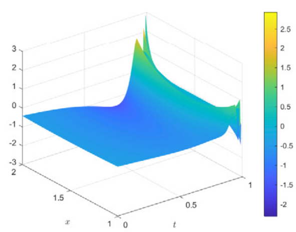
(a) The control constructed with the minimizer of J in V_{20} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0, 1]$.



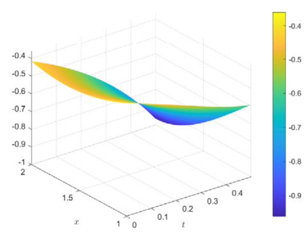
(b) The control constructed with the minimizer of J in V_{20} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0, 0.5]$.



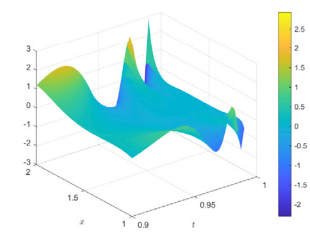
(c) The control constructed with the minimizer of J in V_{20} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0.9, 1]$.



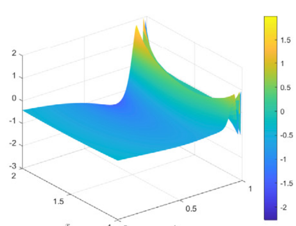
(d) The control constructed with the minimizer of J in V_{50} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0, 1]$.



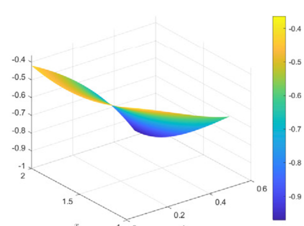
(e) The control constructed with the minimizer of J in V_{50} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0, 0.5]$.



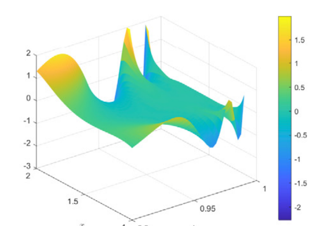
(f) The control constructed with the minimizer of J in V_{50} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0.9, 1]$.



(g) The control constructed with the minimizer of J in V_{100} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0, 1]$.



(h) The control constructed with the minimizer of J in V_{100} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0, 0.5]$.



(i) The control constructed with the minimizer of J in V_{100} for $y^0 = \frac{1}{2}$, $\rho = 1_{(1,2)}$ and $t \in [0.9, 1]$.

Fig. 2. Graphs of the controls for $\rho = 1_{(1,2)}$ and $y^0 = \frac{1}{2}$ constructed with the minimizer of the functional J in V_{20}, V_{50} and V_{100}. In the left column we illustrate the whole graphs, in the middle column we illustrate the graphs with the time variable zoomed in $[0, 1/2]$, and in the left column zoomed in $[0.9, 1]$.

is continuous with respect to the norm:

$$\varphi^T \mapsto \sqrt{\iint_{(0,T)\times G_0} \left(\int_0^{+\infty} \varphi(t, x; \alpha)\rho(\alpha) d\alpha \right)^2 dxdt}.$$

If the initial value is independent of α, then

$$\int_0^{+\infty} \langle y_0^\alpha, \varphi(0, \cdot; \alpha)\rangle_{L^2(G)} \rho(\alpha) d\alpha = \left\langle y_0, \int_0^{+\infty} \varphi(0, \cdot; \alpha)\rho(\alpha) d\alpha \right\rangle_{L^2(G)},$$

so the equivalent observability inequality is

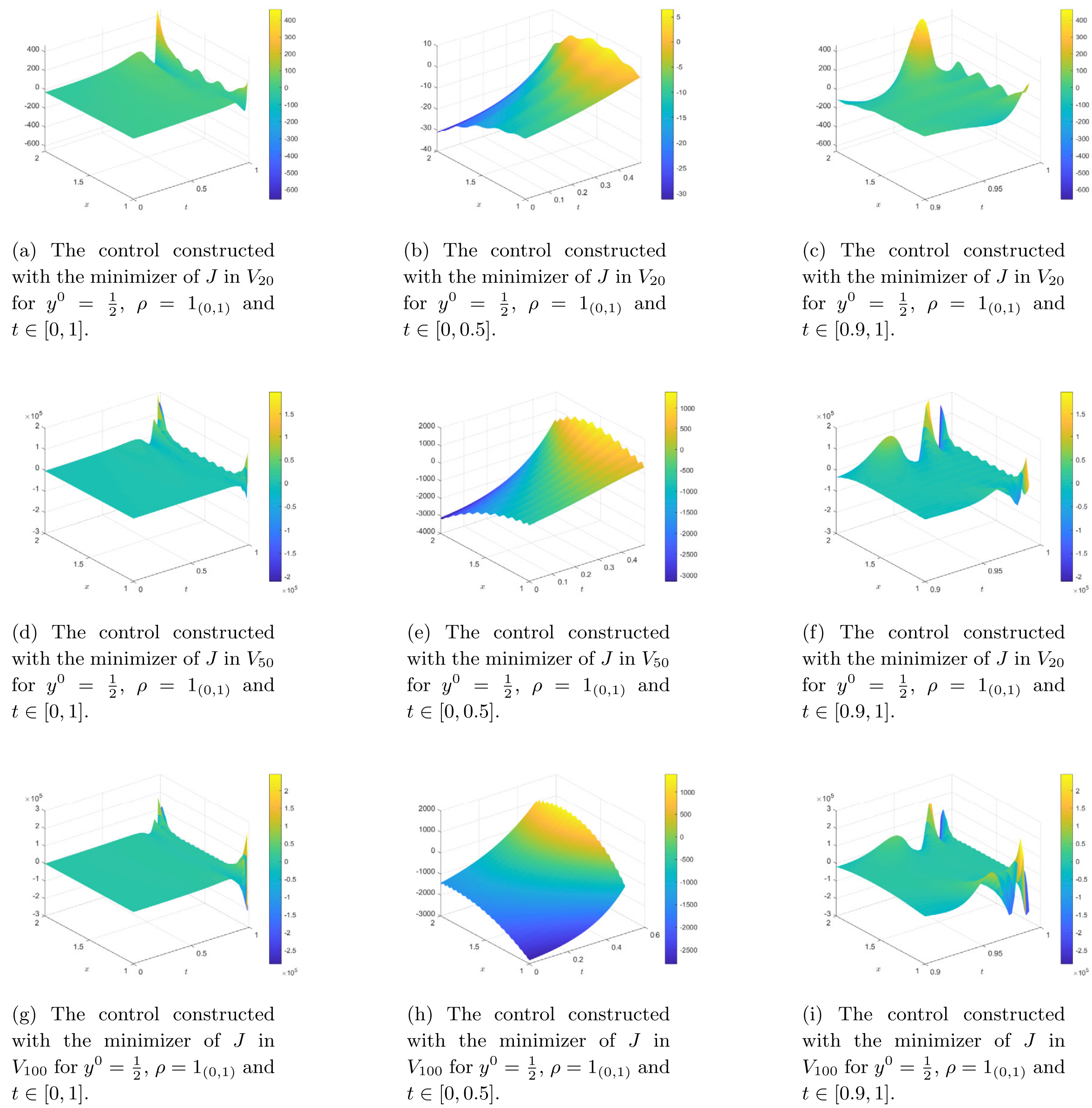
$$\|\tilde{\varphi}(0, \cdot)\|_{L^2(G)} \leq C \|\tilde{\varphi}\|^2_{L^2((0,T)\times G_0)},$$

but if the initial value depends on α, we need to prove that:

$$\int_0^{+\infty} \|\varphi(0, x; \alpha, \varphi^T)\|^2_{L^2(G)} \rho(\alpha) d\alpha \leq C \|\tilde{\varphi}\|^2_{L^2((0,T)\times G_0)}. \tag{8.1}$$

The main obstacle to prove (8.1) by replicating the proof of Theorem 2.6 is to prove an analogous result for (6.5). In recent papers where random initial values are considered the difficulty of a random initial value is bypassed by using exact averaged controllability (see, for instance, [31]), which is satisfied by finite dimensional or hyperbolic systems but not by parabolic ones, or by assuming that there is no randomness in the dynamics, just on the initial value (see, for instance, [32]). We highlight that understanding the controllability properties of (7.1) when y_α^0 is a random initial value is an open problem whose resolution would help to have a more complete picture.

- **Acting on the boundary.** We may prove analogous controllability results to Theorems 2.3, 2.4 and 2.6 when the control acts on the boundary. This can be done by repeating the proofs almost step by step, with the only difference of using [10, Theorem 9] instead of Lemma 3.2. We also remark that we cannot use an extension reduction technique as in [33, Section 3.3] since the trace on the boundary would depend on the parameter, and by the same argument of the first remark in Section 8, averaging the traces does not suffice.



(a) The control constructed with the minimizer of J in V_{20} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0, 1]$.

(b) The control constructed with the minimizer of J in V_{20} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0, 0.5]$.

(c) The control constructed with the minimizer of J in V_{20} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0.9, 1]$.

(d) The control constructed with the minimizer of J in V_{50} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0, 1]$.

(e) The control constructed with the minimizer of J in V_{50} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0, 0.5]$.

(f) The control constructed with the minimizer of J in V_{20} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0.9, 1]$.

(g) The control constructed with the minimizer of J in V_{100} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0, 1]$.

(h) The control constructed with the minimizer of J in V_{100} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0, 0.5]$.

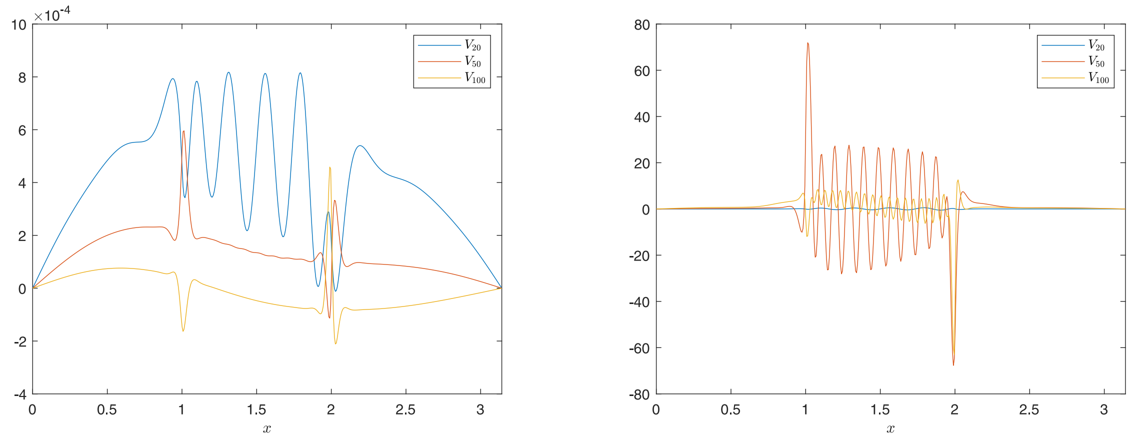
(i) The control constructed with the minimizer of J in V_{100} for $y^0 = \frac{1}{2}$, $\rho = 1_{(0,1)}$ and $t \in [0.9, 1]$.

Fig. 3. Graphs of the controls for $\rho = 1_{(0,1)}$ and $y^0 = \frac{1}{2}$ constructed with the minimizer of the functional J in V_{20}, V_{50} and V_{100}. In the left column we illustrate the whole graphs, whereas in the right column we illustrate the graphs with the time variable zoomed in $[0, 1/2]$, and in the left column zoomed in $[0.9, 1]$.

- **Neumann boundary conditions.** We have analogous results of Theorems 2.3 and 2.6 for the controllability of the averaged solutions of the heat equation with random diffusion and Neumann boundary conditions. Indeed, we can repeat the proof step by step of those theorems since (3.3) is also true for Neumann boundary conditions (see [19, Theorem 2]). However, whether the analogous of Theorem 2.4 is true remains an open question since we do not have an analogous result of Lemma 3.4 for Neumann boundary conditions.
- **More general random variables.** Even if all the results in this paper have been stated for random variables with a density function, they are true for any random variable whose law satisfies the analogous inequalities of (2.3) and (2.4). Indeed, the proofs can be replicated step by step.
- **More regular norms.** Even if we have obtained all the results in this paper for the final state in $L^2(G)$ and we have made the observation in $L^2((0,T) \times G_0)$, analogous results are valid for final states in $H^{s_1}(G)$ and the observation in $H^{s_2}((0,T) \times G_0)$ (for any $s_1, s_2 \in \mathbb{R}^+$) for a domain G sufficiently regular. Indeed, the proofs are very similar with the only difference of some polynomial factors of N or λ. We recall that Lemma 3.2 can be adapted to observe a higher norm with the L^2-norm. In fact, for any function $\phi = \sum_{i\in\Lambda_\lambda} a_i e_i$ we have:

$$\|\phi\|_{H_0^1(G)} = \left(\sum_{i\in\Lambda_\lambda} a_i^2\lambda_i\right)^{1/2} \leq \sqrt{\lambda}Ce^{C\sqrt{\lambda}}\left\|\sum_{i\in\Lambda_\lambda} a_i e_i\right\|_{L^2(G_0)} \leq Ce^{(C+1)\sqrt{\lambda}}\|\phi\|_{L^2(G_0)}.$$

- **Analyticity of the space variable.** If (α, G) satisfies the hypotheses of Theorem 2.6, we can easily prove as in [34, Theorem 1] that the free averaged solutions of the heat equation preserve the analyticity with respect to the spatial variable in the interior of the domain.



(a) The controlled solution in $t = 1$ with $\rho = 1_{(1,2)}$ (b) The controlled solution in $t = 1$ with $\rho = 1_{(0,1)}$

Fig. 4. The state at time $t = 1$ of the averaged solutions of the heat equation after applying the controls constructed with the minimizer of J in V_{20}, V_{50} and V_{100} with $y^0 = \frac{1}{2}$. In the left figure we have considered $\rho = 1_{(1,2)}$ and we observe that those functions are controls that take the solutions almost to equilibrium. The computations are obtained by approximating the density functions with Dirac masses, which explains why the solution does not converge exactly to 0. In the right one, $\rho = 1_{(0,1)}$ and we see that the minimizer of J do not take the solutions to equilibrium, so the functions that we have obtained cannot be considered controls.

- **Diffusion with negative values.** Regarding the cases where the diffusion takes strictly negative values we do not have null averaged controllability. Indeed, under that hypothesis we can easily prove that:

$$\lim_{n\to\infty} \frac{\|\tilde{u}(T,\cdot;e_n)\|_{L^2(G)}}{\|\tilde{u}(\cdot;e_n)\|_{L^2((0,T)\times G)}} = +\infty.$$

- **Some density functions in which the problem remains open.** There are some density functions which satisfy neither (2.3) for some $r > 1/2$ nor (2.4) for some $r < 1/2$. For those density functions their (non-)observability properties are still unproved, for instance, those satisfying (1.6). It is an infinite dimensional class since it contains all functions provided by (2.7) for $k = 1$.
- **Unique continuation.** It would be interesting to have a proof of the unique continuation property for the averaged dynamics of any random variable α, even when it takes negative values. Indeed, there are some random variables whose density functions do not satisfy (2.3) for any $r > 0$ (for instance, $\rho(\alpha) = 2\alpha 1_{(0,1)}(\alpha)$), so their unique continuation is still unproved. In particular, we wonder if the unique continuation is preserved when ρ is too irregular, as a counter-example would probably be of such type.
- **Measurable control domains.** The observability properties proved in Theorem 2.6 can be extended to sets of the type $E \times G_0$, for E a measurable set. Indeed, we can use the approach of [2,3,10,35], which complement the ideas of [11] with some results from Measure Theory.
- **More general heat equation.** An interesting problem that remains open is the study of the averaged observability properties of the random heat equation when the lower terms are also random terms, as:

$$y_t - \operatorname{div}(\sigma(x,\alpha)\nabla y) + A(x,\alpha)\cdot\nabla y + a(x,\alpha)y = 0.$$

In particular, this is interesting when the averaged convection operator and the averaged diffusion operator do not commute. Unfortunately, the techniques presented in this paper do not help in that direction since they rely on being the eigenfunctions associated to the elliptic operator independent of α to ensure that the averages of the respective eigenfunctions remain orthogonal. Consequently, they can only be applied to equations of the type $y_t - \alpha\mathcal{L}y$, for $\mathcal{L}$ an elliptic self-adjoint operator of compact resolvent. Thus, a theoretical or an in-depth numerical study would be of high interest for those operators.
- **Numerics and simulations.** It is an interesting question to develop rigorously the numerics for (7.1), including the speed of convergence. Determining precise methods that reproduce accurately the optimal control and the equation is also an open problem. Moreover, it would be interesting to illustrate with high accuracy that when $\rho = 1_{(0,1)}$ the norms of the optimal controls for taking the average to a distance of at most ε of the origin explodes. In addition, it would be interesting to perform similar simulations with $\rho = e^{-\alpha}1_{(0,+\infty)}$ and $\rho = \sqrt{2}1_{(0,1)}$ with the objective of determining numerically if with those density functions the approximate controllability in average holds.
- **Other random equations.** There are many other interesting questions involving random PDEs such as Schrödinger, wave, beam or Stokes equations:
  - The Schrödinger equations with random diffusions satisfying the uniform, exponential, Laplace, normal, Chi-squared and Cauchy distributions were studied in [2]. There, the authors show that the averaged dynamics may be conservative or diffusive depending on the probability density, which leads to averaged controllability properties of very different kind. They consider the uniform distribution in any segment of $\mathbb{R}$ and the exponential distribution in $[1,+\infty)$, though their proof is valid in any segment of the type $[K,+\infty)$ for any $K \in \mathbb{R}$. However, the problem of determining the dynamics and controllability properties of the averaged Schrödinger equations with arbitrary distributions is still open.
  - The wave equation with random discrete diffusion was studied in [36] and more abstractly in [37], whereas understanding the general case is still an open challenge.
  - Another interesting equation for which we can consider randomness in the higher order term is the Beam equation. In fact, in [38] an optimization problem involving the cost of the control and the average of the square of the mass at a final time T is studied, but

it is an open question to determine its exact (or null) controllability.
- The Stokes equation with random diffusion has not been studied in the literature. However, we can get analogous versions of Theorems 2.3 and 2.6 for the Stokes equation with random diffusion as of the heat equation by considering [39, Theorem 3.1]. Nonetheless, determining if the analogous of Theorem 2.4 is true remains an open problem because a lack of a comparison theorem prevent from using analogous arguments.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was started in a scientific visit of the first author to the Friedrich-Alexander-Universität Erlangen-Nürnberg, funded by the Research Grants - Short-Term Grants, 2020 (57507441) of the German Academic Exchange Service. This project has also received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement NO: 694126-DyCon). The work of the first author was also supported by grants from Région Ile-de-France. The work of the second author is also partially supported by the Air Force Office of Scientific Research (AFOSR), United States under Award NO: FA9550-18-1-0242, by the Grant MTM2017-92996-C2-1-R COSNET of MINECO (Spain), by the Alexander von Humboldt-Professorship program, the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No.765579-ConFlex, and the Transregio 154 Project "Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks" of the German DFG, Germany.

Appendix. Proof of Proposition 5.1

As in Section 5, C denotes a sufficiently large positive constant that may be different each time it appears and which just depends on G, G_0 and ρ. In particular, it does not depend on the index N. Similarly, $\tilde{C}$ is a constant sufficiently large that just depends on G, G_0 and ρ and $\overline{C}$ is a sufficiently large constant depending on those parameters and $\tilde{C}$. Finally, $\lfloor\cdot\rfloor$ denotes the floor function of a real number.

Let us fix $q=(q_1,\dots,q_d)$ and $\ell>0$ such that:

$$K:=[q_1,q_1+\ell]\times\cdots\times[q_d,q_d+\ell]\subset\subset G\setminus G_0.$$

We also fix a positive non-trivial function $\varsigma\in\mathcal{D}(B_{\mathbb{R}^d}(0,1))$. We define for $(\gamma_1,\dots,\gamma_d)\in[0,1]^d$:

$$p(\gamma_1,\dots,\gamma_d):=q+\ell(\gamma_1,\dots,\gamma_d),$$

which is a parametrization of K. With this in mind, we define the functions:

$$\phi_N(x):=\sum_{i_1,\dots,i_d=0}^{\lfloor\tilde{C}\sqrt{N}\rfloor}c_{i,N}\,\varsigma_{i,N}(x),$$

$$\text{for } \varsigma_{i,N}(x):=\varsigma\left(3\tilde{C}\sqrt{N}\,\frac{x-p\left(\frac{i}{\lfloor\tilde{C}\sqrt{N}\rfloor}\right)}{\ell}\right),\tag{A.1}$$

for $i:=(i_1,\dots,i_d)$ and $c_{i,N}$ and $\tilde{C}$ a large constant to be defined later on (see Fig. 5 for an illustration of how K and the support of ϕ_1 may look like). Let us check that for some $\tilde{C}$ and $c_{i,N}$ the sequence ϕ_N given by (A.1) satisfies (5.1)–(5.3):

- We have that:

$$\operatorname{supp}(\phi_N)\subset\left\{x:d(x,K)<\frac{\ell}{3\tilde{C}\sqrt{N}}\right\}.\tag{A.2}$$

Since the right-hand side of (A.2) is a decreasing sequence of sets and since $K\subset\subset G\setminus G_0$ we can easily prove (5.1) for $\tilde{C}$ large enough.
- In order to have (5.2) we just need to find a non-trivial solution of the system:

$$\langle\phi_N,e_i\rangle_{L^2(G)}=0,\quad\forall i\in\Lambda_N.\tag{A.3}$$

We remark that the system (A.3) is a linear homogeneous system with $\left\lfloor\tilde{C}\sqrt{N}\right\rfloor^d$ unknowns (the constants $(c_{i,N})_i$) and $|\Lambda_N|$ equations, so from Weyl's law (see Lemma 3.1) and by taking $\tilde{C}$ large enough we obtain that there are more unknowns than equations, which implies that (A.3) has a non-trivial solution. In particular, we can fix $(c_{i,N})_i$ a non-null tuple such that ϕ_N is a solution of (A.3).
- In order to prove (5.3) it suffices to prove that for $\overline{C}>0$ large enough and all $N\in\mathbb{N}$ we have that:

$$\|\Delta\phi_N\|_{L^2(G)}\le\frac{\overline{C}N}{2}\|\phi_N\|_{L^2(G)}.\tag{A.4}$$

Indeed, from (A.4) we obtain that:

$$\|\phi_N\|_{L^2(G)}\ge\frac{2\|\Delta\phi_N\|_{L^2(G)}}{\overline{C}N}\ge\frac{2\|\mathcal{P}^{\perp}_{\overline{C}N}\Delta\phi_N\|_{L^2(G)}}{\overline{C}N}\ge2\|\mathcal{P}^{\perp}_{\overline{C}N}\phi_N\|_{L^2(G)},$$

so we find that:

$$\|\mathcal{P}_{\overline{C}N}\phi_N\|^2_{L^2(G)}=\|\phi_N\|^2_{L^2(G)}-\|\mathcal{P}^{\perp}_{\overline{C}N}\phi_N\|^2_{L^2(G)}\ge\frac{3}{4}\|\phi_N\|^2_{L^2(G)},$$

which is (5.3) squared. So, let us prove (A.4). We clearly have for all $i,\tilde{i}\in\left\{0,\dots,\left\lfloor\tilde{C}\sqrt{N}\right\rfloor\right\}^d$ satisfying $i\neq\tilde{i}$ that $\operatorname{supp}(\varsigma_{i,N})\cap\operatorname{supp}(\varsigma_{\tilde{i},N})=\emptyset$. Thus, we have that:

$$\begin{aligned}\|\Delta\phi_N\|^2_{L^2(G)}&=\sum_{i_1,\dots,i_d=0}^{\lfloor\tilde{C}\sqrt{N}\rfloor}c^2_{i,N}\left(\frac{3\tilde{C}\sqrt{N}}{\ell}\right)^4\\&\quad\times\int_G|\Delta\varsigma|^2\left(3\tilde{C}\sqrt{N}\,\frac{x-p\left(\frac{i}{\lfloor\tilde{C}\sqrt{N}\rfloor}\right)}{\ell}\right)dx\\&=\left(\sum_{i_1,\dots,i_d=0}^{\lfloor\tilde{C}\sqrt{N}\rfloor}c^2_{i,N}\right)\left(\frac{3\tilde{C}\sqrt{N}}{\ell}\right)^3\|\Delta\varsigma\|^2_{L^2(B(0,1))}\\&\le C\left(\sum_{i_1,\dots,i_d=0}^{\lfloor\tilde{C}\sqrt{N}\rfloor}c^2_{i,N}\right)\left(\frac{3\tilde{C}\sqrt{N}}{\ell}\right)^3\|\varsigma\|^2_{L^2(B(0,1))}\end{aligned}$$

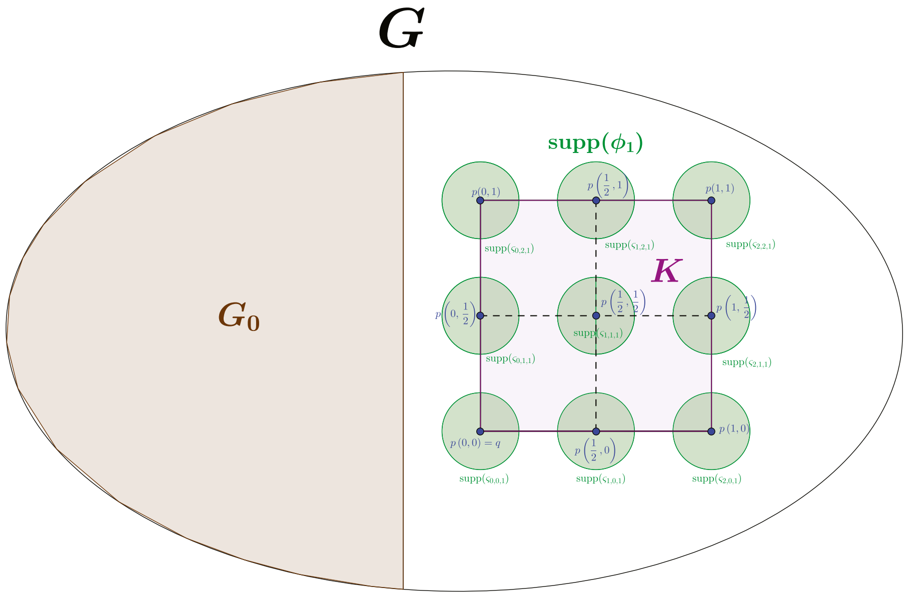


Fig. 5. An illustration of the support of ϕ_1 in a domain belonging to $\mathbb{R}^2$.

$$= C \sum_{i_1,\ldots,i_d=0}^{\lfloor \tilde{C}\sqrt{N} \rfloor} c_{i,N}^2 \left(\frac{3\tilde{C}\sqrt{N}}{\ell}\right)^4 \int_G |\varsigma|^2 \left(3\tilde{C}\sqrt{N}\frac{x - p\left(\frac{i}{\lfloor \tilde{C}\sqrt{N}\rfloor}\right)}{\ell}\right) dx$$

$$\leq C\tilde{C}^4 N^2 \sum_{i_1,\ldots,i_d=0}^{\lfloor \tilde{C}\sqrt{N} \rfloor} c_{i,N}^2 \int_G |\varsigma|^2 \left(3\tilde{C}\sqrt{N}\frac{x - p\left(\frac{i}{\lfloor \tilde{C}\sqrt{N}\rfloor}\right)}{\ell}\right) dx$$

$$= CN^2 \|\phi_N\|^2_{L^2(G)}. \tag{A.5}$$

Consequently, the sequence ϕ_N satisfies (A.4) for $\overline{C}$ large enough depending on $\tilde{C}$, and hence it also satisfies (5.3).

References

[1] E. Zuazua, Stable observation of additive superpositions of partial differential equations, Systems Control Lett. 93 (2016) 21–29.
[2] Q. Lü, E. Zuazua, Averaged controllability for random evolution partial differential equations, J. Math. Pure. Appl. 105 (3) (2016) 367–414.
[3] J. Coulson, B. Gharesifard, A.-R. Mansouri, On average controllability of random heat equations with arbitrarily distributed diffusivity, Automatica 103 (2019) 46–52.
[4] H.O. Fattorini, D.L. Russell, Exact controllability theorems for linear parabolic equations in one space dimension, Arch. Ration. Mech. Anal. 43 (4) (1971) 272–292.
[5] S. Micu, E. Zuazua, On the controllability of a fractional order parabolic equation, SIAM J. Control Optim. 44 (6) (2006) 1950–1972.
[6] L. Miller, On the controllability of anomalous diffusions generated by the fractional Laplacian, Math. Control Signal. 18 (3) (2006) 260–271.
[7] U. Biccari, M. Warma, E. Zuazua, Local elliptic regularity for the Dirichlet fractional Laplacian, Adv. Nonlinear Stud. 17 (2) (2017) 387–409.
[8] U. Biccari, V. Hernández-Santamaría, Controllability of a one-dimensional fractional heat equation: theoretical and numerical aspects, IMA J. Math. Control I. 36 (4) (2019) 1199–1235.
[9] E. Borel, Sur les zéros des fonctions entières, Acta Math. 20 (1897) 357–396.
[10] J. Apraiz, L. Escauriaza, G. Wang, C. Zhang, Observability inequalities and measurable sets, J. Eur. Math. Soc. 16 (11) (2014) 2433–2475.
[11] L. Miller, A direct Lebeau-Robbiano strategy for the observability of heat-like semigroups, Discrete Cont. Dyn. B 14 (2010) 1465–1485.
[12] L. Miller, Geometric bounds on the growth rate of null-controllability cost for the heat equation in small time, J. Differential Equations 204 (1) (2004) 202–226.
[13] L. Miller, The control transmutation method and the cost of fast controls, SIAM J. Control. Optim. 45 (2) (2006) 762–772.
[14] G. Tenenbaum, M. Tucsnak, New blow-up rates for fast controls of Schrödinger and heat equations, J. Differential Equations 243 (1) (2007) 70–100.
[15] S. Ervedoza, E. Zuazua, Sharp observability estimates for heat equations, Arch. Ration. Mech. Anal. 202 (3) (2011) 975–1017.
[16] C. Laurent, M. Léautaud, Observability of the heat equation, geometric constants in control theory, and a conjecture of Luc Miller, 2018, arXiv preprint arXiv:1806.00969.
[17] C.M. Bender, S.A. Orszag, Advances Mathematical Methods for Scientists and Engineers, Mcgraw-Hill book company, 1978.
[18] V. Ivrii, 100 years of Weyl's law, B. Math. Sci. 6 (3) (2016) 379–452.
[19] Q. Lü, A lower bound on local energy of partial sum of eigenfunctions for Laplace-Beltrami operators*, ESAIM: COCV 19 (1) (2013) 255–273.
[20] G. Lebeau, L. Robbiano, Contrôle exact de l'équation de la chaleur, Commun. Part. Diff. Eq. 20 (1–2) (1995) 335–356.
[21] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, New York, 1983.
[22] J.-M. Coron, S. Guerrero, Singular optimal control: a linear 1-D parabolic–hyperbolic example, Asymptotic Anal. 44 (3, 4) (2005) 237–257.
[23] S.G. Krantz, H.R. Parks, A Primer of Real Analytic Functions, Springer Science & Business Media, 2002.
[24] E. Fernández-Cara, E. Zuazua, The cost of approximate controllability for heat equations: the linear case, Adv. Differential Equations 5 (4–6) (2000) 465–514.
[25] E. Zuazua, Averaged control, Automatica 50 (12) (2014) 3077–3087.
[26] M. Abdelli, C. Castro, Numerical approximation of the averaged controllability for the wave equation with unknown velocity of propagation, ESAIM: COCV 27 (64) (2021) 1–26.
[27] R. Glowinski, J.L. Lions, Exact and approximate controllability for distributed parameter systems, Acta Numer. 1 (1994) 269–378.
[28] A. Münch, E. Zuazua, Numerical approximation of null controls for the heat equation: ill-posedness and remedies, Inverse Problems 26 (8) (2010) 085018.
[29] E. Fernández-Cara, A. Münch, Strong convergent approximations of null controls for the 1D heat equation, SÉMA J. 61 (1) (2013) 49–78.
[30] Q. Lü, X. Zhang, A concise introduction to control theory for stochastic partial differential equations, 2021, arXiv preprint arXiv:2101.10678.
[31] M. Lazar, J. Lohéac, Control of parameter dependent systems, 2020, Hal-03035494.
[32] Y. Privat, E. Trélat, E. Zuazua, Optimal shape and location of sensors for parabolic equations with random initial data, Arch. Ration. Mech. Anal. 216 (3) (2015) 921–981.
[33] J.A. Bárcena-Petisco, Null controllability of the heat equation in pseudo-cylinders by an internal control, ESAIM: COCV 26 (122) (2020) 1–34.
[34] L. Escauriaza, S. Montaner, C. Zhang, Analyticity of solutions to parabolic evolutions and applications, SIAM J. Math. Anal. 49 (5) (2017) 4064–4092.
[35] K.D. Phung, G. Wang, An observability estimate for parabolic equations from a measurable set in time and its applications, J. Eur. Math. Soc. 15 (2) (2013) 681–703.

[36] M. Lazar, E. Zuazua, Averaged control and observation of parameter-depending wave equations, C. R. Math. 352 (6) (2014) 497–502.
[37] J. Lohéac, E. Zuazua, Averaged controllability of parameter dependent conservative semigroups, J. Differential Equations 262 (3) (2017) 1540–1574.
[38] F.J. Marín, J. Martínez-Frutos, F. Periago, Robust averaged control of vibrations for the Bernoulli–Euler beam equation, J. Optim. Theory Appl. 174 (2) (2017) 428–454.
[39] F.W. Chaves-Silva, G. Lebeau, Spectral inequality and optimal cost of controllability for the Stokes system, ESAIM:COCV 22 (4) (2016) 1137–1162.