OPTIMAL SHAPE AND LOCATION OF SENSORS OR ACTUATORS IN PDE MODELS

EMMANUEL TRÉLAT

Abstract

We report on a series of works done in collaboration with Y. Privat and E. Zuazua, concerning the problem of optimizing the shape and location of sensors and actuators for systems whose evolution is driven by a linear partial differential equation. This problem is frequently encountered in applications where one wants to optimally design sensors in order to maximize the quality of the reconstruction of solutions by using only partial observations, or to optimally design actuators in order to control a given process with minimal efforts. For example, we model and solve the following informal question: what is the optimal shape and location of a thermometer?

Note that we want to optimize not only the placement but also the shape of the observation or control subdomain over the class of all possible measurable subsets of the domain having a prescribed Lebesgue measure. By probabilistic considerations we model this optimal design problem as the one of maximizing a spectral functional interpreted as a randomized observability constant, which models optimal observability for random initial data.

Solving this problem strongly depends on the operator in the PDE model and requires fine knowledge on the asymptotic properties of eigenfunctions of that operator. For parabolic equations like heat, Stokes or anomalous diffusion equations, we prove the existence and uniqueness of a best domain, proved to be regular enough, and whose algorithmic construction depends in general on a finite number of modes. In contrast, for wave or Schrödinger equations, relaxation may occur and our analysis reveals intimate relations with quantum chaos, more precisely with quantum ergodicity properties of the Laplacian eigenfunctions.

MSC2010: primary 93B07; secondary 49K20, 49Q10, 35P20, 58J51.
Keywords: observability, controllability, partial differential equations, shape optimization, spectral inequalities, quantum ergodicity.
1 Introduction and modeling

Our objective is to address the problem of optimizing the shape and location of sensors and actuators for processes modeled by a linear partial differential equation. Such questions are frequently encountered in engineering applications in which one aims at placing optimally, for instance, some given sensors on a system in order to achieve then the best possible reconstruction from observed signals. Here we also want to optimize the shape of sensors, without prescribing any a priori restriction on their regularity. Such problems have been little treated from the mathematical point of view. Our aim is to provide a relevant and rigorous mathematical model and setting in which the question can be addressed. Since controllability and observability are dual notions, we essentially focus on observability. The equations that we will investigate are mainly the wave equation

\[ \partial_{tt} y = \Delta y \]

or the Schrödinger equation

\[ \partial_t y = i \Delta y \]

or general parabolic equations

\[ \partial_t y = Ay \]

like heat-like, Stokes and anomalous diffusion equations for instance, settled on some open bounded connected subset \( \Omega \) of a Riemannian manifold, with various possible boundary conditions that can be Dirichlet, Neumann, mixed or Robin.

1.1 Spectral optimal design formulation. The first question arising is the one of formulating the problem in a relevant way. There are indeed several possible approaches to model the optimal observation problem; in particular we have to make precise the meaning of optimality here.

Informal considerations. To begin with, let us focus on the wave equation (1) with Dirichlet conditions on \( \partial \Omega \) (considerations for (2) and (3) are similar). The domain \( \Omega \) may represent a cavity in which some signals are propagating, in which we want to design and place some sensors that will then perform some measurements over a certain horizon of time, in view of a reconstruction inverse problem aiming at getting full information on the wave signals from the knowledge of these partial measurements.

We want to settle a relevant and appropriate mathematical formulation of the question of knowing what is the best possible shape and location of sensors, achieving the “best possible” observation in some sense.
A first obvious but important remark is that, in the absence of any constraint, certainly
the best strategy consists of observing the solutions over the whole domain $\Omega$, that is, place sensors everywhere. This is however clearly not reasonable and in practice the subdomain covered by sensors is limited, due for instance to cost considerations. From
the mathematical point of view, this constraint is taken into account by considering as the set of unknowns, the set of all possible measurable subsets $\omega$ of $\Omega$ that are of Lebesgue measure $|\omega| = L|\Omega|$, where $L \in (0, 1)$ is some fixed real number.

Given such a subset $\omega$ representing the sensors (and that we will try to optimize), we observe the restriction $y|_\omega$ of solutions of (1) over a certain time interval $[0, T]$ for some fixed $T > 0$, while wanting that these observations be enough to be indeed able to reconstruct the whole solutions in the most efficient way. This injectivity property is usually called observability.

**Observability inequality.** We recall that the wave equation (1) is observable on $\omega$ in time $T$ if there exists $C > 0$ such that

$$C \|(y(0, \cdot), \partial_t y(0, \cdot))\|_{L^2(\Omega) \times H^{-1}(\Omega)}^2 \leq \int_0^T \int_{\omega} |y(t, x)|^2 \, dx \, dt,$$

for all solutions $y$ of (1). This is called an observability inequality.

It is well known that, for $\omega$ open and $\partial \Omega$ smooth, observability holds if the pair $(\omega, T)$ satisfies the Geometric Control Condition (GCC) in $\Omega$ (see Bardos, Lebeau, and Rauch [1992]), according to which every geodesic ray that propagates in $\Omega$ at unit speed and reflects on its boundary according to the laws of geometric optics (like in a billiard) should intersect $\omega$ within time $T$ (note that this result has been extended to the case of time-varying domains $\omega(t)$ in Le Rousseau, Lebeau, Terpolilli, and Trélat [2017]). On Figure 1, on the

![Illustration of the Geometric Control Condition](image)

Figure 1: Illustration of the Geometric Control Condition

right, GCC is not satisfied because of the existence of trapped rays: there are solutions of
the wave equation that can never be observed on \( \omega \). Note that GCC is necessary if there is no geodesic ray grazing \( \omega \) (see Humbert, Privat, and Trélat [2016]).

The observability constant \( C_T(\chi_\omega) \), defined as

\[
C_T(\chi_\omega) = \inf \left\{ \int_0^T \int_\Omega \chi_\omega(x)|y(t,x)|^2 \, dx \, dt \mid y \text{ is solution of (1)},
\right.
\]

\[
\| (y(0, \cdot), \partial_t y(0, \cdot)) \|_{L^2(\Omega) \times H^{-1}(\Omega)} = 1
\]

is the largest nonnegative constant \( C_T(\chi_\omega) \) such that (4) holds true. Here, the notation \( \chi_\omega \) stands for the characteristic function of \( \omega \). We have observability if \( C_T(\chi_\omega) > 0 \). The observability constant is defined in a similar way for the Schrödinger equation (2) and for the general parabolic equation (3) (see Privat, Trélat, and Zuazua [2015b, 2016a]). The constant \( C_T(\chi_\omega) \) measures the well-posedness of the inverse problem of reconstructing the whole solutions of (1) from partial measurements on \([0, T] \times \omega\).

At first sight it seems therefore relevant to model the problem of maximizing observability as the optimal design problem

(5)

\[
\sup_{\chi_\omega \in \mathcal{U}_L} C_T(\chi_\omega)
\]

where

\[
\mathcal{U}_L = \{ \chi_\omega \mid \omega \subset \Omega \text{ measurable}, \ |\omega| = L|\Omega| \}.
\]

We stress that, in this problem, we want to optimize not only the placement but also the shape of \( \omega \) over all possible measurable subsets of \( \Omega \) having a prescribed measure. We do not put any restriction on the a priori regularity of \( \omega \): the search is over subsets that do not have a prescribed shape, that are not necessarily BV, etc. This lack of compactness shall naturally raise important mathematical difficulties.

Anyway, modeling optimal observability as the problem (5) of maximizing the deterministic observability constant leads to a mathematical problem that is difficult to handle from the theoretical point of view, and more importantly, that is not fully relevant in view of practical issues. Let us explain these two difficulties and let us then explain how to adopt a slightly different model.

The first difficulty is due to the emergence of crossed terms in the spectral expansion of solutions. More precisely, let us fix in what follows a Hilbert basis \( (\phi_j)_{j \in \mathbb{N}^*} \) of \( L^2(\Omega) \) consisting of eigenfunctions of the Dirichlet-Laplacian operator \(-\Delta\) on \( \Omega \), associated with the positive eigenvalues \((\lambda_j)_{j \in \mathbb{N}^*}\) with \( \lambda_1 \leq \cdots \leq \lambda_j \to +\infty \). Since any solution \( y \) of (1) can be expanded as

\[
y(t,x) = \sum_{j=1}^{+\infty} \left(a_j e^{i \sqrt{\lambda_j} t} + b_j e^{-i \sqrt{\lambda_j} t}\right) \phi_j(x)
\]
where the coefficients \( a_j \) and \( b_j \) account for initial data, it follows that

\[
C_T(\chi_\omega) = \frac{1}{2} \inf_{(a_j), (b_j) \in \ell^2(\mathbb{C})} \sum_{j=1}^{+\infty} \left( |a_j e^{i \sqrt{\lambda_j} t} + b_j e^{-i \sqrt{\lambda_j} t}| \phi_j(x) \right)^2 dx dt,
\]

and then maximizing this functional over \( \mathcal{U}_L \) appears to be very difficult from the theoretical point of view, due to the crossed terms \( \int_\omega \phi_j \phi_k dx \) measuring the interaction over \( \omega \) between distinct eigenfunctions. The difficulty is similar to the one appearing in the well known problem of determining what are the best constants in Ingham’s inequalities (see Jaffard and Micu [2001], Jaffard, Tucsnak, and Zuazua [1997], and Privat, Trélat, and Zuazua [2013b]).

The second difficulty with the model (5) is its lack of practical relevance. Indeed, the observability constant \( C_T(\chi_\omega) \) is deterministic and provides an account for the worst possible case: in this sense, it is a pessimistic constant. In practice, we perform the optimal design of sensors a priori, once for all, in view then of realizing a large number of measures (i.e., for many initial conditions). While performing many measurements, it may be expected that the worst case does not occur so often, and one would like that the observation be optimal for most of experiments. This leads us to consider rather an averaged version of the observability inequality over random initial data.

**Randomized observability constant.** We define what we call the randomized observability constant by

\[
C_{T, \text{rand}}(\chi_\omega) = \frac{1}{2} \inf \left\{ \mathbb{E} \int_0^T \int_\omega \left| \sum_{j=1}^{+\infty} \left( \beta_{1,j}^v a_j e^{i \sqrt{\lambda_j} t} + \beta_{2,j}^v b_j e^{-i \sqrt{\lambda_j} t} \right) \phi_j(x) \right|^2 dx dt \right\},
\]

where \( (\beta_{1,j}^v)_{j \in \mathbb{N}^*} \) and \( (\beta_{2,j}^v)_{j \in \mathbb{N}^*} \) are two sequences of i.i.d. random laws (for instance, Bernoulli) on a probability space \((\mathcal{X}, \mathcal{G}, \mathbb{P})\), and \( \mathbb{E} \) is the expectation over the \( \mathcal{X} \) with respect to the probability measure \( \mathbb{P} \). Definitions are similar for other equations (Schrödinger, heat, Stokes, etc). The constant \( C_{T, \text{rand}}(\chi_\omega) \) corresponds to the largest nonnegative constant of an averaged version of the observability inequality over random initial data. Indeed, with respect to the previous expression, the Fourier coefficients of the initial data have been randomized. In turn, by independence and taking the expectation, all crossed terms disappear and we obtain the following explicit expression of \( C_{T, \text{rand}}(\chi_\omega) \), for any of
the equations (1), (2) and (3). For the latter, we assume that \((\phi_j)_{j \in \mathbb{N}^*}\) is a Hilbert basis of \(L^2(\Omega, \mathbb{C})\) consisting of (complex-valued) eigenfunctions of the operator \(-A\), associated with the (complex) eigenvalues \((\lambda_j)_{j \in \mathbb{N}^*}\) such that \(\text{Re}(\lambda_1) \leq \cdots \leq \text{Re}(\lambda_j) \leq \cdots\).

**Theorem 1** (Privat, Trélat, and Zuazua [2015b, 2016a]). For every measurable subset \(\omega\) of \(\Omega\), we have

\[
C_{T,\text{rand}}(\chi_\omega) = T \inf_{j \in \mathbb{N}^*} \gamma_j(T) \int_\omega |\phi_j(x)|^2 \, dx
\]

where

\[
\gamma_j(T) = \begin{cases} 
1/2 & \text{for the wave equation (1),} \\
1 & \text{for the Schrödinger equation (2),} \\
e^{2\text{Re}(\lambda_j)T - \frac{1}{2\text{Re}(\lambda_j)}} & \text{for the parabolic equation (3).}
\end{cases}
\]

Note that we always have \(C_T(\chi_\omega) \leq C_{T,\text{rand}}(\chi_\omega)\) and that the inequality is strict for instance in each of the following cases:

- 1D Dirichlet waves on \(\Omega = (0, \pi)\), whenever \(T\) is not an integer multiple of \(\pi\) (see Privat, Trélat, and Zuazua [2013b]);
- multi-D Dirichlet waves on \(\Omega\) stadium-shaped, when \(\omega\) contains an open neighborhood of the wings (in that case, \(C_T(\chi_\omega) = 0\); see Privat, Trélat, and Zuazua [2016a]).

**Formulation of the optimal observability problem.** Taking into account the fact that, in practice, it is expected that a large number of measurements is to be done, rather than (5), we finally choose to model the problem of best observability as the problem of maximizing the functional \(\chi_\omega \mapsto C_{T,\text{rand}}(\chi_\omega)\) over the set \(\mathcal{U}_L\), that is:

\[
(6) \quad \sup_{\chi_\omega \in \mathcal{U}_L} \inf_{j \in \mathbb{N}^*} \gamma_j(T) \int_\omega |\phi_j(x)|^2 \, dx
\]

This is a spectral optimal design problem.

Note that the randomized observability constant \(C_{T,\text{rand}}(\chi_\omega)\) can also be interpreted as a time-asymptotic observability constant (see Privat, Trélat, and Zuazua [ibid.]).

**Remark 1.** Note that, in (5) or in (6) we take an infimum over all (randomized) initial data. In contrast, if we fix some given initial data, maximizing the functional \(\chi_\omega \mapsto \int_0^T \int_\omega |y(t,x)|^2 \, dx \, dt\) over \(\mathcal{U}_L\) is a problem that can be easily solved thanks to a decreasing rearrangement argument (see Privat, Trélat, and Zuazua [2015a]), showing that there always exists (at least) one optimal set \(\omega^*\). The regularity of \(\omega^*\) depends on the initial data. We can show that it may be a Cantor set of positive measure, even for smooth
data. Of course, in practice designing optimal sensors depending on initial data would make no sense and this is why we consider in our model an infimum over all (or almost all) initial data.

**Remark 2.** As already underlined, in our search of the best possible subset $\omega$, we do not impose any restriction to $\omega$ but its measurability. If we restrict the search to subsets having uniformly bounded (by some $A > 0$) perimeter or total variation or satisfying the $1/A$-cone property, or if we restrict ourselves to subsets parametrized by some compact or finite-dimensional set, then quite straightforwardly there exists (at least) one optimal set $\omega^*$. But then the complexity of $\omega^*$ may then increase with $A$ (spillover phenomenon). We will observe this phenomenon when considering, further in the paper, a truncated version of (6) with a finite number of spectral modes.

Imposing no restriction on $\omega$ is our choice here because we want to address the mathematical question of knowing if there is a ”very best” subdomain over all possible measurable subsets $\omega$ such that $|\omega| = L|\Omega|$.  

### 1.2 Related problems and existing results.

We first mention that the optimal observability problem on which we have focused up to now is related, by duality, to the problem of determining what is the best control domain for controlling to rest, for instance, the wave equation with internal control

$$\partial_{tt} y - \Delta y = \chi_\omega u.$$ 

We have addressed such best actuator problems in Privat, Trélat, and Zuazua [2013a, 2016b, 2017] with a similar randomization approach.

Another closely related problem is that of finding the best possible domain to stabilize the equation

$$\partial_{tt} y - \Delta y = -k \chi_\omega \partial_t y$$

thanks to a localized damping (see Privat and Trélat [2015] for results). Best means here that one may want to design $\omega$ (over $\mathcal{U}_L$) such that exponential decrease of solutions of the above locally damped wave equation is maximal. Historically, up to our knowledge, the first papers addressing this problem were Hébrard and Henrot [2003, 2005], in which the authors studied this problem in 1D and provided complete characterizations of the optimal set whenever it exists, for the problem of determining the best possible shape and position of the damping subdomain of a given measure.

Due to their relevance in engineering applications, optimal design problems for placing sensors or actuators for processes modeled by partial differential equations have been investigated in a large number of papers. Difficulties come from the facts that solutions live in infinite-dimensional spaces and that the class of admissible designs is not closed for the
standard and natural topology. Very few works take into consideration those aspects. In most of existing contributions, numerical tools are developed to solve a simplified version of the optimal design problem where either the PDE has been replaced with a discrete approximation, or the class of optimal designs is replaced with a compact finite dimensional set – see for example Kumar and Seinfeld [1978], Morris [2011], Uciński and Patan [2010], van de Wal and de Jager [2001], and Wouwer, Point, Porteman, and Remy [2000] where the aim is most often to optimize the number, the place and the type of sensors in order to improve the estimation of the state of the system. Sensors often have a prescribed shape (for instance, balls with a prescribed radius) and then the problem consists of placing optimally a finite number of points (the centers of the balls) and thus is finite-dimensional. Of course, the resulting optimization problem is already challenging. Here we want to optimize also the shape of the observation set without making any a priori restrictive assumption to the class of shapes (such as bounded variation) and the search is made over all possible measurable subsets.

From the mathematical point of view, the issue of studying a relaxed version of optimal design problems for shape and position of sensors or actuators has been investigated in a series of articles. In Bellido and Donoso [2007] the authors investigate the problem modeled in Sigmund and Jensen [2003] of finding the best possible distributions of two materials (with different elastic Young modulus and different density) in a rod in order to minimize the vibration energy in the structure. The authors of Allaire, Aubry, and Jouve [2001] also propose a convexification formulation of eigenfrequency optimization problems applied to optimal design. In Fahroo and Ito [1996] are discussed several possible criteria for optimizing the damping of abstract wave equations and derive optimality conditions for a certain criterion related to a Lyapunov equation. In Münch and Periago [2011], the authors study a homogenized version of the optimal location of controllers for the heat equation problem for fixed initial data, noticing that such problems are often ill-posed. In Allaire, Münch, and Periago [2010], the authors consider a similar problem and study the asymptotic behavior as the final time $T$ goes to infinity of the solutions of the relaxed problem; they prove that optimal designs converge to an optimal relaxed design of the corresponding two-phase optimization problem for the stationary heat equation. We also mention Fernández-Cara and Münch [2012] where, still for fixed initial data, numerical investigations are used to provide evidence that the optimal location of null-controllers of the heat equation problem is an ill-posed problem.
2 Study of the optimal design problem

To address the optimal design problem (6), we distinguish between parabolic equations (3) (like heat, Stokes or anomalous diffusion equations) on the one part and the hyperbolic equations (1) and (2) on the other.

As a first remark, since the infimum in (6) involves all spectral modes $j \in \mathbb{N}^*$, solving the problem will require some knowledge on the asymptotic behavior of the squares $|\phi_j|^2$ of the eigenfunctions as $j \to +\infty$. Note also that, because of the weights $\gamma_j(T)$ in (6), there is a strong difference between the parabolic case where $\gamma_j(T)$ is exponentially increasing as $j \to +\infty$ and the hyperbolic (wave and Schrödinger) case where $\gamma_j(T)$ remains constant.

2.1 The parabolic case. For parabolic equations (3), the situation is particularly nice and we have the following result, under several quite general assumptions on the operator $A$, which are satisfied for heat and Stokes equations and also for anomalous diffusion equations, i.e., $A = -(-\Delta)^{\alpha}$, with $\alpha > 1/2$. Note that anomalous diffusion equations provide relevant models in many problems encountered in physics (plasma with slow or fast diffusion, aperiodic crystals, spins, etc), in biomathematics, in economy or in imaging sciences.

Theorem 2 (Privat, Trélat, and Zuazua [2015b]). Let $T > 0$ be arbitrary. Assume that $\partial \Omega$ is piecewise $C^1$. There exists a unique$^1$ optimal observation domain $\omega^*$ solving (6). Moreover $\omega^*$ is open and semi-analytic; in particular, it has a finite number of connected components. Additionally, we have $C_T (\chi_{\omega^*}) < C_{T, \text{rand}} (\chi_{\omega^*})$.

Note that this existence and uniqueness result holds for every fixed orthonormal basis of eigenfunctions of the operator but the optimal set depends on the specific choice of the Hilbert basis.

This result (of which one can find an even more general version in Privat, Trélat, and Zuazua [ibid.]) gives a short and satisfactory positive answer to the question of knowing if there is a “very best” observation domain among all possible measurable subsets. Moreover, we are going to see further that there even exists a nice algorithmic procedure to compute the optimal set $\omega^*$, which happens to be fully characterized by a finite number of modes only.

The fact that the optimal set $\omega^*$ is semi-analytic is a strong (and desirable) regularity property. In addition to the fact that $\omega^*$ has a finite number of connected components, this implies also that $\omega^*$ is Jordan measurable, that is, $|\partial \omega^*| = 0$. This is in contrast with

\footnotesize
$^1$Here, it is understood that the optimal set $\omega^*$ is unique within the class of all measurable subsets of $\Omega$ quotiented by the set of all measurable subsets of $\Omega$ of zero measure.

Let us explain shortly why Theorem 2 applies to (3) with $A = -(\Delta)^\alpha$ (power of the Dirichlet-Laplacian) for every $\alpha > 1/2$. It is instrumental in the proof to use the fine lower estimates of Apraiz, Escauriaza, Wang, and Zhang [2014, Theorem 5], stating that

$$\int_{\omega} |\phi_j(x)|^2 \geq C e^{-C \sqrt{\mu_j}} \quad \forall j \in \mathbb{N}^*$$

(here, the $\mu_j$’s are the eigenvalues of $-\Delta$) where the constant $C > 0$ is uniform with respect to $\chi_{\omega} \in \mathcal{U}_L$. This uniform property is remarkable and particularly useful here in our context. The requirement $\alpha > 1/2$ comes from a balance between the above lower estimate and the exponential weight $\gamma_j(T) \sim e^{\mu_j T}$, yielding in that case a favorable coercivity property, itself implying compactness features that are crucial in the proof. Another instrumental tool in the proof is then a refined minimax theorem due to Hartung [1982].

In the critical case $\alpha = 1/2$, the conclusion of Theorem 2 holds true as well provided that the time $T$ is moreover large enough.

Furthermore, still considering $A = -(\Delta)^\alpha$, it is proved in Privat, Trélat, and Zuazua [2016a] that:

- in the Euclidean square $\Omega = (0, \pi)^2$, when considering the usual Hilbert basis of eigenfunctions consisting of products of sine functions, for every $\alpha > 0$ there exists a unique optimal set in $\mathcal{U}_L$ (as in the theorem), which is moreover open and semi-analytic (whatever the value of $\alpha > 0$ may be);

- in the Euclidean disk $\Omega = \{x \in \mathbb{R}^2 \mid \|x\| < 1\}$, when considering the usual Hilbert basis of eigenfunctions parametrized in terms of Bessel functions, for every $\alpha > 0$ there exists a unique optimal set $\omega^*$ (as in the theorem), which is moreover open, radial, with the following additional property:
  - if $\alpha > 1/2$ then $\omega^*$ consists of a finite number of concentric rings that are at a positive distance from the boundary (see Figure 2);
  - if $\alpha < 1/2$ (or if $\alpha = 1/2$ and $T$ is small enough) then $\omega^*$ consists of an infinite number of concentric rings accumulating at the boundary.

This quite surprising result shows that the complexity of the optimal shape does not only depend on the operator but also depends on the geometry of the domain $\Omega$. The proof of these properties is difficult in the case $\alpha < 1/2$; it requires involved estimates for Bessel functions combined with the use of quantum limits in the disk (like in the hyperbolic case in the next section) and analyticity considerations.
2.2 The hyperbolic case. For the wave equation (1) and the Schrödinger equation (2), since all weights $\gamma_j(T)$ in (6) are equal (and, in turn, the time $T$ thus plays no role), in particular highfrequencies play an important role. Setting $\mu_j = |\phi_j|^2 \, dx$, we see that getting knowledge on the asymptotic behavior of $\mu_j$ as $j \to +\infty$ is now required. Noting that $\mu_j$ is a probability measure for every $j \in \mathbb{N}^*$, the weak limits of the sequence $(\mu_j)_{j \in \mathbb{N}^*}$ now enter into consideration.

**Theorem 3** (Privat, Trélat, and Zuazua [ibid.]). Assume that the sequence of probability measures $\mu_j = |\phi_j|^2(x) \, dx$ converges weakly to the uniform measure $\frac{1}{|\Omega|} \, dx$ (assumption called Quantum Unique Ergodicity on the base) and that there exists $p \in (2, +\infty]$ such that the sequence of eigenfunctions $(\phi_j)_{j \in \mathbb{N}^*}$ is uniformly bounded in $L^p(\Omega)$. Then

$$
(7) \quad \sup_{a \in \overline{U}_L} \inf_{j \in \mathbb{N}^*} \int_{\omega} |\phi_j(x)|^2 \, dx = L \quad \forall L \in (0, 1).
$$

To prove this result, we define $J(\chi_\omega) = \inf_{j \in \mathbb{N}^*} \int_{\omega} |\phi_j(x)|^2 \, dx$ and we introduce a convexified version of the optimal design problem (5) (“relaxation” procedure in shape optimization), by considering the convex closure of the set $U_L$ for the $L^\infty$ weak star topology, that is $\overline{U}_L = \{a \in L^\infty(\Omega, [0, 1]) \mid \int_{\Omega} a(x) \, dx = L|\Omega|\}$. The convexified problem then consists of maximizing the functional $a \mapsto J(a) = \inf_{j \in \mathbb{N}^*} \int_{\Omega} a(x)\phi_j(x)^2 \, dx$ over $\overline{U}_L$. Clearly, a maximizer does exist, and it is easily seen by using Cesàro means of squares of eigenfunctions that the constant function $a(\cdot) = L$ is a maximizer. But since the functional $J$ is not lower semi-continuous it is not clear whether or not there may be a gap between the problem (5) and its convexified version. **Theorem 3** above shows that, under appropriate spectral assumptions, there is no gap. The proof consists of a kind of homogenization procedure which consists of building a maximizing sequence of subsets for the problem of maximizing $J$, showing that it is always possible to increase the values
of $J$ by considering subsets of measure $L|\Omega|$ having an increasing number of connected components. The construction strongly uses the assumption that the sequence $(\mu_j)_{j \in \mathbb{N}^*}$ has a unique weak limit (QUE on the base), which is very strong as we explain below.

**Link with quantum chaos.** Let us comment on the spectral assumptions done in the theorem.

They are satisfied in 1D: for instance in $\Omega = (0, \pi)$, the Dirichlet eigenfunctions $\phi_j(x) = \sqrt{\frac{2}{\pi}} \sin(jx)$ are uniformly bounded in $L^\infty(\Omega)$ and their squares weakly converge to $1/\pi$.

In multi-D, the assumptions are very strong and actually, except in the 1D case, we are not aware of domains $\Omega$ for which the assumptions are satisfied. Firstly, in general the eigenfunctions are not uniformly bounded in $L^\infty(\Omega)$ but, to the best of our knowledge, nothing seems to be known in general on the uniform $L^p$-boundedness property for some $p > 2$. Secondly the probability measures $\mu_j = |\phi_j|^2 \, dx$ may have several weak limits. This question is related with deep open questions in mathematical physics and semi-classical analysis where one of the most fascinating open questions is to determine what can be these weak limits, called *quantum limits* or *semi-classical measures*. The famous Shnirelman theorem (see Colin de Verdière [1985], Gérard and Leichtnam [1993], Šnirelman [1974], and Zelditch and Zworski [1996]) states that, seeing the domain $\Omega$ as a billiard, if the Riemannian geodesic flow is ergodic (for the canonical measure) then there exists a subsequence of $(\mu_j)_{j \in \mathbb{N}^*}$ of density one converging vaguely to the uniform measure $\frac{1}{|\Omega|} \, dx$ (Quantum Ergodicity, in short QE – still on the base, here). This result however lets open the possibility of having an exceptional subsequence of measures $\mu_j$ converging vaguely to some other measure, for instance, the Dirac measure along a closed geodesic (scar$\text{s}$ in quantum physics, see Faure, Nonnenmacher, and De Bièvre [2003]). The QUE assumption mentioned above consists of assuming that the whole sequence $(\mu_j)_{j \in \mathbb{N}^*}$ converges vaguely to the uniform measure. It is likely that QUE holds true on a negatively curved compact manifold (QUE conjecture, see Sarnak [2011] for a survey).

The idea is here that QUE ensures a delocalization property of the energy of high-frequency eigenfunctions. The quantity $\int_\Omega \phi_j^2(x) \, dx$ is interpreted as the probability of finding the quantum state of energy $\lambda_j^2$ in $\omega$. The functional $J(\chi_\omega)$ considered above can be viewed as a measure of eigenfunction concentration, which we seek to maximize over $\mathcal{U}_L$.

**Theorem 3** thus reveals intimate connections between domain optimization and asymptotic spectral properties or quantum ergodicity properties of $\Omega$ (quantum chaos theory). It is interesting to notice that such a relationship was suggested in the early work Chen, Fulling, Narcowich, and Sun [1991] concerning the exponential decay properties of dissipative wave equations.
To end with these remarks on asymptotic properties on eigenfunctions, we note that the weak convergence of the measures $\mu_j$ which is established in the several results mentioned above is however weaker than the convergence of the functions $\phi_j^2$ for the weak topology of $L^1(\Omega)$ that we need in our context. Indeed, the weak convergence of measures may fail to capture sets whose measure of the boundary is positive (such as Cantor of positive measure). This is why we also assume the $L^p$ uniform boundedness property with $p > 2$ because then, by the Portmanteau theorem and since $\Omega$ is bounded, both notions of convergence coincide.

**The assumptions are not sharp.** The spectral assumptions made in Theorem 3 are sufficient but are not necessary. It is indeed proved in Privat, Trélat, and Zuazua [2016a] that (7) is still satisfied if $\Omega$ is a 2D square (with the usual eigenfunctions consisting of products of sine functions) or if $\Omega$ is a 2D disk (with the usual eigenfunctions parametrized by Bessel functions), although, in the latter case, the eigenfunctions do not equidistribute as the eigenfrequencies increase, as illustrated by the well-known whispering galleries effect (see Figure 3): from the mathematical point of view, there exists a subsequence of $\left(\mu_j\right)_{j \in \mathbb{N}^*}$ converging to the Dirac along the boundary of the disk.

**On the existence of an optimal set.** By Theorems 1 and 3, the maximal possible value of $C_{T,\text{rand}}(\chi_\omega)$ over the set $\mathcal{U}_L$ is equal to $TL/2$. We now comment on the problem of existence of an optimal set: is the supremum reached in (7)? By compactness of the convexified set $\overline{\mathcal{U}}_L$, it is easy to see that the maximum of $J$ over $\overline{\mathcal{U}}_L$ is reached (in general in an infinite number of ways), but since $\mathcal{U}_L$ is not compact for any appropriate topology, the question of the reachability of the supremum of $J$ over $\mathcal{U}_L$, that is, the existence of an
optimal classical set, is a difficult question in general. In particular cases it can however be addressed using harmonic analysis (see Privat, Trélat, and Zuazua [2013b, 2016a]):

- In 1D, assume that $\Omega = (0, \pi)$, with the usual Hilbert basis of Dirichlet eigenfunctions made of sine functions. The supremum of $J$ over $\mathcal{U}_L$ (which is equal to $L$) is reached if and only if $L = 1/2$. In that case, it is reached for all measurable subsets $\omega \subset (0, \pi)$ of measure $\pi/2$ such that $\omega$ and its symmetric image $\omega' = \pi - \omega$ are disjoint and complementary in $(0, \pi)$.

- In the 2D square $\Omega = (0, \pi)^2$, with the usual basis of Dirichlet eigenfunctions made of products of sine functions, the supremum of $J$ over the more specific class of all possible subsets $\omega = \omega_1 \times \omega_2$ of Lebesgue measure $L \pi^2$, where $\omega_1$ and $\omega_2$ are measurable subsets of $(0, \pi)$, is reached if and only if $L \in \{1/4, 1/2, 3/4\}$. In that case, it is reached for all such sets $\omega$ satisfying

$$\frac{1}{4}(\chi_\omega(x,y) + \chi_\omega(\pi-x,y) + \chi_\omega(x,\pi-y) + \chi_\omega(\pi-x,\pi-y)) = L$$ for almost all $(x, y) \in [0, \pi^2]$ (see Figure 4).

- In the 2D disk $\Omega = \{x \in \mathbb{R}^2 \mid \|x\| < 1\}$, with the usual Hilbert basis of eigenfunctions defined in terms of Bessel functions, the supremum of $J$ (which is equal to $L$) over the class of all possible subsets $\omega = \{(r, \theta) \in [0, 1] \times [0, 2\pi] \mid r \in \omega_r, \theta \in \omega_\theta\}$ such that $|\omega| = L \pi$, where $\omega_r$ is any measurable subset of $[0, 1]$ and $\omega_\theta$ is any measurable subset of $[0, 2\pi]$, is reached if and only if $L = 1/2$. In that case, it is reached for all subsets $\omega = \{(r, \theta) \in [0, 1] \times [0, 2\pi] \mid \theta \in \omega_\theta\}$ of measure $\pi/2$, where $\omega_\theta$ is any measurable subset of $[0, 2\pi]$ such that $\omega_\theta$ and its symmetric image $\omega_\theta' = 2\pi - \omega_\theta$ are disjoint and complementary in $[0, 2\pi]$.

In general, the question of the existence of an optimal set is completely open. In view of the partial results above and in view of the results of the next section, we conjecture that, for generic domains $\Omega$ and for generic values of $L \in (0, 1)$, the supremum in (7) is not
reached and hence there does not exist any optimal set. We have no clue how to address this conjecture in general.

3 Spectral approximation of the optimal design problem

Motivated by the probable absence of optimal set for wave and Schrödinger equations as explained previously, and motivated by the objective of building the optimal set for parabolic equations, it is natural to consider the following finite-dimensional spectral approximation of the problem (6), namely:

\[
\sup_{\chi_\omega \in \mathcal{U}_L} \min_{1 \leq j \leq N} \gamma_j(T) \int_{\omega} |\phi_j(x)|^2 \, dx
\]

for any \( N \in \mathbb{N}^* \). This is a spectral truncation where we keep only the \( N \) first modes. We have the following easy result.

**Theorem 4** (Privat, Trélat, and Zuazua [2015b, 2016a]). Let \( T > 0 \) be arbitrary. There exists a unique optimal observation domain \( \omega^N \) solving (8). Moreover \( \omega^N \) is open and semi-analytic and thus it has a finite number of connected components.

Actually, since there is only a finite number of modes in (8), existence and uniqueness of an optimal set \( \omega^N \) is not difficult to prove (by a standard minimax argument), as well as a \( \Gamma \)-convergence property of \( J_N \) towards \( J \) for the weak star topology of \( L^\infty \), where we have set \( J_N(\chi_\omega) = \min_{1 \leq j \leq N} \gamma_j(T) \int_{\omega} |\phi_j(x)|^2 \, dx \). In particular, the sets \( \omega^N \) constitute a maximizing sequence for the (convexified) problem of maximizing \( J \) over \( \overline{\mathcal{U}_L} \), and this, without geometric or ergodicity assumptions on \( \Omega \) (under the assumptions of Theorem 3, these sets constitute a maximizing sequence for the problem of maximizing \( J \) over \( \mathcal{U}_L \)).

Let us now analyze how \( \omega^N \) behaves as \( N \) increases, by distinguishing between the parabolic case (3) and the hyperbolic case (1) and (2).

3.1 The parabolic case. For parabolic equations (3), under general assumptions on the operator \( A \), which are satisfied for heat, Stokes equations and anomalous diffusion equations with \( \alpha > 1/2 \), remarkably, the sequence of optimal sets \( (\omega^N)_{N \in \mathbb{N}^*} \) is stationary.

**Theorem 5** (Privat, Trélat, and Zuazua [2015b]). For every \( T > 0 \) there exists \( N_0(T) \in \mathbb{N}^* \) such that

\[
\omega^{N_0(T)} = \omega^* = \omega \quad \forall N \geq N_0(T).
\]

As a consequence, the optimal observation set \( \omega^* \) whose existence and uniqueness has been stated in Theorem 2 can actually be built from a finite-dimensional spectral approximation, by keeping only a finite number of modes. This stationarity property is illustrated
on Figure 5 where we compute, as announced in the abstract, the “optimal thermometer in the square”. For this example, we have $N_0(0.05) = 16$, i.e., for $T = 0.05$ the optimal domain is computed thanks to the 16 first eigenmodes.

It is also proved in Privat, Trélat, and Zuazua [2015b] that the function $T \mapsto N_0(T) \in \mathbb{N}^*$ is nonincreasing and that if $\text{Re}(\lambda_1) < \text{Re}(\lambda_2)$ then $N_0(T) = 1$ as soon as $T$ is large enough, which means that the optimal set $\omega^*$ is entirely determined by the first eigenfunction if the observation time $T$ is large.

### 3.2 The hyperbolic case.

In contrast to the previous parabolic case, for wave and Schrödinger equations, the fact that all eigenmodes have the same weight ($\gamma_j(T)$ remains constant) causes a strong instability of the optimal sets $\omega^N$, whose complexity increases drastically as $N$ increases.

Moreover, the sets $\omega^N$ have a finite number of connected components, expected to increase in function of $N$. The numerical simulations of Figures 6 and 7 show the shapes of these sets. Their increasing complexity (number of connected components) which can be observed as $N$ increases is in accordance with the conjecture of the nonexistence of an optimal set for (6).

Of course, however, up to some subsequence the sequence of maximizers $\chi_{\omega^N}$ of $J_N$ converges (in weak-star topology) to some maximizer $a \in \overline{U}_L$ of $J$. 

![Figure 5: Dirichlet heat equation on $\Omega = (0, \pi)^2$, $L = 0.2$, $T = 0.05$. Row 1, from left to right: optimal domain $\omega^N$ (in green) for $N = 1, 4, 9$. Row 2, from left to right: optimal domain $\omega^N$ (in green) for $N = 16, 25, 36.\]
In the 1D case $\Omega = (0, \pi)$ with Dirichlet boundary conditions, it can be proved that, for $L > 0$ sufficiently small, the optimal set $\omega^N$ maximizing $J_N$ is the union of $N$ intervals concentrating around equidistant points and that $\omega^N$ is actually the worst possible subset for the problem of maximizing $J_{N+1}$: in other words, the optimal domain for $N$ modes is the worst possible one when considering the truncated problem with $N + 1$ modes. This is the spillover phenomenon, noticed in Hébrard and Henrot [2005] and proved in Privat, Trélat, and Zuazua [2013b] (the proof is highly technical).

**Weighted observability inequalities.** This intrinsic instability is due to the fact that in (6) all modes have the same weight. This is so in the mathematical definition of the (deterministic or randomized) observability constant. One could argue that high frequencies are difficult to observe and, trying to reflect the Heisenberg uncertainty principle of quantum physics, this leads to the intuition that lower frequencies should be in some sense more weighted than higher ones. One can then introduce a weighted version of the observability inequality (4), by considering, for instance the (equivalent) inequality

$$C_{T,\sigma}(\chi_\omega) \left( \|y^0, y^1\|_{L^2 \times H^{-1}}^2 + \sigma \|y^0\|^{-2}_{H^{-1}} \right) \leq \int_0^T \int_\omega |y(t, x)|^2 \, dx \, dt$$
where $\sigma \geq 0$ is some weight. We have $C_{T, \sigma}(\chi_\omega) \leq C_T(\chi_\omega)$, and considering as before an averaged version of this weighted observability inequality over random initial data leads to

$$C_{T, \sigma, \text{rand}}(\chi_\omega) = \frac{T}{2} \inf_{j \in \mathbb{N}^*} \frac{\lambda_j^2}{\sigma + \lambda_j^2} \int_\omega \phi_j(x)^2 \, dx$$

where now the weights are an increasing sequence of positive real numbers converging to 1. Actually, if $\frac{\lambda_1^2}{\sigma + \lambda_1^2} < L < 1$ then highfrequencies do not play any role in the problem of maximizing $C_{T, \sigma, \text{rand}}$ over $U_L$ and we have the following result for not too small values of $L$.

**Theorem 6 (Privat, Trélat, and Zuazua [2016a]).** Assume that the whole sequence of probability measures $\mu_j = \phi_j^2(x) \, dx$ converges vaguely to the uniform measure $\frac{1}{|\Omega|} \, dx$ and that the sequence of eigenfunctions $\phi_j$ is uniformly bounded in $L^\infty(\Omega)$. Then for every $L \in \left(\frac{\lambda_1^2}{\sigma + \lambda_1^2}, 1\right)$ there exists $N_0 \in \mathbb{N}^*$ such that

$$\max_{\chi_\omega \in U_L} \inf_{j \in \mathbb{N}^*} \frac{\lambda_j^2}{\sigma + \lambda_j^2} \int_\omega \phi_j(x)^2 \, dx = \max_{\chi_\omega \in U_L} \inf_{1 \leq j \leq N} \frac{\lambda_j^2}{\sigma + \lambda_j^2} \int_\omega \phi_j(x)^2 \, dx \leq \frac{\lambda_1^2}{\sigma + \lambda_1^2} < L \quad \forall N \geq N_0.$$

In particular, the problem of maximizing $C_{T, \sigma, \text{rand}}$ over $U_L$ has a unique solution $\chi_{\omega_{N_0}}$ and moreover the set $\omega_{N_0}$ is open and semi-analytic.

This result says that, when highfrequencies are weighted as above, there exists a unique optimal observation set if $L$ is large enough, i.e., if one is allowed to cover a fraction of the whole domain $\Omega$ that is large enough. This is similar to what we have obtained in the parabolic case. Moreover the optimal set can then be computed from a finite number of modes because the sequence of optimal sets $\omega^N$ of the truncated problem is stationary. The threshold value $\frac{\lambda_1^2}{\sigma + \lambda_1^2}$ becomes smaller when $\sigma$ increases, in accordance with physical intuition. We do not know what may happen when $L \leq \frac{\lambda_1^2}{\sigma + \lambda_1^2}$ but we suspect that the situation is the same as when $\sigma = 0$ (spillover phenomenon and, probably, nonexistence of an optimal set); this conjecture is supported by some numerical simulations for the truncated problem (see Privat, Trélat, and Zuazua [ibid.]) which show that, when $L$ is small, the optimal domains have an increasing complexity as $N$ increases.

As before, we can notice that the assumptions of the above result, which are very strong, are not necessary and one can prove that the conclusion still holds true in a hypercube with Dirichlet boundary conditions when one considers the usual Hilbert basis made of products of sine functions.
4 Conclusion

We have modeled the problem of optimal shape and location of the observation or control domain having a prescribed measure, in terms of maximizing a spectral functional over all measurable subsets of fixed Lebesgue measure. This spectral functional can be interpreted as a randomized version of the observability constant over random initial data. For parabolic equations, we have existence and uniqueness of an optimal set, which can be determined from a finite number of modes. For wave and Schrödinger equations, the optimal observability problem is closely related to quantum chaos, in particular, asymptotic properties of eigenfunctions and we have seen that, generically, an optimal set should not exist, in accordance with the spillover phenomenon. In all cases, developing knowledge on concentration or delocalization properties of highfrequency eigenfunctions is crucial in order to address optimal observability issues.

We have seen that a way to avoid spillover and to recover existence and uniqueness of an optimal set for wave and Schrödinger equations is to consider weighted observability inequalities in which highfrequencies are penalized. Certainly, other approaches are possible, exploiting the physics of the problem.

Optimal boundary observability. In the paper we have focused on internal observation or control subdomains. Similar studies can be led for boundary subdomains. Optimal observability can be modeled by the optimal design problem

$$\sup_{|\omega|=L|\partial \Omega|} \inf_{j \in \mathbb{N}^*} \gamma_j(T) \int_\omega \frac{1}{\lambda_j} \left( \frac{\partial \phi_j}{\partial v} \right)^2 d \mathcal{H}^{n-1}$$

where now the Neumann traces of the Dirichlet-Laplacian eigenfunctions play a prominent role (in particular, their asymptotic properties). This problem, which interestingly can be interpreted as a spectral shape sensitivity problem, is studied in Privat, Trélat, and Zuazua [2018].

On the deterministic observability constant. We have let untouched the problem of maximizing the deterministic observability constant $C_T(\chi_\omega)$ over $\mathcal{U}_L$. Although we have explained that this problem is certainly less relevant in practice where a large number of measurements is performed, it is anyway very interesting from the mathematical point of view. The crossed terms (which we have ruled out by randomization) are then expected to have an important role. A first remark is that, extending the functional $C_T$ to $\overline{\mathcal{U}}_L$, for 1D wave equations the constant density $a \equiv L$ is not a maximizer of $C_T$ if $T \notin \pi \mathbb{N}^*$. Knowing if there is a relaxation phenomenon or not is an open problem.
Another remark is the following. It is proved in Humbert, Privat, and Trélat [2016] that, for the wave equation (1), given any measurable subset \( \omega \) of \( \Omega \), we have

\[
\lim_{T \to +\infty} \frac{C_T(\chi_\omega)}{T} = \frac{1}{2} \min \left( \inf_{j \in \mathbb{N}^*} \int_\omega \phi_j(x)^2 \, dx, \lim_{T \to +\infty} \inf_{\gamma \in \Gamma} \frac{1}{T} \int_0^T \chi_\omega(\gamma(t)) \, dt \right)
\]

where \( \Gamma \) is the set of all geodesic rays on \( \Omega \), provided that \( \omega \) has no grazing ray, i.e., provided that there exists no \( \gamma \in \Gamma \) such that \( \gamma(t) \in \partial \omega \) over a set of times of positive measure. This equality says that, in large time, the deterministic observability constant \( C_T(\chi_\omega) \) is the minimum of two quantities: the first one is exactly \( C_{T,\text{rand}}(\omega) \), which is the functional we have focused on throughout the paper; the second one is of a geometric nature and provides an account for the average time spent by geodesic rays in the observation subset. Although this result is only valid asymptotically in time, it gives the intuition that geodesic rays play an important role. In order to address the problem of maximizing \( C_T(\chi_\omega) \) over \( \mathcal{U}_L \), one should first try solve, for any \( T > 0 \),

\[
\sup_{\chi_\omega \in \mathcal{U}_L} \inf_{\gamma \in \Gamma} \frac{1}{T} \int_0^T \chi_\omega(\gamma(t)) \, dt.
\]

This is an interesting optimal design problem.

**Discretization issues.** In the search of an optimal observation domain for a PDE model, certainly the most usual approach in engineering applications is to discretize the PDE (for instance by means of finite elements), thus obtaining a family of equations in finite dimension, indexed by some \( h > 0 \) which can be thought as the size of the mesh. Given some fixed \( h \), one then performs an optimal design procedure to find, if it exists, an optimal observation set \( \omega^h \). The question is then natural to ask whether \( \omega^h \) converges, as \( h \to 0 \), to the (if it exists and is unique) optimal observation set \( \omega^* \) of the complete model. In other words, do the numerical optimal designs converge to the continuous optimal design as the mesh size tends to 0? Under which assumptions do the optimal designs commute with discretization schemes?

We have seen with the spectral truncation (which is a particular discretization method) that the answer is certainly negative for wave and Schrödinger equations but is positive for parabolic equations. The question is open for general discretization schemes and is of great interest in view of practical applications, all the more than discrete or semi-discrete models are often employed.
References


Sorbonne Université
Université Paris-Diderot SPC
CNRS, Inria
Laboratoire Jacques-Louis Lions, équipe CAGE
F-75005 Paris
France
emmanuel.trelat@upmc.fr
emmanuel.trelat@sorbonne-universite.fr