Averaged and Greedy Control

Enrique Zuazua

FAU -AvH
enrique.zuazua@fau.de

April 3, 2020

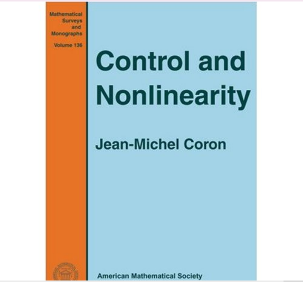
Table of Contents

1. In past decades controllability theory for PDE has evolved significantly.
2. Some of the most paradigmatic models are by now well understood: Wave and heat equations, in particular.
3. But theory lacks of unity. Often times rather different analytical tools are required to tackle different models/problems.
4. Practical applications need of **robust** control theoretical results and **fast numerical solvers**.
5. One of the key issues to be addressed in that direction is the controllability of **PDE models depending on parameters**, that represent uncertain or unknown quantities.
6. In this lecture we present some basic elements of the implementation of the **greedy methods** in this context and formulate some challenging open problems.
7. This leads to a new class of **Inverse Problems**.

What's known?

Many fundamental questions are by now well-understood (under the influence of the pioneering works of **D. Russell, J.-L. Lions** among others)

1. **Wave equations** by means of Microlocal techniques starting with the pioneering work of Bardos-Lebeau-Rauch (1988).
2. **Heat equations** by means of Carleman inequalities: Fursikov-Imanuvilov (1992); Lebeau-Robbiano (1995).
3. Control of **nonlinear models**: The return method, J.- M. Coron (1994), Steady-state control, J.-M. Coron - E. Trélat (2004).



What about numerics?

Much less is known!

- Pioneering works by **R. Glowinski and J. L. Lions** (*Acta Numerica* (1994)).
- Numerics and high frequency filtering for **wave equations**: S. Ervedoza & E. Zuazua, SpringerBriefs (2013), M. Tucsnak et al., E. Fernández-Cara & A. Münch, M. Asch - G. Lebeau - M. Nodet,....
- Numerics for **heat-like equations** based on Carleman inequalities, F. Boyer - F. Hubert - M. Morancey - J. Le Rousseau ...

Significant work remains to be done to bring the numerical theory to the same level as the PDE one.

And, overall, **robust numerical methods** are needed.

What about parameter-depending problems?

- **Singular perturbations**: From wave to heat (López-Zhang-Zuazua (2000)), viscous to inviscid conservation laws (Coron-Guerrero (2005), Guerrero-Lebeau (2007))
- **Homogenisation** (Castro-Zuazua (1997), G. Lebeau (1999), López-Zuazua (2002), Alessandrini-Escauriaza (2008)):

$$y_{tt} - div(a(x/\varepsilon)\nabla y) = 0; \quad y_t - div(a(x/\varepsilon)\nabla y) = 0.$$

- $T \to 0$ for heat equations ($\exp(-c/T)$): L. Miller (2004), G. Tenenbaum - M. Tucsnak (2007), P. Lissy (2015).

The analysis of these singular perturbation problems needs of significant *ad hoc* arguments and exhibits the lack of unified treatment.

Some of the most fundamental issues are still badly understood: Controllability for the heat equation with rapidly oscillating coefficients in multi-d? Cost of control as $T \to 0$ (What is c_L)?

Regular dependence on parameters

The issue of developing robust and efficient numerical solvers for the controllability of parameter-dependent problems is still poorly understood.

The state of the art: For each individual realisation of the relevant parameters check controllability and apply the corresponding numerical solver.

Limited validity and high computational cost!

Think for example on

$$y_{tt} - div(a(x, \nu))\nabla y) = 0$$

For each value of the parameter ν one should check whether the Geometric Control Condition holds and then develop the corresponding numerical algorithm on well adapted meshes, filtering high frequencies, etc.

Table of Contents

Consider the finite dimensional linear control system (possibly obtained from a PDE control problem after space discretisation)

$$\begin{cases} x'(t) = A(\nu)x(t) + Bu(t), \; 0 < t < T, \\ x(0) = x^0. \end{cases} \tag{1}$$

In (**??**):

- The (column) vector valued function $x(t,\nu) = \big(x_1(t,\nu), \ldots, x_N(t,\nu)\big) \in \mathbb{R}^N$ is the state of the system,
- ν is a multi-parameter living in a compact set K of $\mathbb{R}^d$,
- $A(\nu)$ is a $N \times N-$matrix,
- $u = u(t)$ is a M-component control vector in $\mathbb{R}^M$, $M \leq N$.

Given a control time $T > 0$ and a final target $x^1 \in \mathbb{R}^N$ we look for a control u such that the solution of (**??**) satisfies the averaged control property:

$$\int_K x(T,\nu)d\nu = x^1. \tag{2}$$

Theorem

[a] *Averaged controllability holds if and only the following rank condition is satisfied:*

$$rank\Big[B, \int_0^1 [A(\nu)]d\nu B, \int_0^1 [A(\nu)]^2 d\nu B, \dots\Big] = N. \tag{3}$$

[a] E. Zuazua, Automatica, 2014.

Drawbacks:

1. Nothing is said about the efficiency of the control for specific realisations of ν.
2. Complex (**and interesting !**) in the PDE setting. [1]
   Consider the transport equation with unknown velocity v,
   $$f_t + v f_x = 0,$$
   and take averages with respect to v. Then
   $$g(x,t) = \int f(x,t;v)\rho(v)dv$$
   then, for the Gaussian density ρ:
   $$\rho(v) = (4\pi)^{-1/2}\exp(-v^2/4)$$
   $$g(x,t) = h(x,t^2); \quad h_t - h_{xx} = 0.$$
   One can then employ parabolic techniques based on Carleman inequalities.

[1] Q. Lü & E. Z. Average Controllability for Random Evolution Equations, JMPA, 2016. Linked to averaging Lemmas (Golse - Lions - Perthame - Sentis)

Table of Contents

[2] Assume that the system depends on a parameter $\nu \in K \subset \mathbf{R}^d$, $d \geq 1$, K being a compact set, and controllability being fulfilled for all values of ν.

$$\begin{cases} x'(t) = A(\nu)x(t) + Bu(t), \; 0 < t < T, \\ x(0) = x^0. \end{cases} \tag{4}$$

Controls $u(t, \nu)$ are chosen to be of minimal norm satisfying the controllability condition:

$$x(T, \nu) = x^1, \tag{5}$$

and lead to a manifold of dimension d in $[L^2(0, T)]^M$:

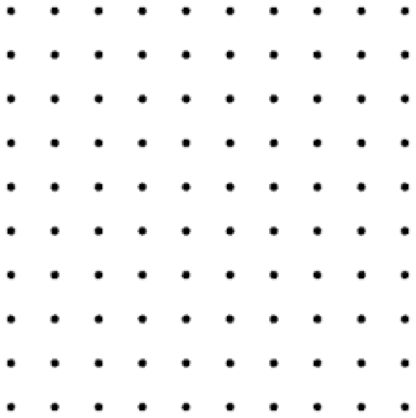
$$\nu \in K \subset \mathbf{R}^d \to u(t, \nu) \in [L^2(0, T)]^M.$$

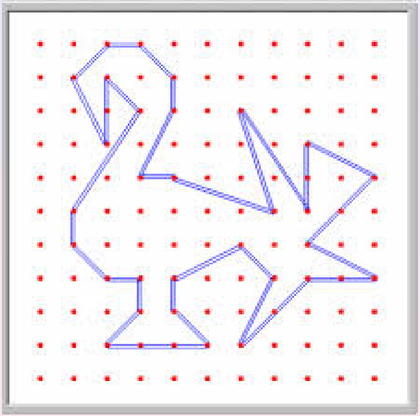
This manifold inherits the regularity of the mapping $\nu \to A(\nu)$.

To diminish the computational cost we look for the very distinguished values of ν that yield the best possible approximation of this manifold.

[2] M. Lazar & E. Zuazua, Greedy controllability of finite dimensional linear systems, Automatica, to appear.

Naive versus smart sampling of K





Our work relies on recent ones on greedy algorithms and reduced bases methods:

A. Cohen, R. DeVore, *Kolmogorov widths under holomorphic mappings*, IMA Journal on Numerical Analysis, to appear

A. Cohen, R. DeVore, *Approximation of high-dimensional parametric PDEs*, arXiv preprint, 2015.

Y. Maday, O. Mula, A. T. Patera, M. Yano, *The generalized Empirical Interpolation Method: stability theory on Hilbert spaces with an application to the Stokes equation*, submitted

M. A. Grepl, M Kärche, *Reduced basis a posteriori error bounds for parametrized linear-quadratic elliptic optimal control problems*, CRAS Paris, 2011.

S. Volkwein, PDE-Constrained Multiobjective Optimal Control by Reduced-Order Modeling, IFAC CPDE2016, Bertinoro.

Description of the Method

We look for the realisations of the parameter ν ensuring the best possible approximation of the manifold of controls

$$\nu \in K \subset \mathbf{R}^d \to u(t, \nu) \in [L^2(0, T)]^M$$

(of dimension d in $[L^2(0, T)]^M$) in the sense of the Kolmogorov width.[3]

Greedy algorithms **search for the values of ν leading to the most distinguished controls** $u(t, \nu)$, those that are farther away one from each other.

Given an error ε, the goal is to find $\nu_1, ..., \nu_{n(\epsilon)}$, so that for all parameter values ν the corresponding control $u(t, \nu)$ can be approximated by a linear combination of $u(t, \nu_1), ..., u(t, \nu_{n(\epsilon)})$ with an error $\leq \epsilon$.

An of course to do it with a minimum number $n(\epsilon)$.

[3] Ensure the optimal rate of approximation by means of all possible finite-dimensional subspaces.

Step 1. *Characterization of minimal norm controls by adjoints*

The adjoint system depends also on the parameter ν:

$$-\varphi'(t) = A^*(\nu)\varphi(t), t \in (0,\ T);\ \varphi(T) = \varphi^0. \tag{6}$$

The control is

$$u(t,\nu) = B^*\varphi(t,\nu),$$

where $\varphi(t,\nu)$ is the solution of the adjoint system associated to the minimizer of the following quadratic functional in $\mathbf{R}^{\mathbf{N}}$:

$$J_\nu\left(\varphi^0(\nu)\right) = \frac{1}{2}\int_0^T |B^*\varphi(t,\nu)|^2\,dt - < x^1, \varphi^0 > + < x^0, \varphi(0,\nu) > .$$

The functional is continuous and convex, and its coercivity is guaranteed by the Kalman rank condition that we assume to be satisfied for all ν.

Step 2. *Controllability distance*

Given two parameter values ν_1 and ν_2, how can we measure the distance between $u(t,\nu_1)$ and $u(t,\nu_2)$?

Of course the issue relies on the fact that these two controls are unknown!!!

Roughly: Compute the residual

$$||x(T,\nu_2) - x^1||$$

for the solution of the state equation ν_2 achieved by the control $u(t,\nu_1)$.

More precisely: Solve the Optimality System (OS):

$$-\varphi'(t) = A^*(\nu_2)\varphi(t) \; t \in (0,\ T);\; \varphi(T) = \varphi_1^0.$$

$$x'(t) = A(\nu_2)x(t) + BB^*\varphi(t,\nu_2),\, 0 < t < T,\, x(0) = x^0.$$

Then

$$\left|\nabla J_{\nu_2}(\varphi_1^0)\right| = ||x(T,\nu_2) - x^1|| \sim ||\varphi_1^0 - \varphi_2^0||.$$

Offline algorithm

Step 3. *Initialisation of the weak-greedy algorithm.* Choose any ν in K, $\nu = \nu_1$, and compute the minimizer of J_{ν_1}. This leads to φ_1^0.

Step 4. *Recursive choice of $\nu's$.*
Assuming we have $\nu_1, ..., \nu_p$, we choose ν_{p+1} as the maximiser of

$$\max_{\nu \in K} \min_{\phi \in span[\varphi_j^0, j=1,...,p]} |\nabla J_\nu(\phi)|$$

We take ν_{p+1} as the one realizing this maximum.
Note that

$$|\nabla J_\nu(\phi)| = ||x(T, \nu) - x^1||.$$

$x(T, \nu)$ being the solution obtained by means of the control $u = B^* \phi(t, \nu)$, ϕ being the solution of the adjoint problem associated to the initial datum ϕ^0 in $span[\varphi_j^0, j = 1, ..., p]$.

Step 5. *Stopping criterion.* Stop if the max $\leq \epsilon$.

Online part

Step 6. For a specific realisation of ν solve the finite-dimensional reduced minimisation problem:

$$\min_{\phi \in span[\varphi_j^0, j=1,\ldots,p]} |\nabla J_\nu(\phi)|.$$

This minimiser yields:

$$u(t,\nu) = B^*\varphi(t,\nu),$$

$\varphi(t,\nu)$ being the solution of the adjoint problem with datum ϕ at $t = T$.

The same applies for infinite-dimensional systems when A and B are bounded operators.

Theorem

The weak-greedy algorithm above leads to an optimal approximation method.
More precisely, if the set of parametres K is finite-dimensional, and the map $\nu \to A(\nu)$ is analytic, for all $\alpha > 0$ there exists $C_\alpha > 0$ such that for all other values of ν the control $u(\cdot, \nu)$ can be approximated by linear combinations of the weak-greedy ones as follows:

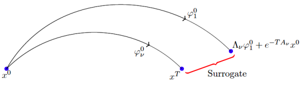
$$dist(u(\cdot, \nu); span[u(\cdot; \nu_j) : j = 1, ..., k]) \leq C_\alpha k^{-\alpha}.$$

4

[4] The approximation of the controls has to be understood in the sense above: Taking the control given by the corresponding adjoint solution, achieved by minimising the functional J over the finite-dimensional subspace generated by the adjoints for the distinguished parameter-values.

Potential improvements

1. **Find cheaper surrogates.** Is there a reduced model leading to lower bounds on controllability distances without solving the full Optimality System?



$$||x(T,\nu) - x_1|| \geq ??????$$

2. All this depends on the initial and final data: x_0, x_1.
   Can the search of the most relevant parameter-values ν be done independent of x_0, x_1?
   In other words, get **lower bounds on the controllability distances** between (A_1, B_1) and (A_2, B_2).

As we shall see this leads to Inverse Problems of a non-standard form

Table of Contents

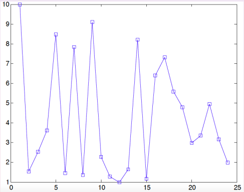
Semi-discrete wave equation

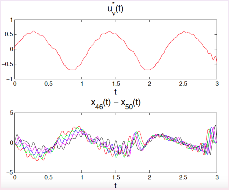
1. Finite difference approximation of the $1-d$ wave equation with 50 nodes in the space-mesh.
2. Unknown velocity v ranging within $[1, \sqrt{10}]$.
3. Discrete parameters taken over an equi-distributed set of 100 values
4. Boundary control
5. Sinusoidal initial data given: $y_0 = \sin(\pi x); y_1 \equiv 0$. Null final target.
6. Time of control $T = 3$.
7. Approximate control with error 0.5 in the energy.
8. Weak-greedy requires 24 snapshots $(\nu_1, ..., \nu_{24})$.
9. Offline time: 2.312 seconds (personal notebook with a 2.7 GHz processor and DDR3 RAM with 8 GB and 1,6 GHz).
10. Online time for one realisation ν: 7 seconds
11. Computational time for one single parameter value with standard methods: 51 seconds.

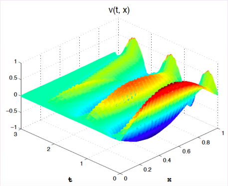
Choose a number at random within $[1, 10]$

Thank you for choosing π !

The greedy algo leads to:

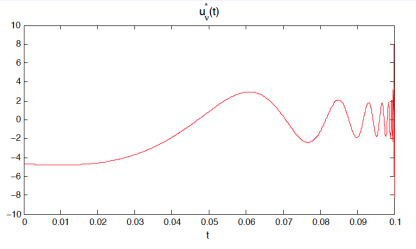


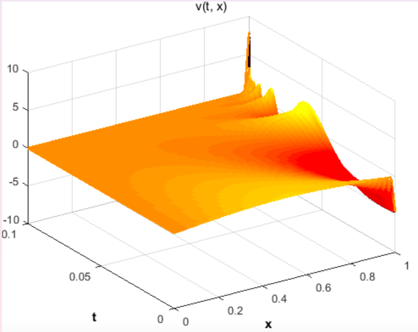




Semi-discrete heat equation

1. Finite difference approximation of the $1 - d$ heat equation with 50 nodes in the space-mesh.
2. Unknown diffusivity v ranging within $[1, 2]$.
3. Discrete parameters taken over an equi-distributed set of 100 values
4. Boundary control
5. Sinusoidal initial data given: $y_0 = \sin(\pi x)$. Null final target.
6. Time of control $T = 0.1$.
7. Weak-greedy requires 20 snapshots.
8. Approximate control with error 10^{-4} in each component.
9. The algo stops after 3 iterations: $\nu = 1.00, 1.18, 1.45$.
10. Offline time: 213 seconds.
11. Online time for one realisation $\nu = \sqrt{2}$: 1.5 seconds
12. Computational time for one single parameter value with standard methods: 37 seconds.


$u_v^*(t)$
10
8
6
4
2
0
−2
−4
−6
−8
−10
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1
t


v(t, x)
10
5
0
-5
-10
0.1
0.05
0
t
0
0.2
0.4
0.6
0.8
1
x

Open problems and perspectives

- The method be extended to PDE. But **analyticity of controls with respect to parameters** has to be ensured to guarantee optimal Kolmogorov widths. This typically holds for elliptic and parabolic equations. But not for wave-like equations.
  Indeed, solutions of
  $$y_{tt} - v^2 y_{xx} = 0$$
  do not depend analytically on the coefficient v.
  One expects this to be true for heat equations in the context of null-controllability. But this needs to be rigorously proved.
- **Cheaper surrogates** need to be found so to make the recursive choice process of the various $\nu's$ faster.
  1. For wave equations in terms of distances between the dynamics of the Hamiltonian systems of bicharacteristic rays?
  2. For $1-d$ wave equations in terms of spectral distances?
  3. For heat equations?

Table of Contents

Problem formulation

[5]

To better understand the complexity of the problem of applying the greedy methodology for control systems, independently of the initial and final data under consideration, it is natural to consider the following **diffusive equation as a model problem**.

Let Ω be a bounded domain of $\mathbb{R}^n$, $n \geq 1$. Fix $0 < \sigma_0 < \sigma_1$ and consider the class of scalar diffusivity coefficients

$$\Sigma = \{\sigma \in L^\infty(\Omega);\ \sigma_0 \leq \sigma \leq \sigma_1 \text{ a.e. in } \Omega\}.$$

For $\sigma \in \Sigma$, let $A_\sigma : H_0^1(\Omega) \to H^{-1}(\Omega)$ be the bounded operator given by

$$A_\sigma u = -\text{div}(\sigma \nabla u).$$

The inverse or resolvent operator $R_\sigma : H^{-1}(\Omega) \to H_0^1(\Omega)$.

The goal is to implement the greedy algo in the class of resolvent operators.

[5] Joint work with M. Choulli

The existing theory gives the answer for a given right hand side term:

$$-\text{div}(\sigma \nabla u) = f.$$

But we are interested on searching the most representative realisations of the resolvents as operators, independently of the value of f.
The analog at the control theoretical level would be to do it for the inverse of the Gramian operators rather than proceeding as above, for each specific data to be controlled.

The question under consideration is. How to find a surrogate (lower bound) for

$$dist(R_\sigma, span[R_{\sigma_1}, ..., R_{\sigma_k}])$$

?

The question is easy to solve when dealing with two resolvents R_1 and R_2.
But seems to become non-trivial in the general case.

This leads to a new class of Inverse Problems

Distance between two resolvents

It is easy to get a surrogate for the distance between two resolvents R_1 and R_2 corresponding to two different diffusivities σ_1 and σ_2:

$$A_1 - A_2 = A_1(R_2 - R_1)A_2,$$

$$\Big|A_1 - A_2\Big| \leq \sigma_1^2 \Big|R_1 - R_2\Big|.$$

$$\langle (A_1 - A_2)u, u\rangle_{-1,1} = \int_\Omega (\sigma_1 - \sigma_2)|\nabla u|^2 dx,$$

$$\int_\Omega (\sigma_1 - \sigma_2)|\nabla u|^2 dx \leq \Big|A_1 - A_2\Big| \Big|u\Big|^2_{H_0^1(\Omega)} \leq \sigma_1^2 \Big|R_1 - R_2\Big| \Big|u\Big|^2_{H_0^1(\Omega)}.$$

Now taking $u = u_\epsilon$ so that $|\nabla u_\epsilon|^2$ **constitutes an approximation of the identity** (for each $x_0 \in \Omega$) we get

$$||\sigma_1 - \sigma_2||_\infty \leq \sigma_1^2 \Big|R_1 - R_2\Big|.$$

This can be understoof in the context of Inverse Problems: The resolvent determines the diffusivity, with Lipschitz continuous dependence.

1d

Unfortunately, this argument does not seem to apply for estimating the distance to a subspace

$$R_1 - \sum_{j=1}^{k} \alpha_j R_j.$$

This is a non-standard inverse problems. We are dealing with linear combinations of $k+1$ resolvents and not only 2 as in classical identification problems

In $1-d$ the problem can be solved, thanks to the explicit representation of solutions[6]

$$-(\sigma(x)u_x)_x = f \text{ in } (0,1), \quad u_x(0) = 0 \text{ and } u(1) = 0. \tag{7}$$

$$u_x(x) = -\frac{1}{\sigma(x)} \int_0^x f(t)dt = -T_\sigma f \text{ a.e. } (0,1). \tag{8}$$

[6] Very much as in the context of homogenisation

$$||R_\sigma - R_{\widetilde{\sigma}}||_* = \Big|\frac{1}{\widetilde{\sigma}(x)} - \frac{1}{\sigma(x)}\Big|_{L^\infty((0,1))}.$$

$$\left(R_\tau f - \sum_{i=1}^N a_i R_i f\right)_x = \left(\sum_{i=1}^N \frac{a_i}{\sigma_i(x)} - \frac{1}{\tau(x)}\right) \int_0^x f(t)dt \text{ a.e. } (0,1) \quad (9)$$

$$\Big|R_\tau - \sum_{i=1}^N a_i R_i\Big|_* = \Big|\sum_{i=1}^N \frac{a_i}{\sigma_i(x)} - \frac{1}{\tau(x)}\Big|_{L^\infty((0,1))}. \quad (10)$$

This means that, in this $1d$ context, it suffices (?) to apply the greedy algo in L^∞ within the class of coefficients $1/\sigma(x)$.

Multi-dimensional extension?

Table of Contents

Consider control systems of the form

$$\begin{cases} x'(t) = A_j x(t) + Bu(t), \; 0 < t < T, \\ x(0) = x^0, \end{cases} \tag{11}$$

$j = 1, ..., k$.
Control operators:

$$P_j(x^0) = u_j(t), j = 1, ..., K.$$

Find a surrogate for

$$dist(P_j, span[P_\ell; \ell \neq j]) = sup_{||x^0||=1} dist(u_j(t), span[u_\ell(t) : \ell \neq j]).$$

We want an equivalent measure, but easier to be computed.

For two operators ($||P_1 - P_2||$) it suffices to consider the inverses, the Gramians: $||\Lambda_1 - \Lambda_2||$.

$$-\varphi_j'(t) = A_j^* \varphi_j(t) \ t \in (0,\ T);\ \varphi_j(T) = \phi.$$

$$x_j'(t) = A_j x_j(t) + BB^* \varphi_j(t),\ 0 < t < T,\ x_j(0) = 0,$$

$$\Lambda_j(\phi) = x_j(T).$$

What about

$$||x_1(T) - x_2(T)||?$$

Easier for PDE? For instance, for wave equations, take ϕ a Gaussian wave packet so that φ_j are Gaussian beams following the corresponding bicharacteristic rays. Then solve the controlled system. The output $x_j(T)$ should be close to a Gaussian wave packet as well.

Can we recover out of the distance $||x_1(T) - x_2(T)||$ the distances between coefficients?

This is an **Inverse Problem.**

But it seems NOT to be the case for distances to subspaces... How to handle this more complex and fundamental issue?