Optimal control of ODEs with dynamics uncertainty

Alessandro Scagliotti

TUM, Munich
MCML – Munich Center for Machine Learning

INRIA – McTAO Group Seminar

Ensembles of control systems: motivations

Ensembles of control systems: motivations

Let us consider the following model for chemotherapy:

$$\begin{cases} \dot{x}_1 = \xi_1 x_1 \left(1 - \frac{x_1 + x_2}{M}\right) - \mu u x_1 & \text{(sensitive population)} \\ \dot{x}_2 = \xi_2 x_2 \left(1 - \frac{x_1 + x_2}{M}\right) & \text{(resistant population)} \end{cases}$$

where

- $u \in [0, u_{\max}]$ is the control;

Ensembles of control systems: motivations

Let us consider the following model for chemotherapy:

$$\begin{cases} \dot{x}_1 = \xi_1 x_1 \left(1 - \frac{x_1 + x_2}{M}\right) - \mu u x_1 & \text{(sensitive population)} \\ \dot{x}_2 = \xi_2 x_2 \left(1 - \frac{x_1 + x_2}{M}\right) & \text{(resistant population)} \end{cases}$$

where

- $u \in [0, u_{\max}]$ is the control;
- $\xi_1, \xi_2, M, \mu > 0$ are **unknown** parameters.

Ensembles of control systems: motivations

Let us consider the following model for chemotherapy:

$$\begin{cases} \dot{x}_1 = \xi_1 x_1 \left(1 - \frac{x_1 + x_2}{M}\right) - \mu u x_1 & \text{(sensitive population)} \\ \dot{x}_2 = \xi_2 x_2 \left(1 - \frac{x_1 + x_2}{M}\right) & \text{(resistant population)} \end{cases}$$

where

- $u \in [0, u_{\max}]$ is the control;
- $\xi_1, \xi_2, M, \mu > 0$ are **unknown** parameters.

Classical strategy in medicine

Maximal dose: $u(t) \equiv u_{\max}$ until the tumor starts growing again.

Ensembles of control systems: motivations

Let us consider the following model for chemotherapy:

$$\begin{cases} \dot{x}_1 = \xi_1 x_1 \left(1 - \frac{x_1+x_2}{M}\right) - \mu u x_1 & \text{(sensitive population)} \\ \dot{x}_2 = \xi_2 x_2 \left(1 - \frac{x_1+x_2}{M}\right) & \text{(resistant population)} \end{cases}$$

where

- $u \in [0, u_{\max}]$ is the control;
- $\xi_1, \xi_2, M, \mu > 0$ are **unknown** parameters.

Classical strategy in medicine

Maximal dose: $u(t) \equiv u_{\max}$ until the tumor starts growing again.
Then, when possible, change drug (2nd line treatment) and use it at the maximal dose.

Ensembles of control systems: motivations

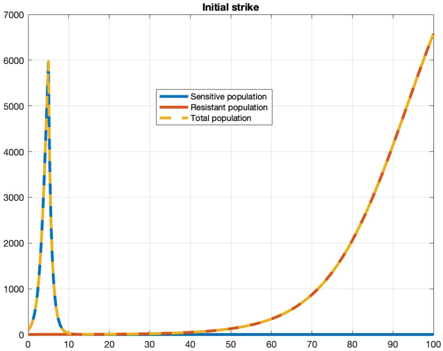

Figure: Strategy $u(t) \equiv u_{\max}$. The sensitive population is rapidly extincted by the treatment. After some time, a resistant tumor returns.

Ensembles of control systems: motivations

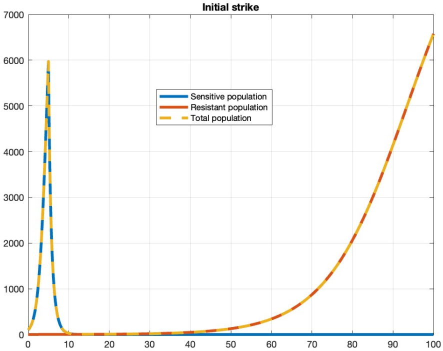

Figure: Strategy $u(t) \equiv u_{\max}$. The sensitive population is rapidly extincted by the treatment. After some time, a resistant tumor returns.

This strategy does not require the knowledge of ξ_1, ξ_2, M, μ.

Ensembles of control systems: motivations

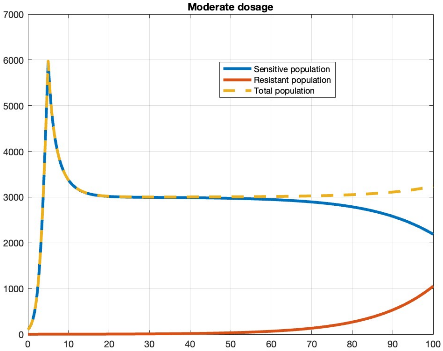

Figure: Strategy $u(t) \equiv \bar{u} < u_{\max}$. The tumor never disappears, but it is stabilized. The sensitive cells are delaying the growth of the resistant population.

Ensembles of control systems: motivations

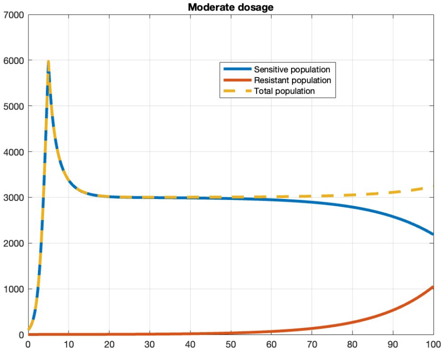

Figure: Strategy $u(t) \equiv \bar{u} < u_{\max}$. The tumor never disappears, but it is stabilized. The sensitive cells are delaying the growth of the resistant population.

This strategy **depends on** ξ_1, ξ_2, M, μ.

Ensembles of control systems: motivations

Control of a qubit (Schrödinger equation):

$$i\frac{d\psi}{dt} = \begin{pmatrix} E+\alpha & u(t) \\ u(t) & -E-\alpha \end{pmatrix}\psi,$$

where α represents the uncertainty affecting the *resonance frequency*.

Ensembles of control systems: motivations

Control of a qubit (Schrödinger equation):

$$i\frac{d\psi}{dt} = \begin{pmatrix} E + \alpha & u(t) \\ u(t) & -E - \alpha \end{pmatrix} \psi,$$

where α represents the uncertainty affecting the *resonance frequency*.

In [Robin, Augier, et al., J.Diff.Eq., 2022] a strategy for *uniform ensemble controllability* is proposed (steer $\psi(0) = (0,1)^T$ to the target $(1,0)^T$). They can do the job when $\alpha \in [\alpha_{\min}, \alpha_{\max}] \subset (-0.5, 0.5)$.

Ensembles of control systems: motivations

Control of a qubit (Schrödinger equation):

$$i\frac{d\psi}{dt} = \begin{pmatrix} E+\alpha & u(t) \\ u(t) & -E-\alpha \end{pmatrix} \psi,$$

where α represents the uncertainty affecting the *resonance frequency*.

In [Robin, Augier, et al., J.Diff.Eq., 2022] a strategy for *uniform ensemble controllability* is proposed (steer $\psi(0) = (0,1)^T$ to the target $(1,0)^T$). They can do the job when $\alpha \in [\alpha_{\text{min}}, \alpha_{\text{max}}] \subset (-0.5, 0.5)$.

Is it possible to find *optimal strategies*?
Optimal could be *on average* on the ensemble, or *uniformly*.

Constructed vs. optimal controls: shape

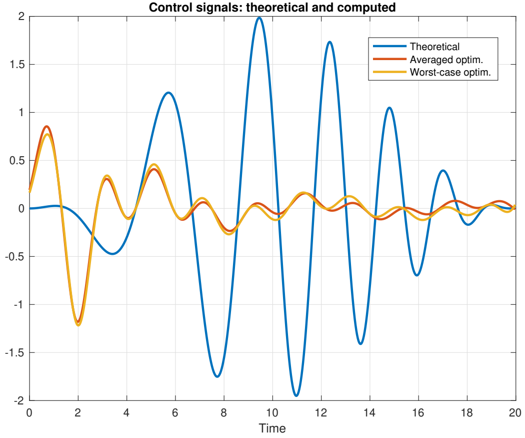

Figure: Comparison between the controls used for the proof of controllability, and the computed optimal controls.

Constructed vs. optimal controls: performances

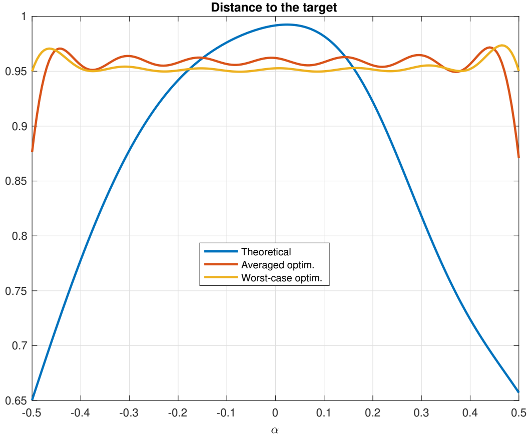

Figure: *x*-axis: Value of the unknown parameter α (resonance frequency). *y*-axis: Distance to the target state $(1, 0)^T$.

Constructed vs. optimal controls: shape

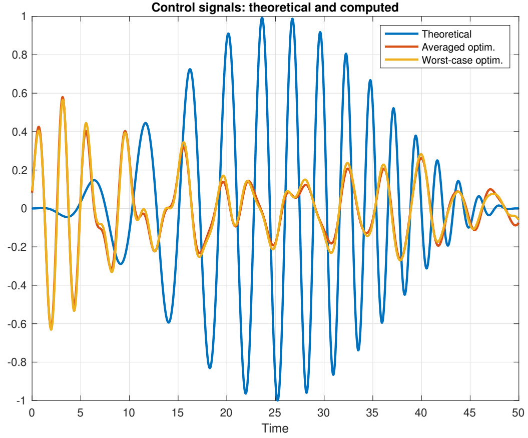

Figure: Comparison between the controls used for the proof of controllability, and the computed optimal controls.

Constructed vs. optimal controls: performances

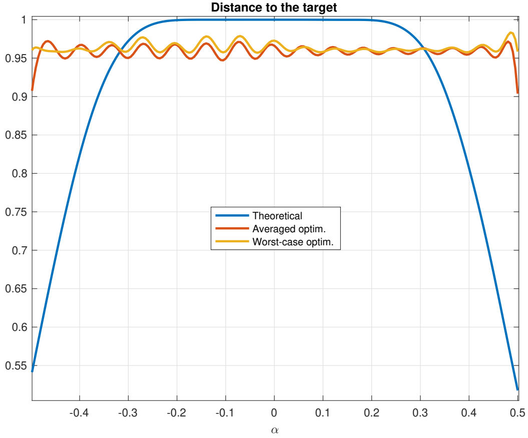

Figure: x-axis: Value of the unknown parameter α (resonance frequency). y-axis: Distance to the target state $(1, 0)^T$.

Ensembles of affine-control systems

Ingredients

- Compact set of parameters $\Theta \subset \mathbb{R}^d$;
- Dynamics in $\mathbb{R}^n$ on the time interval $[0, T]$:

$$\dot{x}^\theta = b^\theta(x) + A^\theta(x)u, \quad x^\theta(0) = x_0^\theta;$$

- a *simultaneous* control $u \in \mathcal{U} := L^p([0, T], \mathbb{R}^m)$, $1 < p < \infty$.

Ensembles of affine-control systems

Ingredients

- Compact set of parameters $\Theta \subset \mathbb{R}^d$;
- Dynamics in $\mathbb{R}^n$ on the time interval $[0, T]$:

$$\dot{x}^\theta = b^\theta(x) + A^\theta(x)u, \quad x^\theta(0) = x_0^\theta;$$

- a *simultaneous* control $u \in \mathcal{U} := L^p([0, T], \mathbb{R}^m)$, $1 < p < \infty$.

Aim: Modelling data uncertainty in the dynamics and/or the initial datum.

Ensembles of affine-control systems

Ingredients

- Compact set of parameters $\Theta \subset \mathbb{R}^d$;
- Dynamics in $\mathbb{R}^n$ on the time interval $[0, T]$:

$$\dot{x}^\theta = b^\theta(x) + A^\theta(x)u, \quad x^\theta(0) = x_0^\theta;$$

- a *simultaneous* control $u \in \mathcal{U} := L^p([0, T], \mathbb{R}^m)$, $1 < p < \infty$.

Aim: Modelling data uncertainty in the dynamics and/or the initial datum.

Technical assumptions

- $(x, \theta) \mapsto b^\theta(x) \in \mathbb{R}^n$, $(x, \theta) \mapsto A^\theta(x) \in \mathbb{R}^{n \times m}$ Lipschitz-continuous;
- $\theta \mapsto x_0^\theta$ is Lipschitz-continuous.

Optimal control of ensembles: weighted problems

θ-specific problem

For every $\theta \in \Theta$, we would like to solve

$$\ell^\theta(x_u^\theta(T)) + \beta \int_0^T f(u(s))\, ds \to \min,$$

with $\beta > 0$, and where

- $\ell^\theta : \mathbb{R}^n \to \mathbb{R}_+$ is the end-point cost ($\ell : \mathbb{R}^n \times \Theta \to \mathbb{R}_+$ continuous)

Optimal control of ensembles: weighted problems

θ-specific problem

For every $\theta \in \Theta$, we would like to solve

$$\ell^\theta(x_u^\theta(T)) + \beta \int_0^T f(u(s))\, ds \to \min,$$

with $\beta > 0$, and where

- $\ell^\theta : \mathbb{R}^n \to \mathbb{R}_+$ is the end-point cost ($\ell : \mathbb{R}^n \times \Theta \to \mathbb{R}_+$ continuous);
- $f : \mathbb{R}^m \to \mathbb{R}$ convex, continuous, and $f(u) \geq c(1 + |u|_2^p)$.

Optimal control of ensembles: weighted problems

θ-specific problem

For every $\theta \in \Theta$, we would like to solve

$$\ell^\theta(x_u^\theta(T)) + \beta \int_0^T f(u(s))\, ds \to \min,$$

with $\beta > 0$, and where

- $\ell^\theta : \mathbb{R}^n \to \mathbb{R}_+$ is the end-point cost ($\ell : \mathbb{R}^n \times \Theta \to \mathbb{R}_+$ continuous);
- $f : \mathbb{R}^m \to \mathbb{R}$ convex, continuous, and $f(u) \geq c(1 + |u|_2^p)$.

Idea

We use a probability measure $\mu \in \mathcal{P}(\Theta)$ to *describe our knowledge on* θ:

$$\mathcal{G}_\mu(u) = \int_\Theta \ell(x_u^\theta(T), \theta)\, d\mu(\theta) + \beta \int_0^T f(u(s))\, ds \to \min.$$

Existence of minimizers

Let us consider

$$\mathcal{G}_\mu(u) = \int_\Theta \ell(x_u^\theta(T), \theta)\, d\mu(\theta) + \beta \int_0^T f(u(s))\, ds.$$

Proposition

There exists $\hat{u} \in \mathcal{U}$ *such that*

$$\mathcal{G}_\mu(\hat{u}) = \inf_{\mathcal{U}} \mathcal{G}_\mu.$$

Moreover, for every $\hat{u} \in \arg\min \mathcal{G}_\mu$*, we have* $\|\hat{u}\|_{L^p} \leq C(\beta)$*.*

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{G}_\mu$ requires:

- the exact knowledge of the probability measure μ;

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{G}_\mu$ requires:

- the exact knowledge of the probability measure μ;
- handling simultaneously an infinite number of control systems.

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{G}_\mu$ requires:

- the exact knowledge of the probability measure μ;
- handling simultaneously an infinite number of control systems.

Data-driven approach: In practice, we may have access to empirical measurements $\theta^1, \ldots, \theta^M \in \Theta$, independently sampled from μ.
We define

$$\mu^M := \frac{1}{M} \sum_{j=1}^{M} \delta_{\theta^j}$$

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{G}_\mu$ requires:

- the exact knowledge of the probability measure μ;
- handling simultaneously an infinite number of control systems.

Data-driven approach: In practice, we may have access to empirical measurements $\theta^1, \dots, \theta^M \in \Theta$, independently sampled from μ.
We define

$$\mu^M := \frac{1}{M} \sum_{j=1}^{M} \delta_{\theta^j},$$

and we assume that $\lim_{M\to\infty} \langle \mu^M, \phi \rangle = \langle \mu, \phi \rangle, \quad \forall \phi \in C_b(\Theta)$,
i.e., $\mu^M \rightharpoonup^* \mu$ as $M \to \infty$.

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{G}_\mu$ requires:

- the exact knowledge of the probability measure μ;
- handling simultaneously an infinite number of control systems.

Data-driven approach: In practice, we may have access to empirical measurements $\theta^1, \ldots, \theta^M \in \Theta$, independently sampled from μ.
We define

$$\mu^M := \frac{1}{M} \sum_{j=1}^{M} \delta_{\theta^j},$$

and we assume that $\lim_{M\to\infty} \langle \mu^M, \phi \rangle = \langle \mu, \phi \rangle, \quad \forall \phi \in C_b(\Theta)$,
i.e., $\mu^M \rightharpoonup^* \mu$ as $M \to \infty$.

Similar setup as in **supervised ML**!

Reduction to finite ensembles: Γ-convergence

For every $M \geq 1$, we consider $\mathcal{G}_{\mu^M} : \mathcal{U} \to \mathbb{R}$ defined as

$$\begin{aligned}\mathcal{G}_{\mu^M}(u) &= \sum_{j=1}^{M} \frac{1}{M} \ell\big(x_u^{\theta_j}(T), \theta_j\big) + \beta \int_0^T f(u(s))\, ds \\ &= \int_\Theta \ell\big(x_u^{\theta}(T), \theta\big)\, d\mu^M(\theta) + \beta \int_0^T f(u(s))\, ds.\end{aligned}$$

Reduction to finite ensembles: Γ-convergence

For every $M \geq 1$, we consider $\mathcal{G}_{\mu^M} : \mathcal{U} \to \mathbb{R}$ defined as

$$\begin{aligned} \mathcal{G}_{\mu^M}(u) &= \sum_{j=1}^{M} \frac{1}{M} \ell\big(x_u^{\theta_j}(T), \theta_j\big) + \beta \int_0^T f(u(s))\, ds \\ &= \int_\Theta \ell\big(x_u^{\theta}(T), \theta\big)\, d\mu^M(\theta) + \beta \int_0^T f(u(s))\, ds. \end{aligned}$$

Theorem

*For every $M \geq 1$, let us consider $\mathcal{G}_{\mu^M} : \mathcal{U} \to \mathbb{R}$.
Then, the sequence $(\mathcal{G}_{\mu^M})_{M\geq 1}$ is Γ-convergent to the functional $\mathcal{G}_\mu : \mathcal{U} \to \mathbb{R}$ with respect to the weak topology of $\mathcal{U}$.*

Reduction to finite ensembles: Γ-convergence

For every $M \geq 1$, we consider $\mathcal{G}_{\mu^M} : \mathcal{U} \to \mathbb{R}$ defined as

$$\begin{aligned} \mathcal{G}_{\mu^M}(u) &= \sum_{j=1}^{M} \frac{1}{M} \ell\left(x_u^{\theta_j}(T), \theta_j\right) + \beta \int_0^T f(u(s))\, ds \\ &= \int_\Theta \ell\left(x_u^{\theta}(T), \theta\right) d\mu^M(\theta) + \beta \int_0^T f(u(s))\, ds. \end{aligned}$$

Theorem

For every $M \geq 1$, let us consider $\mathcal{G}_{\mu^M} : \mathcal{U} \to \mathbb{R}$.
Then, the sequence $(\mathcal{G}_{\mu^M})_{M\geq 1}$ is Γ-convergent to the functional $\mathcal{G}_\mu : \mathcal{U} \to \mathbb{R}$ with respect to the weak topology of $\mathcal{U}$.

Remark

Here the fact that the systems are affine in the control is crucial!

Γ-convergence: consequences

Convergence of minima.

$$\min_{u\in\mathcal{U}} \mathcal{G}_{\mu^M} \to \min_{u\in\mathcal{U}} \mathcal{G}_{\mu} \quad \text{as } M \to \infty.$$

Γ-convergence: consequences

Convergence of minima.

$$\min_{u\in\mathcal{U}} \mathcal{G}_{\mu^M} \to \min_{u\in\mathcal{U}} \mathcal{G}_{\mu} \quad \text{as } M \to \infty.$$

Convergence of minimizers. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{G}_{\mu^M}$.
Then $(\hat{u}^M)$ is pre-compact in the **weak topology** of L^p, and clusters are minimizers of $\mathcal{G}_{\mu}$.

Γ-convergence: consequences

Convergence of minima.

$$\min_{u\in\mathcal{U}} \mathcal{G}_{\mu^M} \to \min_{u\in\mathcal{U}} \mathcal{G}_{\mu} \quad \text{as } M \to \infty.$$

Convergence of minimizers. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{G}_{\mu^M}$.
Then $(\hat{u}^M)$ is pre-compact in the **weak topology** of L^p, and clusters are minimizers of $\mathcal{G}_\mu$.

Convergence of integral costs. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{G}_{\mu^M}$, and assume that $\hat{u}^M \rightharpoonup \hat{u}$. Then,

$$\lim_{M\to\infty} \int_\Theta a(x^\theta_{\hat{u}^M}(T), \theta)\, d\mu^M(\theta) = \int_\Theta a(x^\theta_{\hat{u}}(T), \theta)\, d\mu(\theta),$$

$$\lim_{M\to\infty} \int_0^T f(\hat{u}^M(s))\, ds = \int_0^T f(\hat{u}(s))\, ds.$$

Γ-convergence: consequences – improved

Convergence of minima.

$$\min_{u\in\mathcal{U}} \mathcal{G}_{\mu^M} \to \min_{u\in\mathcal{U}} \mathcal{G}_{\mu} \quad \text{as } M \to \infty.$$

Convergence of minimizers. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{G}_{\mu^M}$.
Then, if f is **strictly convex**, $(\hat{u}^M)$ is pre-compact in the **strong topology** of L^p, and clusters are minimizers of $\mathcal{G}_\mu$.

Convergence of integral costs. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{G}_{\mu^M}$, and assume that $\hat{u}^M \rightharpoonup \hat{u}$. Then,

$$\lim_{M\to\infty} \sum_{j=1}^{M} \frac{1}{M} \ell\big(x_{\hat{u}^M}^{\theta_j}(T), \theta_j\big) = \int_\Theta \ell(x_{\hat{u}}^{\theta}(T), \theta)\, d\mu(\theta),$$

$$\lim_{M\to\infty} \int_0^T f(\hat{u}^M(s))\, ds = \int_0^T f(\hat{u}(s))\, ds.$$

Γ-convergence and PMP: preliminaries

Idea

Try to use Γ-convergence to establish PMP for infinite ensembles.
When considering μ^M, the problem reduces to a control system in $(\mathbb{R}^n)^M$.

Γ-convergence and PMP: preliminaries

Idea

Try to use Γ-convergence to establish PMP for infinite ensembles.
When considering μ^M, the problem reduces to a control system in $(\mathbb{R}^n)^M$.

Notations

For every $u \in \mathcal{U}$, we define $X_u : [0, T] \times \Theta \to \mathbb{R}^n$ as

$$X_u(t, \theta) := x_u^\theta(t),$$

Γ-convergence and PMP: preliminaries

Idea

Try to use Γ-convergence to establish PMP for infinite ensembles.
When considering μ^M, the problem reduces to a control system in $(\mathbb{R}^n)^M$.

Notations

For every $u \in \mathcal{U}$, we define $X_u : [0, T] \times \Theta \to \mathbb{R}^n$ as

$$X_u(t, \theta) := x_u^\theta(t),$$

and $\Lambda_u : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ (**analogous to back-propagation!**)

Γ-convergence and PMP: preliminaries

Idea

Try to use Γ-convergence to establish PMP for infinite ensembles.
When considering μ^M, the problem reduces to a control system in $(\mathbb{R}^n)^M$.

Notations

For every $u \in \mathcal{U}$, we define $X_u : [0, T] \times \Theta \to \mathbb{R}^n$ as

$$X_u(t, \theta) := x_u^\theta(t),$$

and $\Lambda_u : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ (**analogous to back-propagation!**) as

$$\Lambda_u(t, \theta) := \lambda_u^\theta(t),$$

where

$$\dot{\lambda}_u^\theta = -\lambda_u^\theta \cdot \frac{\partial}{\partial x}\left(b^\theta(x_u) - A^\theta(x_u)u\right), \quad \lambda_u^\theta(T) = -\nabla_x \ell(x_u(T), \theta).$$

Γ-convergence and PMP: finite ensembles

If $\#\mathrm{supp}(\mu^M) = M$, then we have a problem in $(\mathbb{R}^n)^M$.

Γ-convergence and PMP: finite ensembles

If $\#\mathrm{supp}(\mu^M) = M$, then we have a problem in $(\mathbb{R}^n)^M$.

Theorem

Let $\hat{u}^M \in \arg\min \mathcal{G}_{\mu^M}$. *Then, considering* $X_{\hat{u}^M} : [0, T] \times \Theta \to \mathbb{R}^n$ *and* $\Lambda_{\hat{u}^M} : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ *as before, we have that*

$$\hat{u}^M(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}^M}(t, \theta) \cdot A^\theta \big(X_{\hat{u}^M}(t, \theta) \big) v \, d\mu^M(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0, T]$.

Γ-convergence and PMP: finite ensembles

If $\#\mathrm{supp}(\mu^M) = M$, then we have a problem in $(\mathbb{R}^n)^M$.

Theorem

Let $\hat{u}^M \in \arg\min \mathcal{G}_{\mu^M}$. Then, considering $X_{\hat{u}^M} : [0, T] \times \Theta \to \mathbb{R}^n$ and $\Lambda_{\hat{u}^M} : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ as before, we have that

$$\hat{u}^M(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}^M}(t, \theta) \cdot A^\theta \big(X_{\hat{u}^M}(t, \theta) \big) v \, d\mu^M(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0, T]$.

Remark

No need for computing $X_{\hat{u}^M}(t, \theta), \Lambda_{\hat{u}^M}(t, \theta)$ for every $\theta \in \Theta$.
Sufficient for $\theta \in \mathrm{supp}(\mu^M)$.

Γ-convergence and PMP: idea

Consider $\hat{u}^M \in \arg\min \mathcal{G}_{\mu^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u}$

Γ-convergence and PMP: idea

Consider $\hat{u}^M \in \arg\min \mathcal{G}_{\mu^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u} \in \arg\min \mathcal{G}_{\mu}$. We have

$$\begin{cases} \mu^M \rightharpoonup^* \mu \\ X_{\hat{u}^M} \to_{C^0} X_{\hat{u}} \\ \Lambda_{\hat{u}^M} \to_{C^0} \Lambda_{\hat{u}} \\ \hat{u}^M(t) \to \hat{u}(t) \quad \text{a.e. (up to subseq.)} \end{cases} \qquad \text{as } M \to \infty.$$

Γ-convergence and PMP: idea

Consider $\hat{u}^M \in \arg\min \mathcal{G}_{\mu^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u} \in \arg\min \mathcal{G}_\mu$. We have

$$\begin{cases} \mu^M \rightharpoonup^* \mu \\ X_{\hat{u}^M} \to_{C^0} X_{\hat{u}} \\ \Lambda_{\hat{u}^M} \to_{C^0} \Lambda_{\hat{u}} \\ \hat{u}^M(t) \to \hat{u}(t) \quad \text{a.e. (up to subseq.)} \end{cases} \qquad \text{as } M \to \infty.$$

We can pass to the limit here:

$$\hat{u}^M(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}^M}(t,\theta) \cdot A^\theta\big(X_{\hat{u}^M}(t,\theta)\big) v \, d\mu^M(\theta) - \beta f(v) \right\}$$

We use that the subdifferential ∂f has closed graph.

Γ-convergence and PMP: idea

Consider $\hat{u}^M \in \arg\min \mathcal{G}_{\mu^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u} \in \arg\min \mathcal{G}_\mu$.
We have

$$\begin{cases} \mu^M \rightharpoonup^* \mu \\ X_{\hat{u}^M} \to_{C^0} X_{\hat{u}} \\ \Lambda_{\hat{u}^M} \to_{C^0} \Lambda_{\hat{u}} \\ \hat{u}^M(t) \to \hat{u}(t) \quad \text{a.e. (up to subseq.)} \end{cases} \qquad \text{as } M \to \infty.$$

We get:

$$\hat{u}(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}}(t,\theta) \cdot A^\theta\big(X_{\hat{u}}(t,\theta)\big) v \, d\mu(\theta) - \beta f(v) \right\}$$

We use that the subdifferential ∂f has closed graph.

Γ-convergence and PMP: infinite ensembles

Using the Γ-convergence consequences, we can deduce:

Theorem (S., 2023)

Let $\hat{u} \in \arg\min \mathcal{G}_\mu$. Then, considering $X_{\hat{u}} : [0, T] \times \Theta \to \mathbb{R}^n$ and $\Lambda_{\hat{u}} : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ as before, we have that

$$\hat{u}(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}}(t, \theta) \cdot A^\theta\big(X_{\hat{u}}(t, \theta)\big) v \, d\mu(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0, T]$.

Optimal control of ensembles: minimax problems

θ-specific problem

For every $\theta \in \Theta$, we would like to solve

$$\ell^\theta(x_u^\theta(T)) + \beta \int_0^T f(u(s))\, ds \to \min,$$

with $\beta > 0$, and where

- $\ell^\theta : \mathbb{R}^n \to \mathbb{R}_+$ is the end-point cost ($\ell : \mathbb{R}^n \times \Theta \to \mathbb{R}_+$ continuous);
- $f : \mathbb{R}^m \to \mathbb{R}$ convex, continuous, and $f(u) \geq c(1 + |u|_2^p)$.

Optimal control of ensembles: minimax problems

θ-specific problem

For every $\theta \in \Theta$, we would like to solve

$$\ell^\theta(x_u^\theta(T)) + \beta \int_0^T f(u(s))\, ds \to \min,$$

with $\beta > 0$, and where

- $\ell^\theta : \mathbb{R}^n \to \mathbb{R}_+$ is the end-point cost ($\ell : \mathbb{R}^n \times \Theta \to \mathbb{R}_+$ continuous);
- $f : \mathbb{R}^m \to \mathbb{R}$ convex, continuous, and $f(u) \geq c(1 + |u|_2^p)$.

Idea

We want to do *the best in the worst scenario*:

$$\mathcal{F}_\Theta(u) = \sup_{\theta \in \Theta} \ell(x_u^\theta(T), \theta) + \beta \int_0^T f(u(s))\, ds \to \min\,.$$

Existence of minimizers

Let us consider

$$\mathcal{F}_\Theta(u) = \sup_{\theta\in\Theta} \ell(x_u^\theta(T),\theta) + \beta \int_0^T f(u(s))\, ds.$$

Proposition

There exists $\hat{u} \in \mathcal{U}$ such that

$$\mathcal{F}_\Theta(\hat{u}) = \inf_{\mathcal{U}} \mathcal{F}_\Theta.$$

Moreover, for every $\hat{u} \in \arg\min \mathcal{F}_\Theta$, we have $\|\hat{u}\|_{L^p} \leq C(\beta)$.

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{F}_\Theta$ requires:

- handling simultaneously an infinite number of control systems.

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{F}_\Theta$ requires:

- handling simultaneously an infinite number of control systems.

In practice, we may be able to construct $\Theta^M \subset \Theta$ such that:

- $\#\Theta^M < \infty$;
- $d_H(\Theta^M, \Theta) \to 0$ as $M \to \infty$ (Hausdorff distance).

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{F}_\Theta$ requires:

- handling simultaneously an infinite number of control systems.

In practice, we may be able to construct $\Theta^M \subset \Theta$ such that:

- $\#\Theta^M < \infty$;
- $d_H(\Theta^M, \Theta) \to 0$ as $M \to \infty$ (Hausdorff distance).

$$\mathcal{F}_{\Theta^M}(u) = \sup_{\theta \in \Theta^M} a(x^\theta(T), \theta) + \beta \int_0^T f(u(s))\, ds.$$

Reduction to finite ensembles

The problem of finding the minimizers of $\mathcal{F}_\Theta$ requires:

- handling simultaneously an infinite number of control systems.

In practice, we may be able to construct $\Theta^M \subset \Theta$ such that:

- $\#\Theta^M < \infty$;
- $d_H(\Theta^M, \Theta) \to 0$ as $M \to \infty$ (Hausdorff distance).

$$\mathcal{F}_{\Theta^M}(u) = \sup_{\theta \in \Theta^M} a(x^\theta(T), \theta) + \beta \int_0^T f(u(s))\, ds.$$

Theorem

*For every $M \geq 1$, let us consider $\mathcal{F}_{\Theta^M} : \mathcal{U} \to \mathbb{R}$.
Then, the sequence $(\mathcal{F}_{\Theta^M})_{M \geq 1}$ is Γ-convergent to the functional $\mathcal{F}_\Theta : \mathcal{U} \to \mathbb{R}$ with respect to the weak topology of $\mathcal{U}$.*

Γ-convergence: consequences – improved

Convergence of minima.

$$\min_{u \in \mathcal{U}} \mathcal{F}_{\Theta^M} \to \min_{u \in \mathcal{U}} \mathcal{F}_{\Theta} \quad \text{as } M \to \infty.$$

Γ-convergence: consequences – improved

Convergence of minima.

$$\min_{u\in\mathcal{U}} \mathcal{F}_{\Theta^M} \to \min_{u\in\mathcal{U}} \mathcal{F}_{\Theta} \quad \text{as } M \to \infty.$$

Convergence of minimizers. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{F}_{\Theta^M}$.
Then, if f is **strictly convex**, $(\hat{u}^M)$ is pre-compact in the **strong topology** of L^p, and clusters are minimizers of $\mathcal{F}_{\Theta}$.

Γ-convergence: consequences – improved

Convergence of minima.

$$\min_{u\in\mathcal{U}} \mathcal{F}_{\Theta^M} \to \min_{u\in\mathcal{U}} \mathcal{F}_{\Theta} \quad \text{as } M \to \infty.$$

Convergence of minimizers. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{F}_{\Theta^M}$.
Then, if f is **strictly convex**, $(\hat{u}^M)$ is pre-compact in the **strong topology** of L^p, and clusters are minimizers of $\mathcal{F}_{\Theta}$.

Convergence of integral costs. Let $\hat{u}^M \in \arg\min_{\mathcal{U}} \mathcal{F}_{\Theta}$, and assume that $\hat{u}^M \rightharpoonup \hat{u}$. Then,

$$\lim_{M\to\infty} \sup_{\theta\in\Theta^M} \ell(x^{\theta}_{\hat{u}^M}(T), \theta) = \sup_{\theta\in\Theta} \ell(x^{\theta}_{\hat{u}}(T), \theta),$$

$$\lim_{M\to\infty} \int_0^T f(\hat{u}^M(s))\, ds = \int_0^T f(\hat{u}(s))\, ds.$$

PMP for minimax: notations recap

Notations

For every $u \in \mathcal{U}$, we define $X_u : [0, T] \times \Theta \to \mathbb{R}^n$ as

$$X_u(t, \theta) := x_u^\theta(t),$$

PMP for minimax: notations recap

Notations

For every $u \in \mathcal{U}$, we define $X_u : [0, T] \times \Theta \to \mathbb{R}^n$ as

$$X_u(t, \theta) := x_u^\theta(t),$$

and $\Lambda_u : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ (**analogue to back-propagation!**) as

$$\Lambda_u(t, \theta) := \lambda_u^\theta(t),$$

where

$$\dot{\lambda}_u^\theta = -\lambda_u^\theta \cdot \frac{\partial}{\partial x}\left(b^\theta(x_u) - A^\theta(x_u)u\right), \quad \lambda_u^\theta(T) = -\nabla_x \ell(x_u(T), \theta).$$

PMP for minimax problems (finite ensembles)

If $\#\Theta^M < \infty$, then we have a minimax problem in $(\mathbb{R}^n)^{\#\Theta^M}$.

PMP for minimax problems (finite ensembles)

If $\#\Theta^M < \infty$, then we have a minimax problem in $(\mathbb{R}^n)^{\#\Theta^M}$.

Theorem (in Vinter, Minimax Optimal Control, 2005)

Let $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$. Then, considering $X_{\hat{u}^M} : [0, T] \times \Theta \to \mathbb{R}^n$ and $\Lambda_{\hat{u}^M} : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ as before, there exists a probability measure $\nu^M \in \mathcal{P}(\Theta^M)$

PMP for minimax problems (finite ensembles)

If $\#\Theta^M < \infty$, then we have a minimax problem in $(\mathbb{R}^n)^{\#\Theta^M}$.

Theorem (in Vinter, Minimax Optimal Control, 2005)

Let $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$. *Then, considering* $X_{\hat{u}^M} : [0,T] \times \Theta \to \mathbb{R}^n$ *and* $\Lambda_{\hat{u}^M} : [0,T] \times \Theta \to (\mathbb{R}^n)^*$ *as before, there exists a probability measure* $\nu^M \in \mathcal{P}(\Theta^M)$ *such that*

$$\hat{u}^M(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}^M}(t,\theta) \cdot A^\theta\big(X_{\hat{u}^M}(t,\theta)\big) v \, d\nu^M(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0,T]$

PMP for minimax problems (finite ensembles)

If $\#\Theta^M < \infty$, then we have a minimax problem in $(\mathbb{R}^n)^{\#\Theta^M}$.

Theorem (in Vinter, Minimax Optimal Control, 2005)

Let $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$. *Then, considering* $X_{\hat{u}^M} : [0,T] \times \Theta \to \mathbb{R}^n$ *and* $\Lambda_{\hat{u}^M} : [0,T] \times \Theta \to (\mathbb{R}^n)^*$ *as before, there exists a probability measure* $\nu^M \in \mathcal{P}(\Theta^M)$ *such that*

$$\hat{u}^M(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}^M}(t,\theta) \cdot A^\theta\big(X_{\hat{u}^M}(t,\theta)\big) v \, d\nu^M(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0,T]$, *and* $\bar{\theta} \in \mathrm{supp}(\nu^M) \implies \bar{\theta} \in \arg\max_{\Theta^M} a(x^\theta_{\hat{u}^M}(T), \theta)$.

PMP for minimax problems (finite ensembles)

If $\#\Theta^M < \infty$, then we have a minimax problem in $(\mathbb{R}^n)^{\#\Theta^M}$.

Theorem (in Vinter, Minimax Optimal Control, 2005)

Let $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$. Then, considering $X_{\hat{u}^M} : [0, T] \times \Theta \to \mathbb{R}^n$ and $\Lambda_{\hat{u}^M} : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ as before, there exists a probability measure $\nu^M \in \mathcal{P}(\Theta^M)$ such that

$$\hat{u}^M(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}^M}(t,\theta) \cdot A^\theta\big(X_{\hat{u}^M}(t,\theta)\big) v \, d\nu^M(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0, T]$, and $\bar{\theta} \in \operatorname{supp}(\nu^M) \implies \bar{\theta} \in \arg\max_{\Theta^M} a(x^\theta_{\hat{u}^M}(T), \theta)$.

In other words, $\hat{u}^M$ is as well an extremal for

$$\mathcal{G}_{\nu^M}(u) = \int_\Theta \ell(x_u^\theta(T), \theta) \, d\nu^M(\theta) + \beta \int_0^T f(u(s)) \, ds.$$

Γ-convergence and PMP (again)

Consider $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u} \in \arg\min \mathcal{F}_{\Theta}$.

Γ-convergence and PMP (again)

Consider $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u} \in \arg\min \mathcal{F}_{\Theta}$. We have

$$\begin{cases} \nu^M \rightharpoonup^* \nu & \text{(up to subseq.)} \\ X_{\hat{u}^M} \to_{C^0} X_{\hat{u}} & \\ \Lambda_{\hat{u}^M} \to_{C^0} \Lambda_{\hat{u}} & \\ \hat{u}^M(t) \to \hat{u}(t) & \text{a.e. (up to subseq.)} \end{cases} \qquad \text{as } M \to \infty.$$

Γ-convergence and PMP (again)

Consider $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u} \in \arg\min \mathcal{F}_{\Theta}$. We have

$$\begin{cases} \nu^M \rightharpoonup^* \nu & \text{(up to subseq.)} \\ X_{\hat{u}^M} \to_{C^0} X_{\hat{u}} & \\ \Lambda_{\hat{u}^M} \to_{C^0} \Lambda_{\hat{u}} & \\ \hat{u}^M(t) \to \hat{u}(t) & \text{a.e. (up to subseq.)} \end{cases} \qquad \text{as } M \to \infty.$$

We can pass to the limit here:

$$\hat{u}^M(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_{\Theta} \Lambda_{\hat{u}^M}(t,\theta) \cdot A^{\theta}\big(X_{\hat{u}^M}(t,\theta)\big) v \, d\nu^M(\theta) - \beta f(v) \right\}$$

We use that the subdifferential ∂f has closed graph.

Γ-convergence and PMP (again)

Consider $\hat{u}^M \in \arg\min \mathcal{F}_{\Theta^M}$, and assume that $\hat{u}^M \to_{L^p} \hat{u} \in \arg\min \mathcal{F}_{\Theta}$. We have

$$\begin{cases} \nu^M \rightharpoonup^* \nu & \text{(up to subseq.)} \\ X_{\hat{u}^M} \to_{C^0} X_{\hat{u}} & \\ \Lambda_{\hat{u}^M} \to_{C^0} \Lambda_{\hat{u}} & \\ \hat{u}^M(t) \to \hat{u}(t) & \text{a.e. (up to subseq.)} \end{cases} \qquad \text{as } M \to \infty.$$

We get:

$$\hat{u}(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_{\Theta} \Lambda_{\hat{u}}(t,\theta) \cdot A^\theta\big(X_{\hat{u}}(t,\theta)\big) v \, d\nu(\theta) - \beta f(v) \right\}$$

We use that the subdifferential ∂f has closed graph.

PMP for minimax problems

Theorem (S., 2024)

Let $\hat{u} \in \arg\min \mathcal{F}_\Theta$*. Then, considering* $X_{\hat{u}} : [0, T] \times \Theta \to \mathbb{R}^n$ *and* $\Lambda_{\hat{u}} : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ *as before, there exists a probability measure* $\nu \in \mathcal{P}(\Theta)$ *such that*

$$\hat{u}(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}}(t,\theta) \cdot A^\theta\big(X_{\hat{u}}(t,\theta)\big) v \, d\nu(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0, T]$*, and* $\bar{\theta} \in \operatorname{supp}(\nu) \implies \bar{\theta} \in \arg\max_{\theta \in \Theta} a(x^\theta_{\hat{u}}(T), \theta)$.

PMP for minimax problems

Theorem (S., 2024)

Let $\hat{u} \in \arg\min \mathcal{F}_\Theta$. Then, considering $X_{\hat{u}} : [0, T] \times \Theta \to \mathbb{R}^n$ and $\Lambda_{\hat{u}} : [0, T] \times \Theta \to (\mathbb{R}^n)^*$ as before, there exists a probability measure $\nu \in \mathcal{P}(\Theta)$ such that

$$\hat{u}(t) \in \arg\max_{v \in \mathbb{R}^m} \left\{ \int_\Theta \Lambda_{\hat{u}}(t, \theta) \cdot A^\theta \big(X_{\hat{u}}(t, \theta) \big) v \, d\nu(\theta) - \beta f(v) \right\}$$

for a.e. $t \in [0, T]$, and $\bar{\theta} \in \operatorname{supp}(\nu) \implies \bar{\theta} \in \arg\max_{\theta \in \Theta} a(x_{\hat{u}}^\theta(T), \theta)$.

In other words, $\hat{u}$ is as well an extremal for

$$\mathcal{G}_\nu(u) = \int_\Theta \ell(x_u^\theta(T), \theta) \, d\nu(\theta) + \beta \int_0^T f(u(s)) \, ds.$$

Ensemble Optimal Control: numerical methods

General strategy

Reduce to finite ensembles using Γ-convergence, and solve it as a finite-dimensional problem.

Ensemble Optimal Control: numerical methods

General strategy

Reduce to finite ensembles using Γ-convergence, and solve it as a finite-dimensional problem.

Taking advantage of the reduction to finite ensembles, we can use numerical methods for Optimal Control problems

Ensemble Optimal Control: numerical methods

General strategy

Reduce to finite ensembles using Γ-convergence, and solve it as a finite-dimensional problem.

Taking advantage of the reduction to finite ensembles, we can use numerical methods for Optimal Control problems:

- Projected Gradient Flow

Ensemble Optimal Control: numerical methods

General strategy

Reduce to finite ensembles using Γ-convergence, and solve it as a finite-dimensional problem.

Taking advantage of the reduction to finite ensembles, we can use numerical methods for Optimal Control problems:

- Projected Gradient Flow;
- Iterative methods based on Pontryagin Maximum Principle ([Sakawa & Shindo, 1980], [Chernousko & Lyubushin, 1982]).

Ensemble Optimal Control: numerical methods

General strategy

Reduce to finite ensembles using Γ-convergence, and solve it as a finite-dimensional problem.

Taking advantage of the reduction to finite ensembles, we can use numerical methods for Optimal Control problems:

- Projected Gradient Flow;
- Iterative methods based on Pontryagin Maximum Principle ([Sakawa & Shindo, 1980], [Chernousko & Lyubushin, 1982]).

Remark

The minimax problem is harder: the measures ν^M are not explicitly given, they should be *adaptively guessed* during the approximation of the optimal control.

Thanks for the attention!

References:

- A. S. Optimal control of ensembles of dynamical systems. *ESAIM: COCV*, 2023.
- A. S. Minimax problems for ensembles of control-affine systems. *SIAM J Control Optim*, accepted in November 2024.
- C. Cipriani, A. S., T. Wöhrer. A minimax optimal control approach for robust neural ODEs. *European Control Conference ECC24*, 2024.