Numerical Approximations for Minimax Markov Control Problems

Tomás Prieto-Rumeau. UNED, Madrid, Spain

Joint work with François Dufour. INRIA, Bordeaux, France

Stochastic control, computational methods, and applications
IMA, Minneapolis, May 2018
Motivation

- This talk is part of a “general approach” to obtain numerical approximations for control problems, developed by FD and TPR since 2010.
- The goal is to provide explicit numerical approximations for a large class of MDPs.

⚠️ Most results in the literature address MDPs with discrete (sometimes, compact) state or action space.

- Our results cover
  - discrete-time and continuous-time discounted MDPs, constrained and unconstrained.
  - discrete-time and continuous-time average MDPs.
  - games against nature, zero sum-games.

- Our ideas are based on discrete approximations to “general probability measures”: quantization, concentration inequalities, approximations in the Wasserstein metric...
Motivation

- This talk is part of a “general approach” to obtain numerical approximations for control problems, developed by FD and TPR since 2010.
- The goal is to provide explicit numerical approximations for a large class of MDPs.
- Most results in the literature address MDPs with discrete (sometimes, compact) state or action space.
- Our results cover
  - discrete-time and continuous-time discounted MDPs, constrained and unconstrained.
  - discrete-time and continuous-time average MDPs.
  - games against nature, zero sum-games.
- Our ideas are based on discrete approximations to “general probability measures”: quantization, concentration inequalities, approximations in the Wasserstein metric...
Statement of the problem

- Approximate the value function and an optimal policy of a minimax Markov control process (a.k.a. robust control problem or game against nature).
- Control model: general state and action spaces $X, A, B$, discrete-time, discounted cost criterion.
- When using the DP approach to the infinite horizon discounted minimax problem, the Bellman equation reads:

$$v(x) = \min_{a \in A} \max_{b \in B} \{c(x, a, b) + \alpha \int v(y) Q(dy|x, a, b)\}$$

for $x \in X$.
- The optimal value function is $v$. An action $a \in A$ attaining the min is optimal.
- It is not possible to solve, in general, this equation.
Statement of the problem

- Approximate the value function and an optimal policy of a minimax Markov control process (a.k.a. robust control problem or game against nature).
- Control model: general state and action spaces $X, A, B$, discrete-time, discounted cost criterion.
- When using the DP approach to the infinite horizon discounted minimax problem, the Bellman equation reads:

$$v(x) = \min_{a \in A} \max_{b \in B} \left\{ c(x, a, b) + \alpha \int v(y) Q(dy|x, a, b) \right\}$$

for $x \in X$.
- The optimal value function is $v$. An action $a \in A$ attaining the min is optimal.
- It is not possible to solve, in general, this equation.
Motivation

- The idea is to discretize $X$, $A$, and $B$, so that

$$v(x) = \min_{a \in A} \max_{b \in B} \left\{ c(x, a, b) + \alpha \int v(y) Q(dy|x, a, b) \right\}, \ x \in X$$

becomes

$$v'(x) = \min_{a \in A'} \max_{b \in B'} \left\{ c(x, a, b) + \alpha \int v'(y) Q'(dy|x, a, b) \right\}, \ x \in X'.$$

- Basic questions:
  - How to discretize the elements of the control problem?
  - Is it possible to get bounds on $||v - v'||$?
**Our approach**

- Discretization of the action spaces $A, B$ is made by finite approximations $A', B'$ in the Hausdorff metric (a **geometric** criterion).
- Discretization of the state space is based on approximating probability measures $Q$ by p.m.'s $Q'$ with finite support (a **probabilistic** criterion).
- We show that

  $$\|v - v'\| \sim d_H(A, A') + d_H(B, B') + d(Q, Q').$$

- The choice of the Hausdorff metric is quite “natural”. What about measuring the distance between two p.m.'s?
Our approach

- Discretization of the action spaces $A, B$ is made by finite approximations $A', B'$ in the Hausdorff metric (a geometric criterion).

- Discretization of the state space is based on approximating probability measures $Q$ by p.m.'s $Q'$ with finite support (a probabilistic criterion).

- We show that

$$||v - v'|| \sim d_H(A, A') + d_H(B, B') + d(Q, Q').$$

- The choice of the Hausdorff metric is quite “natural”. What about measuring the distance between two p.m.'s?
The Wasserstein metric

- **1-Wasserstein metric.** For probability measures in $\mathcal{P}_1(X)$ with finite first moment

$$W_1(\lambda, \mu) = \inf_{\{\nu : \nu_1 = \lambda, \nu_2 = \mu\}} \int_{X \times X} \rho(x_1, x_2) \nu(dx_1, dx_2).$$

- The dual Kantorovich-Rubinstein characterization gives

$$W_1(\lambda, \mu) = \sup_{f \in L^1(X)} \left| \int f \, d\mu - \int f \, d\lambda \right|$$

for all 1-Lipschitz continuous functions.

- $W_1(\mu_n, \mu) \to 0$ if and only if $\mu_n \Rightarrow \mu$ (weak convergence) and

$$\int_X \rho(x, x_0) \mu_n(dx) \to \int_X \rho(x, x_0) \mu(dx)$$

for some (and then for all) $x_0 \in X$. 

Tomás Prieto-Rumeau

Approximation of minimax MCPs
Why use the Wasserstein metric?

- Given a function \( f \) and a p.m. measure \( \mu \) on a Polish space \( X \), we want to approximate

  \[
  \int_X f d\mu \quad \text{with} \quad \int_X f d\lambda
  \]

  for some p.m. \( \lambda \) with finite support.

- If \( d_{LP}(\mu, \lambda) < \epsilon \) and \( f \) is, e.g., bounded and continuous,

  \[
  \left| \int_X f d\mu - \int_X f d\lambda \right| < ??
  \]

- If \( W_1(\mu, \lambda) < \epsilon \) and \( f \) is \( L \)-Lipschitz continuous, then

  \[
  \left| \int_X f d\mu - \int_X f d\lambda \right| < L\epsilon.
  \]

- This forces us to “live” in a Lipschitz continuous world.
Why use the Wasserstein metric?

- Given a function $f$ and a p.m. measure $\mu$ on a Polish space $X$, we want to approximate

$$\int_X f d\mu$$

with

$$\int_X f d\lambda$$

for some p.m. $\lambda$ with finite support.

- If $d_{LP}(\mu, \lambda) < \epsilon$ and $f$ is, e.g., bounded and continuous,

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < ??$$

- If $W_1(\mu, \lambda) < \epsilon$ and $f$ is $L$-Lipschitz continuous, then

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < L\epsilon.$$ 

- This forces us to “live” in a Lipschitz continuous world.
Why use the Wasserstein metric?

- Given a function $f$ and a p.m. measure $\mu$ on a Polish space $X$, we want to approximate

$$\int_X f d\mu \text{ with } \int_X f d\lambda$$

for some p.m. $\lambda$ with finite support.

- If $d_{LP}(\mu, \lambda) < \epsilon$ and $f$ is, e.g., bounded and continuous,

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < ??$$

- If $W_1(\mu, \lambda) < \epsilon$ and $f$ is $L$-Lipschitz continuous, then

$$\left| \int_X f d\mu - \int_X f d\lambda \right| < L\epsilon.$$

- This forces us to “live” in a Lipschitz continuous world.
Approximation in the Wasserstein metric

Theorem

Let $\mu \in \mathcal{P}_1(X)$. For every $\epsilon > 0$ there exists a p.m. $\lambda$ on $X$ with finite support such that $W_1(\mu, \lambda) < \epsilon$.

The construction of $\lambda$ is “theoretical” and it can pose a computational challenge.

For $n$ i.i.d. draws with distribution $\mu$, let $\mu_n$ be the empirical probability measure.

Theorem (Boissard, 2011)

Let $\mu \in \mathcal{P}_1(X)$ have a finite exponential moment. For each $\gamma > 0$ there exist $C_1, C_2 > 0$ such that

$$\mathbb{P}\{W_1(\mu_n, \mu) > \gamma\} \leq C_1 \exp\{-C_2 n\} \quad \text{for all } n \geq 1.$$
Approximation in the Wasserstein metric

**Theorem**

Let $\mu \in \mathcal{P}_1(X)$. For every $\epsilon > 0$ there exists a p.m. $\lambda$ on $X$ with finite support such that $W_1(\mu, \lambda) < \epsilon$.

The construction of $\lambda$ is “theoretical” and it can pose a computational challenge.

For $n$ i.i.d. draws with distribution $\mu$, let $\mu_n$ be the empirical probability measure.

**Theorem (Boissard, 2011)**

Let $\mu \in \mathcal{P}_1(X)$ have a finite exponential moment. For each $\gamma > 0$ there exist $C_1, C_2 > 0$ such that

$$\mathbb{P}\{W_1(\mu_n, \mu) > \gamma\} \leq C_1 \exp\{-C_2 n\} \quad \text{for all } n \geq 1.$$
Dynamics of the control model

- Two players, the controller and the opponent (or nature), act on a stochastic dynamic system \( \{x_k\} \).
  - The system is in state \( x_k \),
  - The controller plays first and takes an action \( a_k \),
  - The opponent observes the controller and then takes an action \( b_k \),
  - The controller incurs a cost \( c(x_k, a_k, b_k) \),
  - The system makes a transition \( x_{k+1} \sim Q(\cdot|x_k, a_k, b_k) \).

- The game is played on an infinite time horizon \( t \geq 0 \) with costs discounted at a constant factor \( 0 < \alpha < 1 \).
- The controller wants to optimize the “worst scenario”

\[
\min_{\pi} \sup_{\gamma} \mathbb{E} \left[ \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t, b_t) \right].
\]
Dynamics of the control model

- Two players, the controller and the opponent (or nature), act on a stochastic dynamic system \( \{x_k\} \).
  - The system is in state \( x_k \),
  - The controller plays first and takes an action \( a_k \),
  - The opponent observes the controller and then takes an action \( b_k \),
  - The controller incurs a cost \( c(x_k, a_k, b_k) \),
  - The system makes a transition \( x_{k+1} \sim Q(\cdot|x_k, a_k, b_k) \).

- The game is played on an infinite time horizon \( t \geq 0 \) with costs discounted at a constant factor \( 0 < \alpha < 1 \).
- The controller wants to optimize the “worst scenario”

\[
\min_{\pi} \sup_{\gamma} E \left[ \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t, b_t) \right].
\]
The control model $\mathcal{M}$

Consider a control model $\mathcal{M} = (X, A, B, \mathcal{K}_A, \mathcal{K}, Q, c)$ where

- A Borel state space $X$.
- Borel action spaces $A$ and $B$.
- $A(x)$ is the set of available actions for the controller in state $x \in X$,
  \[ \mathcal{K}_A = \{(x, a) : x \in X, a \in A(x)\}. \]
- $B(x, a)$ is the set of available actions for the nature,
  \[ \mathcal{K} = \{(x, a, b), x \in X, a \in A(x), b \in B(x, a)\}. \]
- $Q = Q(B|x, a, b)$ is a stochastic kernel on $X$ given $\mathcal{K}$.
- $c : \mathcal{K} \rightarrow \mathbb{R}$ is the cost function.
Definition of the control model

- **Policies for the controller** $\Pi_A$. Based on the history
  \[(x_0, a_0, b_0, \ldots, x_{t-1}, a_{t-1}, b_{t-1}, x_t)\]
  the controller (randomly) chooses an action in $A(x_t)$.

- **Deterministic stationary policies** $\mathcal{F}_A$ are given by $f : X \rightarrow A$ such that $f(x) \in A(x)$ for $x \in X$.

- **Policies for the opponent** $\Pi_B$. Based on
  \[(x_0, a_0, b_0, \ldots, x_{t-1}, a_{t-1}, b_{t-1}, x_t, a_t)\]
  the opponent (randomly) chooses an action in $B(x_t, a_t)$.

- Given an initial state $x \in X$ and policies $(\pi, \gamma) \in \Pi_A \times \Pi_B$, there exists a unique p.m. $P^{\pi, \gamma, x}$ on $\mathbb{K}^\infty$ modeling the minimax problem.
Definition of the control model

Optimality criterion

- The total expected discounted cost is

\[ V(x, \pi, \gamma) = \mathbb{E}^{\pi, \gamma, x} \left[ \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t, b_t) \right] \]

- The minimax value function \( V^* \) is defined as

\[ V^*(x) = \inf_{\pi \in \Pi_A} \sup_{\gamma \in \Pi_B} V(x, \pi, \gamma) \] for every \( x \in X \).

- A policy \( \pi^* \in \Pi_A \) for the controller is a minimax policy if

\[ \sup_{\gamma \in \Pi_B} V(x, \pi^*, \gamma) = V^*(x) \] for every \( x \in X \).
The action sets $A(x)$ and $B(x,a)$ are compact, and Lipschitz continuous; e.g.,

$$d_H(A(x), A(y)) \leq L_A \cdot \rho(x, y) \quad \text{for all } x, y \in X,$$

There exist a p.m. $\mu \in \mathcal{P}_1(X)$ and $q : X \times \mathbb{K} \rightarrow \mathbb{R}^+$ such that

$$Q(B|x, a, b) = \int_B q(y|x, a, b) \mu(dy)$$

for all $B \in \mathcal{B}(X)$ and $(x, a, b) \in \mathbb{K}$.

For some “weight function” $w : X \rightarrow [1, \infty)$

$$|c(x, a, b)| \leq \bar{c}w(x) \quad \text{for all } (x, a, b) \in \mathbb{K}.$$

$Qw \leq \beta w$ with $\alpha \beta < 1$. 
Hypotheses

**Lipschitz continuity conditions**
- The weight function $w$, the cost function $c$ and the density $q$ are Lipschitz continuous.
- The density function satisfies
  - $q(y|x,a,b) \leq qw(x)$.
  - $y \mapsto w(y)q(y|x,a,b)$ is $Lqw(x)$-Lipschitz-continuous.

**Some notation**
- $u \in \mathbb{B}_w(X)$ if it is measurable and $w$-bounded:
  $$|u(x)| \leq ||u||_w w(x) \quad \text{for } x \in X.$$
- $u \in \mathbb{L}_w(X)$ if it is Lipschitz-continuous and $w$-bounded.
Dynamic programming equation

- Given \( u \in B_w(X) \) define, for \( x \in X \),

\[
Tu(x) = \min_{a \in A(x)} \max_{b \in B(x,a)} \left\{ c(x,a,b) + \alpha Qu(x,a,b) \right\}.
\]

- The dynamic programming equation is

\[
\nu(x) = Tv(x) \quad \text{for all} \ x \in X.
\]

- Regularity of \( T \):
  - If \( u \in B_w(X) \) then \( Tu \in L_w(X) \).
  - \( T \) is an \( \alpha \beta \)-contraction operator on \( B_w(X) \).
Dynamic programming equation

- Given $u \in \mathbb{B}_W(X)$ define, for $x \in X$,

$$Tu(x) = \min_{a \in A(x)} \max_{b \in B(x,a)} \{c(x, a, b) + \alpha Qu(x, a, b)\}.$$

- The **dynamic programming** equation is

$$v(x) = Tv(x) \quad \text{for all } x \in X.$$

- Regularity of $T$:
  - If $u \in \mathbb{B}_W(X)$ then $Tu \in \mathbb{L}_W(X)$.
  - $T$ is an $\alpha \beta$-contraction operator on $\mathbb{B}_W(X)$. 
Theorem

(i) The value function $V^*$ is the unique solution in $\mathbb{B}_w(X)$ of the DP equation.

(ii) $f \in F_A$ is a minimax policy if and only if it “attains the minimum” in the DP equation

$$V^*(x) = \max_{b \in B(x, f(x))} \left\{ c(x, f(x), b) + \alpha QV^*(x, f(x), b) \right\}$$

for every $x \in X$, and such $f$ indeed exists.

(iii) $V^*$ is in $L_w(X)$

$$L_{V^*} = \left( L_c + \frac{\alpha L_q \bar{c} \mu(w)}{1 - \alpha \beta} \right) \cdot (1 + L_A) \cdot (1 + L_B).$$
Dynamic programming equation

Theorem

(i) The value function $V^*$ is the unique solution in $\mathbb{B}_w(X)$ of the DP equation.

(ii) $f \in \mathbb{F}_A$ is a minimax policy if and only if it “attains the minimum” in the DP equation

$$V^*(x) = \max_{b \in B(x,f(x))} \{c(x,f(x),b) + \alpha QV^*(x,f(x),b)\}$$

for every $x \in X$, and such $f$ indeed exists.

(iii) $V^*$ is in $\mathbb{L}_w(X)$

$$LV^* = \left( L_c + \frac{\alpha L_q \bar{c} \mu(w)}{1 - \alpha \beta} \right) \cdot (1 + L_A) \cdot (1 + L_B).$$
Discretization of the action space

For all small $\delta > 0$ construct finite set $A_\delta(x) \subseteq A(x)$ and $B_\delta(x,a) \subseteq B(x,a)$ with

- $x \mapsto A_\delta(x)$ and $(x,a) \mapsto B_\delta(x,a)$ are Lipschitz continuous.
- The finite action sets are close to the original action sets

\[ d_H(A(x), A_\delta(x)) \leq \delta \omega(x) \quad \text{and} \quad d_H(B(x,a), B_\delta(x,a)) \leq \delta \omega(x). \]

- In this way, we obtain the feasible pairs $\mathbb{K}_A^\delta$ and triplets $\mathbb{K}^\delta$. 
Approximation of the control model

Discretization of the state space

Recall that the transition kernel is

$$Q(B|x, a, b) = \int_B q(y|x, a, b)\mu(dy).$$

Replace $\mu \in \mathcal{P}_1(X)$ with $\nu \in \mathcal{P}_1(X)$ with finite support.

Define the transition kernel

$$Q_\nu(B|x, a, b) = \frac{\int_B q(y|x, a, b)\nu(dy)}{\int_X q(y|x, a, b)\nu(dy)}.$$
Approximation of the control model

The approximating control model

- We approximate $\mathcal{M}$ with
  \[
  \mathcal{M}_{V,\delta} = (X, A, B, K^\delta_A, K^\delta, Q_V, c).
  \]
- The discount factor is $\alpha > 0$.
- We can define policies for the controller and the opponent.
- The minimax value function $V_{V,\delta}^*$ is
  \[
  V_{V,\delta}^*(x) = \inf_{\pi \in \Pi_A^\delta} \sup_{\gamma \in \Pi_B^\delta} V_{V,\delta}(x, \pi, \gamma) \quad \text{for every } x \in X.
  \]
- The definition of a minimax strategy is similar to that for $\mathcal{M}$.
The dynamic programming operator is, for $x \in X$,

$$T_{\nu, \delta} v(x) = \min_{a \in A_\delta(x)} \max_{b \in B_\delta(x, a)} \left\{ c(x, a, b) + \alpha Q_\nu v(x, a, b) \right\}.$$ 

The operator $T_{\nu, \delta}$ inherits all properties from $T$ except:

- If $u \in B_w(X)$ then $Tu \in L_w(X)$
- If $u \in B_w(X)$ then $T_{\nu, \delta} u \in C_w(X)$

So, $\mathcal{M}_{\nu, \delta}$ is not a Lipschitz continuous control model!
Approximation of the control model

Solving $\mathcal{M}_{\nu,\delta}$

(i) The value function $V_{\nu,\delta}^*$ is the unique solution in $\mathbb{B}_w(X)$ of the DP equation $\nu = T_{\nu,\delta} \nu$.

(ii) The policy $f \in F^\delta_A$ is a minimax policy if and only if

$$V_{\nu,\delta}^*(x) = \max_{b \in B_{\delta}(x, f(x))} \left\{ c(x, f(x), b) + \alpha Q_{\nu} V_{\nu,\delta}^*(x, f(x), b) \right\}$$

for $x \in X$, and such $f$ indeed exist.
Theorem

There are constants $G_1$ and $G_2$ s.t. for every $\delta > 0$ and every $\nu \in \mathcal{P}_1(X)$ we have

$$\|V^* - V_{\nu,\delta}\|_W \leq G_1 \delta + G_2 W_1(\mu, \nu).$$

The constants are $G_1 = \frac{2}{1 - \alpha \beta}\left(L_c + \frac{\alpha L_q \bar{c}}{1 - \alpha \beta} \mu(w)\right)$ and

$$G_2 = \frac{4\alpha \bar{c}}{(1 - \alpha \beta)^2}(L_{wq} + \bar{q}L_w + L_q \beta) + \frac{4\alpha \bar{q}}{(1 - \alpha \beta)}\left(L_c + \frac{\alpha L_q \bar{c} \mu(w)}{1 - \alpha \beta}\right) \cdot (1 + L_A) \cdot (1 + L_B).$$
Approximation of the value function

Sketch of the proof

- Compare the operators $T$ and $T_{\nu,\delta}$ when applied to a Lipschitz continuous $u \in \mathbb{L}_w(X)$.
- Difficulty: double (nested) approximation of the actions sets.
- Then compare $V^*$ and $V^*_{\nu,\delta}$ taking advantage that $V^* \in \mathbb{L}_w(X)$. 
Approximation of a minimax strategy

Idea

- Find a minimax policy $f^*_{\nu,\delta} \in F^\delta_A$ for the control problem $M_{\nu,\delta}$.
- This policy is admissible for $M$ because $F^\delta_A \subseteq F_A$.
- Plug this policy into the original model $M$.
- Get upper bounds on

$$0 \leq \sup_{\gamma \in \Pi_B} V(x, f^*_{\nu,\delta}, \gamma) - V^*(x) \leq ??$$
**Theorem**

There are constants $K_1$ and $K_2$ s.t. for every $\delta > 0$ and every $\nu \in \mathcal{P}_1(X)$ we have

$$\| V^* - \sup_{\gamma \in \Pi_B} V(\cdot, f^*_\nu, \gamma) \|_w \leq K_1 \delta + K_2 W_1(\mu, \nu).$$

**Sketch of the proof:**

- The value function $V^*_\nu, \delta$ is not Lipschitz continuous.
- We can prove that it is locally Lipschitz continuous.
- We can approximate $V^*_\nu, \delta$ in the $w$-norm with a Lipschitz continuous function.
Consider the dynamics

\[ x_{t+1} = \max \{ x_t + a_t - \xi_t, 0 \} \quad \text{for } t \in \mathbb{N} \]

where

- \( x_t \) is the stock level at the beginning of period \( t \);
- \( a_t \) is the amount ordered at the beginning of period \( t \);
- \( \xi_t \) is the random demand at the end of period \( t \) with distribution \( F(\cdot, b) \).

The capacity of the warehouse is \( M > 0 \). Therefore,

\[ X = A = [0, M] \quad A(x) = [0, M - x] \quad \text{and} \quad B = B(x, a) = [b, \bar{b}] \]
An inventory management system

The controller incurs:
- a buying cost of $b > 0$ for each unit;
- a holding cost $h > 0$ for each period and unit;
- and receives $p > 0$ for each unit that is sold.

The running cost function is

$$c(x, a, b) = ba + h(x + a) - pE[\min\{x + a, \xi\}].$$

**Theorem**

If the $\{\xi_t\}$ are i.i.d. with density function $f(\cdot, \cdot)$, which is Lipschitz continuous on $[0, M] \times B$ with $f(0, b) = 0$, then the inventory management system satisfies our assumptions.
Fix $0 < p < 1$. The probability measure $\mu$ is

$$\mu\{0\} = p \quad \text{and} \quad \mu(B) = \frac{1-p}{M} \lambda(B) \quad \text{for measurable } B \subseteq (0, M].$$

Given $n \geq 1$, approximate it with

$$\nu_n\{0\} = p \quad \text{and} \quad \nu_n\{M \cdot j/n\} = (1-p)/n \quad \text{for } 0 < j \leq n$$

supported on the grid $\Gamma_n$.

The density function of the demand is a $\gamma(1/b, 2)$.

The approximating action sets are

$$A_\delta(x) = \left\{ \frac{(M-x)j}{n-1} : j = 0, 1, \ldots, n-1 \right\}.$$

$$B_\delta(x, a) = \left\{ \frac{(b-a)j}{n-1} : j = 0, 1, \ldots, n-1 \right\}.$$
An inventory management system

For \( n = 20, 100, 200 \) we solve the minimax problem on \( \Gamma_n \) and then we extend it to the whole \( X = [0, M] \).
An inventory management system

We determine a minimax $\tilde{f}_{n,\delta}$ for $M_{n,\delta}$ and we evaluate it for $M$.

Intuitively, $f^*(x) = (M_0 - x)^+$ for some $M_0 \approx 3.8$. 
Conclusions

- We have proposed a general procedure to approximate a continuous state and action minimax control problem.
- We can do this for a “Lipschitz-continuous” control model.
- The approximation error combines Hausdorff (for the actions) and Wasserstein (for the states) distances.
- In the application, our method provides very good approximations.

Thank you for your attention.