Greedy control

Martin Lazar
University of Dubrovnik

Berlin, 2016
7th European Mathematical Congress

Joint work with E: Zuazua, UAM, Madrid

WeConMApp

Outline

- Parametric dependent systems
- Reduced basis methods
- Greedy control
- Numerical example

Parameter dependent problems

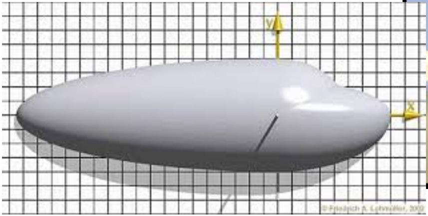
Real life processes depend on (a huge number of) parameters.

Parameter dependent problems

Real life processes depend on (a huge number of) parameters.

humidity

temperature



curvature

velocity

These parameters are variable, subject to uncertainty, undetermined ...
The study of a parameter dependent problems requires robust approach.

Control of parametric dependent system

Consider the finite dimensional linear control system

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

In (1):

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

We assume the system is (uniform) controllable for all $\nu \in \mathcal{N}$.

Controllability

The system (1) is controllable in time $T > 0$ if for any initial datum x^0 there exists a control u_ν such that x$(T) = 0$.
The control is not unique in general.
We restrict to a class of a minimal energy norm which provides uniqueness.

By assumption the system is controllable for any $\nu \in \mathcal{N}$.
As the dynamics depends on ν, so it does the control u_ν.

What does it mean in practice?

You measure the parameter value, and you determine the control by some standard methods.
And you repeat the process each time for any new value of ν.

Controllability

The system (1) is controllable in time $T > 0$ if for any initial datum x^0 there exists a control u_ν such that x$(T) = 0$.
The control is not unique in general.
We restrict to a class of a minimal energy norm which provides uniqueness.

By assumption the system is controllable for any $\nu \in \mathcal{N}$.
As the dynamics depends on ν, so it does the control u_ν.

What does it mean in practice?

You measure the parameter value, and you determine the control by some standard methods.
And you repeat the process each time for any new value of ν.

Can we do it better?

Greedy control

Fix a control time $T > 0$, an arbitrary initial data x^0, and a final target $x^T \in \mathbf{R}^N$.

The controls

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

constitute a manifold $U(\mathcal{N})$ of dimension d in $[L^2(0, T)]^M$.

The idea:

– to determine a finite number of values of ν that yield the best possible approximation of this "control manifold".

We do not do it in a "naive" way by simply taking a uniform mesh on $\mathcal{N}$ and then evaluating the control for each value ν corresponding to the nodes of this mesh. This would be too expensive.

We look for a distinguished parameter values yielding the optimal approximation by the smallest number of points.

Each control can be uniquely determined by the relation

$$\mathsf{u}_{\boldsymbol{\nu}} = \mathbf{B}^* e^{(T-t)\mathbf{A}_{\boldsymbol{\nu}}^*} \varphi_{\boldsymbol{\nu}}^0,$$

where $\varphi_{\boldsymbol{\nu}}^0 \in \mathbf{R}^N$ is the unique minimiser of a quadratic functional associated to the adjoint problem.

This minimiser can be expressed as the solution of the system

$$\mathbf{G}_{\boldsymbol{\nu}} \varphi_{\boldsymbol{\nu}}^0 = \mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}} \mathbf{x}^0,$$

where $\mathbf{G}_{\boldsymbol{\nu}}$ is the is the controllability Gramian

$$\mathbf{G}_{\boldsymbol{\nu}} = \int_0^T e^{(T-t)\mathbf{A}_{\boldsymbol{\nu}}} \mathbf{B}_{\boldsymbol{\nu}} \mathbf{B}_{\boldsymbol{\nu}}^* e^{(T-t)\mathbf{A}_{\boldsymbol{\nu}}^*} dt\,.$$

Greedy control

As we have $1-1$ correspondence

$$\varphi^0_{\boldsymbol{\nu}} \longleftrightarrow u_\nu$$

it is sufficient to get a good approximation of the manifold $\varphi^0(\mathcal{N})$:

$$\nu \in \mathcal{N} \to \varphi^0_\nu \in \mathbf{R}^{\mathbf{N}}.$$

Thus our problem can be formulated as:

The greedy control problem

Given $\varepsilon > 0$ determine a *small* family of parameters $\nu_1, ..., , \nu_n$ in $\mathcal{N}$ so that the corresponding minimisers $\varphi^0_1, ..., \varphi^0_n$, are such that for every $\nu \in \mathcal{N}$ there exists $\varphi^{0*}_\nu \in \mathrm{span}\{\varphi^0_1, ..., \varphi^0_n\}$ satisfying

$$||\varphi^0_\nu - \varphi^{0*}_\nu|| \leq \varepsilon.$$

In order to achieve this goal we rely on greedy algorithms and reduced bases methods for parameter dependent PDEs or abstract equations in Banach spaces.

- 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

The pure greedy method

X – a Banach space
$K \subset X$ – a compact subset.

The method approximates K by a a series of finite dimensional linear spaces V_n (a linear method).

The algorithm

The first step

Choose $x_1 \in K$ such that

$$\|x_1\|_X = \max_{x \in K} \|x\|_X .$$

The general step

Having found $x_1..x_n$, denote $V_n = \mathrm{span}\{x_1, \ldots, x_n\}$.
Choose the next element

$$x_{n+1} := \arg\max_{x \in K} \mathrm{dist}(x, V_n)\,. \tag{2}$$

The algorithm stops

when $\sigma_n(K) := \max_{x \in K} \mathrm{dist}(x, V_n)$ becomes less than the given tolerance ε.

The greedy idea

The greedy idea

Which one you are going to choose?

Sometimes it is hard to solve the maximisation problem (2).

The weak greedy method

– a relaxed version of the pure one.

The algorithm

Fix a constant $\gamma \in \langle 0, 1]$.

The first step

Choose $x_1 \in K$ such that

$$\|x_1\|_X \geq \gamma \max_{x \in K} \|x\|_X .$$

The general step

Having found $x_1..x_n$, denote $V_n = \mathrm{span}\{x_1, \ldots, x_n\}$.
Choose the next element

$$\mathrm{dist}(x_{n+1}, V_n) \geq \gamma \max_{x \in K} \mathrm{dist}(x, V_n) . \tag{3}$$

The algorithm stops

when $\sigma_n(K) := \max_{x \in K} \mathrm{dist}(x, V_n)$ becomes less than the given tolerance ε.

Efficiency

In order to estimate **the efficiency of the (weak) greedy algorithm** we compare its approximation rates $\sigma_n(K)$ with the best possible one.

The Kolmogorov n width, $d_n(K)$

– measures how well K can be approximated by a subspace in X of a fixed dimension n.

$$d_n(K) := \inf_{\dim Y = n} \sup_{x \in K} \inf_{y \in Y} \|x - y\|_X .$$

Thus $d_n(K)$ represents optimal approximation performance that can be obtained by a n-dimensional linear space.
The greedy approximation rates have same decay as the Kolmogorov widths.

Theorem

(Cohen, DeVore '15) [3]
For any $\alpha > 0, C_0 > 0$

$$d_n(K) \le C_0 n^{-\alpha} \quad \Longrightarrow \quad \sigma_n(K) \le C_1 n^{-\alpha}, \quad k \in \mathbf{N},$$

where $C_1 := C_1(\alpha, C_0, \gamma)$.

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

Performance obstacles

- The set K in general consists of infinitely many vectors.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).

Performance obstacles

- The set K in general consists of infinitely many vectors.
  Finite discretisation of K.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).

Performance obstacles

- The set K in general consists of infinitely many vectors.
  Finite discretisation of K.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).
  One uses some **surrogate** value of an uniformly equivalent norm instead of the exact distance appearing in (3).

Performance obstacles

- The set K in general consists of infinitely many vectors.
  Finite discretisation of K.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).
  One uses some **surrogate** value of an uniformly equivalent norm instead of the exact distance appearing in (3).

Practical realisation depends crucially on an existence of an appropriate surrogate .

The vectors chosen by the greedy procedure are the snapshots.

Their computation can be time consuming and computational expensive (offline part).

Los Alamos National Laboratory

The vectors chosen by the greedy procedure are the snapshots.

Their computation can be time consuming and computational expensive (offline part).

Los Alamos National Laboratory

Once having chosen the snapshots, one should easily approximate any value $x \in K$ (online part).

The vectors chosen by the greedy procedure are the snapshots.

Their computation can be time consuming and computational expensive (offline part).

Los Alamos National Laboratory

Once having chosen the snapshots, one should easily approximate any value $x \in K$ (online part).

Greedy control

Recall

The greedy control problem

Given $\varepsilon > 0$ determine a *small* family of parameters $\nu_1, ..., , \nu_n$ in $\mathcal{N}$ so that the corresponding minimisers $\varphi_1^0, ..., \varphi_n^0$ are such that for every $\nu \in \mathcal{N}$ there exists $\varphi_\nu^{0*} \in \mathrm{span}\{\varphi_1^0, ..., \varphi_n^0\}$ satisfying

$$||\varphi_\nu^0 - \varphi_\nu^{0*}|| \leq \varepsilon.$$

The greedy method choose the next snapshot by maximising

$$\mathrm{dist}_{\nu\in\mathcal{N}}(\varphi_{\boldsymbol{\nu}}^0, \Phi_n^0),$$

where $\Phi_n^0 = \mathrm{span}\{\varphi_1^0, \ldots, \varphi_n^0\}$.
Thus one would have to find $\varphi_{\boldsymbol{\nu}}^0$ for every $\nu \in \mathcal{N}$, what is exactly what we want to avoid.

One has to find an appropriate surrogate!

Surrogate choice

Suppose we have chosen φ_1^0. How should we estimate $\mathrm{dist}_{\nu\in\mathcal{N}}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0)$, without knowing $\varphi_{\boldsymbol{\nu}}^0$?

As

$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_{\boldsymbol{\nu}}^0 = \mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0, \tag{4}$$

try φ_1^0 as the solution to (4), i.e. compute

$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0 \qquad\qquad\qquad .$$

Surrogate choice

Suppose we have chosen φ_1^0. How should we estimate $\mathrm{dist}_{\nu \in \mathcal{N}}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0)$, without knowing $\varphi_{\boldsymbol{\nu}}^0$?

As

$$\mathbf{G}_{\boldsymbol{\nu}} \varphi_{\boldsymbol{\nu}}^0 = \mathsf{x}^T - e^{T \mathbf{A}_{\boldsymbol{\nu}}} \mathbf{x}^0, \tag{4}$$

try φ_1^0 as the solution to (4), i.e. compute and check its distance from the target

$$\mathbf{G}_{\boldsymbol{\nu}} \varphi_1^0 - \left(\mathsf{x}^T - e^{T \mathbf{A}_{\boldsymbol{\nu}}} \mathbf{x}^0\right).$$

Surrogate choice

Suppose we have chosen φ_1^0. How should we estimate $\mathrm{dist}_{\nu\in\mathcal{N}}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0)$, without knowing $\varphi_{\boldsymbol{\nu}}^0$?

As

$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_{\boldsymbol{\nu}}^0 = \mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0, \tag{4}$$

try φ_1^0 as the solution to (4), i.e. compute and check its distance from the target

$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0 - \left(\mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0\right).$$

Thus

$$\mathrm{dist}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0) \sim \mathrm{dist}(\mathbf{G}_{\boldsymbol{\nu}}\varphi_{\boldsymbol{\nu}}^0, \mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0) = \underbrace{\mathrm{dist}(\mathsf{x}^T - e^{-T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0, \mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0)}_{surrogate}.$$

Surrogate choice

Suppose we have chosen φ_1^0. How should we estimate $\mathrm{dist}_{\nu\in\mathcal{N}}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0)$, without knowing $\varphi_{\boldsymbol{\nu}}^0$?

As

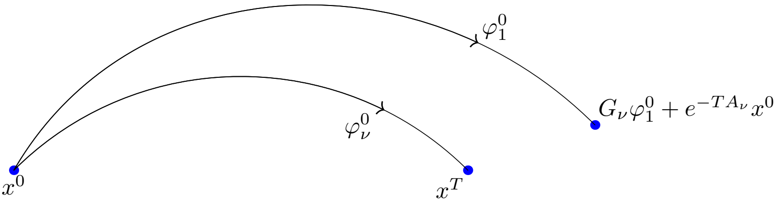
$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_{\boldsymbol{\nu}}^0 = \mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0, \tag{4}$$

try φ_1^0 as the solution to (4), i.e. compute and check its distance from the target

$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0 - \left(\mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0\right).$$

Thus

$$\mathrm{dist}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0) \sim \mathrm{dist}(\mathbf{G}_{\boldsymbol{\nu}}\varphi_{\boldsymbol{\nu}}^0, \mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0) = \underbrace{\mathrm{dist}(\mathsf{x}^T - e^{-T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0, \mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0)}_{surrogate}.$$



Surrogate choice

Suppose we have chosen φ_1^0. How should we estimate $\mathrm{dist}_{\nu\in\mathcal{N}}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0)$, without knowing $\varphi_{\boldsymbol{\nu}}^0$?

As

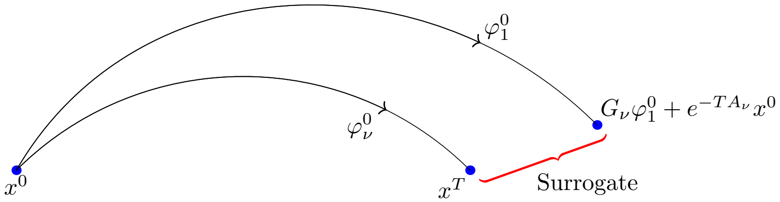
$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_{\boldsymbol{\nu}}^0 = \mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0, \tag{4}$$

try φ_1^0 as the solution to (4), i.e. compute and check its distance from the target

$$\mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0 - (\mathsf{x}^T - e^{T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0)\,.$$

Thus

$$\mathrm{dist}(\varphi_{\boldsymbol{\nu}}^0, \varphi_1^0) \sim \mathrm{dist}(\mathbf{G}_{\boldsymbol{\nu}}\varphi_{\boldsymbol{\nu}}^0, \mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0) = \underbrace{\mathrm{dist}(\mathsf{x}^T - e^{-T\mathbf{A}_{\boldsymbol{\nu}}}\mathbf{x}^0, \mathbf{G}_{\boldsymbol{\nu}}\varphi_1^0)}_{surrogate}\,.$$



Construction of the approximating space
Offline part

As the second snapshot, we choose the value for which φ_1^0 gives the worst approximation.
And so on ...

Theorem

The algorithm stops after the most $n_0 \leq N$ steps, and it fulfils the requirements of the weak greedy theory.

Corollary

The greedy control algorithm leads to an optimal approximation method. More precisely, for all $\alpha > 0$ there exists $C_\alpha > 0$ such that for any ν the minimiser ϕ_ν^0 can be approximated by linear combinations of the weak-greedy ones as follows:

$$dist(\phi_\nu^0; \mathrm{span}\{\phi_j^0 : j = 1, ..., n\}) \leq C_\alpha n^{-\alpha}.$$

The same applies when $\mathcal{N}$ is infinite-dimensional provided its Kolmogorov width decays polynomially.

Construction of the approximating control for a given parameter value
Online part

Having constructed the approximating space Φ_n^0 how do we construct an approximative control $\mathsf{u}^*_{\boldsymbol{\nu}}$ associated to an an arbitrary given value $\boldsymbol{\nu} \in \mathcal{N}$. The exact control is given by

$$\mathsf{u}_{\boldsymbol{\nu}} = B^* e^{-(T-t)\mathbf{A}^*_{\boldsymbol{\nu}}} \varphi^0_{\nu},$$

We construct the approximative one as

$$\mathsf{u}^*_{\boldsymbol{\nu}} = B^* e^{-(T-t)\mathbf{A}^*_{\boldsymbol{\nu}}} \sum_i^k \lambda_i \varphi^0_i, \tag{5}$$

where the coefficients λ_i are determined by the projection of the vector $\mathbf{G}_{\boldsymbol{\nu}} \phi^0_{\nu} = \mathsf{x}^T - e^{-T\mathbf{A}_{\boldsymbol{\nu}}} \mathbf{x}^0$ to the space $\mathbf{G}_{\boldsymbol{\nu}} \Phi^0_n = \mathrm{span}\{\mathbf{G}_{\boldsymbol{\nu}} \varphi^0_1, \ldots, \mathbf{G}_{\boldsymbol{\nu}} \varphi^0_n\}$.

N.B.

$\mathsf{u}^*_{\boldsymbol{\nu}} \notin \mathrm{span}\{u_1, ..., u_n\}$

The first example

We consider the system

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

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & -I \\ \nu(N/2+1)^2\tilde{\mathbf{A}} & \mathbf{0} \end{pmatrix},$$

$$\tilde{\mathbf{A}} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

The system corresponds to the discretisation of the wave equation problem with the control on the right boundary:

$$\begin{cases} \partial_{tt}v - \nu\partial_{xx}v = 0, \qquad (t,x) \in \langle 0,T\rangle \times \langle 0,1\rangle \\ \quad v(t,0) = 0, \quad v(t,1) = u(t) \\ v(0,x) = v_0, \quad \partial_t v(x,0) = v_1\,. \end{cases}$$

The wave equation

We take the following values:

$$T = 3, \ N = 20$$

$$v_0 = \sin(\pi x), \ v_1 = 0$$
$$x^T = 0$$

and we assume

$$\nu \in [1, 10] = \mathcal{N}$$

The system satisfies the Kalman's rank condition for any ν.

The greedy control has been applied with $\varepsilon = 0.5$ and the uniform discretisation of $\mathcal{N}$ in $k = 100$ values.

The offline algorithm stopped after 10 iterations.
10 values (out of 100) were chosen in the following order:

$$10.00 \quad 1.45 \quad 2.44 \quad 6.85 \quad 7.48 \quad 4.51 \quad 1.27 \quad 2.71 \quad 4.87 \quad 1.09$$

The corresponding minimisers have been calculated and saved.

The online part should us give an approximative control for any $\nu \in [1, 10]$.

Let us try!

Results

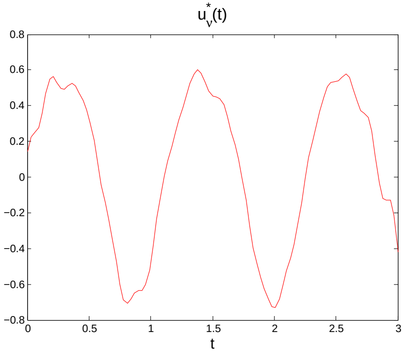
Please, choose a value between 1 and 10.

Results

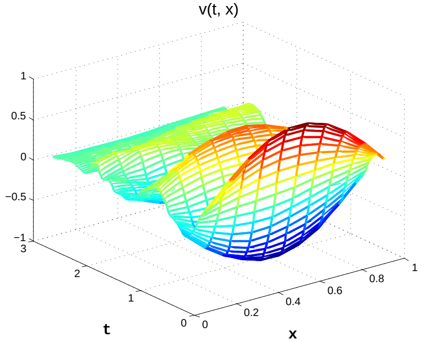
Please, choose a value between 1 and 10.
But please choose $\nu = \pi$.

Results

Please, choose a value between 1 and 10.
But please choose $\nu = \pi$.



The approximate control $u_\nu^\star$



The evolution of the solution
to the semi-discretised wave problem
for $\nu = \pi$.

Open problems and perspectives

- Our work can be extended to systems with the control operator and/or initial conditions depending on the parameter as well.
  Of special interest is the affine dependence case

  $$\mathbf{B}(\nu) = \bar{\mathbf{B}} + \sum_j \nu_j \mathbf{B}_j$$

  with $(\|\mathbf{B}_j\|) \in l^p$ for some $p \leq 1$ which should significantly reduce the computational cost.
- Our work can be extended to PDE but analyticity of controls with respect to parameters has to be ensured. This typically holds for elliptic and parabolic equations. But not for wave-like equations.
  Indeed, solutions of

  $$y_{tt} - \nu^2 y_{xx} = 0$$

  do not depend analytically on the coefficient ν.
- Alternative surrogates need to found so to make the recursive choice process of the various $\nu' s$ (offline part) faster and cheaper.

Open problems and perspectives

- Our work can be extended to systems with the control operator and/or initial conditions depending on the parameter as well.
  Of special interest is the affine dependence case

  $$\mathbf{B}(\nu) = \bar{\mathbf{B}} + \sum_j \nu_j \mathbf{B}_j$$

  with $(\|\mathbf{B}_j\|) \in l^p$ for some $p \leq 1$ which should significantly reduce the computational cost.
- Our work can be extended to PDE but analyticity of controls with respect to parameters has to be ensured. This typically holds for elliptic and parabolic equations. But not for wave-like equations.
  Indeed, solutions of

  $$y_{tt} - \nu^2 y_{xx} = 0$$

  do not depend analytically on the coefficient ν.
- Alternative surrogates need to found so to make the recursive choice process of the various $\nu' s$ (offline part) faster and cheaper.

Thanks for your attention!