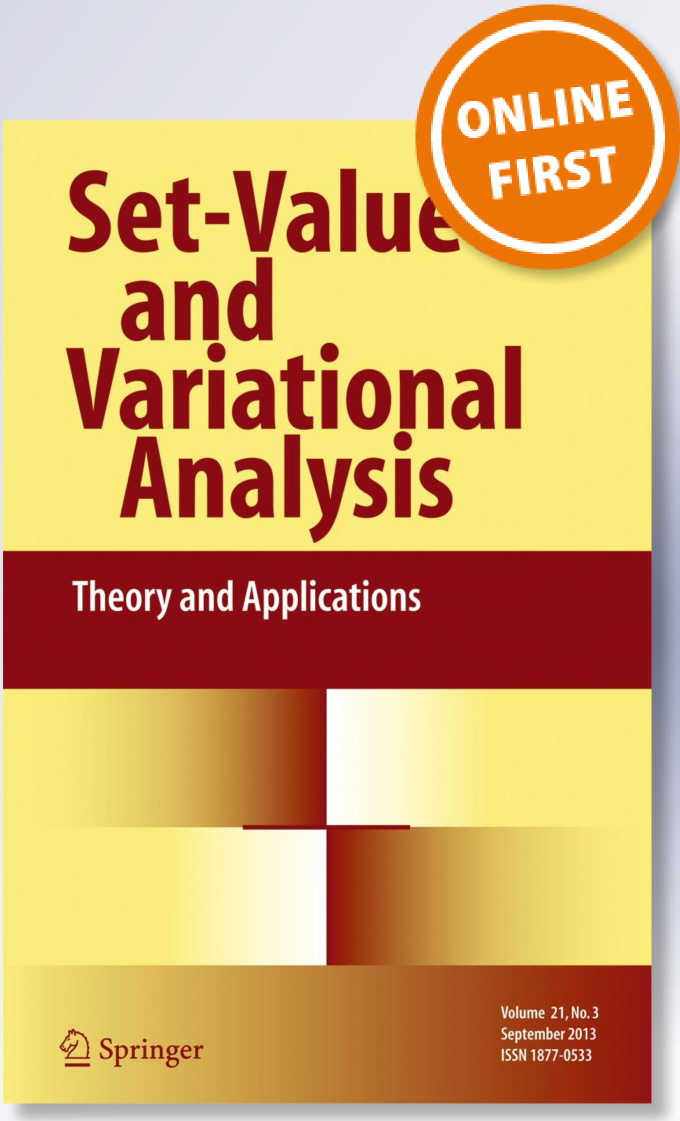
An Approximation Scheme for Uncertain Minimax Optimal Control Problems

Laura S. Aragone, Justina Gianatti, Pablo A. Lotito & Lisandro A. Parente

Set-Valued and Variational Analysis
Theory and Applications

ISSN 1877-0533

Set-Valued Var. Anal
DOI 10.1007/s11228-017-0450-7



Set-Valued Var. Anal
https://doi.org/10.1007/s11228-017-0450-7

An Approximation Scheme for Uncertain Minimax Optimal Control Problems

Laura S. Aragone[1] **· Justina Gianatti**[1] **·**
Pablo A. Lotito[2] **· Lisandro A. Parente**[1]

Received: 30 September 2016 / Accepted: 2 October 2017

Abstract In this work, we address an uncertain minimax optimal control problem with linear dynamics where the objective functional is the expected value of the supremum of the running cost over a time interval. By taking an independently drawn random sample, the expected value function is approximated by the corresponding sample average function. We study the epi-convergence of the approximated objective functionals as well as the convergence of their global minimizers. Then we define an Euler discretization in time of the sample average problem and prove that the value of the discrete time problem converges to the value of the sample average approximation. In addition, we show that there exists a sequence of discrete problems such that the accumulation points of their minimizers are optimal solutions of the original problem. Finally, we propose a convergent descent method to solve the discrete time problem, and show some preliminary numerical results for two simple examples.

Keywords Minimax control problems · Uncertain control problems · Sample average approximation · Epi-convergence · Numerical solutions

This work has been partially supported by the following projects: PICT-2012-2212 ANPCyT and PIP 112-201101-286 CONICET.

✉ Lisandro A. Parente
parente@cifasis-conicet.gov.ar

Laura S. Aragone
aragone@cifasis-conicet.gov.ar

Justina Gianatti
gianatti@cifasis-conicet.gov.ar

Pablo A. Lotito
lotito@exa.unicen.edu.ar

[1] CIFASIS-CONICET-UNR, Ocampo y Esmeralda, S2000EZP, Rosario, Argentina

[2] PLADEMA-UNCPBA-CONICET, Paraje Arroyo Seco, B7000, Tandil, Argentina

Mathematics Subject Classification (2010) 49K35 · 49M25 · 49M37

1 Introduction

In this work we consider an uncertain minimax optimal control problem. We assume that the dynamics is linear, and the involved coefficients as well as the initial condition depend on stochastic parameters. The goal is to minimize the expected value of a supremum over the time interval of a function which depends on the state and also on the stochastic parameter. The aim of this paper is to propose an approximation scheme for this problem, combining some of the results recently presented in [19] and [13].

Minimax optimal control problems differ from those problems usually considered in the optimal control literature where an accumulated cost is minimized. There are many applications where minimizing a maximum arises naturally, as for instance, minimization of the maximum trajectory deviation from what is desired ([10, 11, 17]). These problems were studied in the last decades by several authors. As usual in the optimal control theory there are two main approaches. The first one is based on the Dynamic Programming Principle where the value function is obtained as the unique viscosity solution of the Hamilton-Jacobi-Bellman equation, see for instance [1] and [7] for the deterministic framework and [6] and [8] for the stochastic one. The other approach is based on the Pontryagin Maximum Principle, and we recommend [24] for the case of maximum cost.

Instead of dealing with those approaches, in this work we use the nonsmooth optimization techniques presented in [13]. In that work, a numerical method for deterministic minimax optimal control problems is proposed through the definition of a suitable discrete time approximation. Since the definition of the objective function includes a supremum, in order to obtain (directional) differentiability properties, convexity assumptions are added. Then, a set of optimality conditions for the continuous and discrete time cases are obtained, allowing the design of an easily implementable descent method.

Many applications are naturally (or better) modeled by systems with uncertainties. There are different sources of uncertainty such as measurement errors, uncertain initial data, or unknown parameters and inputs. Usually in the literature, there are two different approaches to dealing with uncertain optimization problems. The first one is to consider the criterion expressed in terms of expectations of the cost (e.g. [15, 19, 22]), and the second one is to consider the worst-case performance criterion (e.g. [18, 23, 27]). When there is information about the behavior of the perturbation (probably available from a statistical analysis), and that is if we assume that the probability distribution of the uncertainties is known, then it makes sense to consider the mean value as objective criterion. Although, depending on the goal of the problem, there are many cases where the other approach is chosen. On the other hand, when no information is available on the expected perturbations, then the worst-case approach is commonly used. An analysis of advantages an disadvantages of using this two options is out of the scope of this paper. In this work we assume the knowledge of the probability distribution of the uncertainties involved, and we deal with the first approach.

In the recent work [19], the authors present a numerical framework to solve an uncertain optimal control problem with Mayer-type objective functional. The main idea is to approximate the expected value with a sample average (see [14, 22]), by taking an independently drawn random sample from the space of stochastic parameters. With strong differentiability assumptions, they provide results about the convergence of the objective functionals in terms of epi-convergence, and they also show that accumulation points of global minimizers of the approximate problems are minimizers of the original problems.

The main difference in the present work is that the functional to be minimize involves a supremum, so assumptions about differentiability cannot be considered. In order to obtain some kind of differentiability of the objective functional, we add some convexity assumptions as in [13]. Also, for the sake of simplicity we assume that the cost function only depends on the state and the random parameter. As in [19], using some previous notions and results about random lower semicontinuity and epi-convergence, we analyze the relationship between our problem and the sample average approximation. With the aim of obtaining a computationally implementable scheme, we define an associated discrete time problem and analyze the convergence of the values of the problems. We also prove the existence of a sequence of discrete time problems such that the accumulation points of their minimizers are optimal solutions of the original problem. We provide optimality conditions for the three problems and propose a convergent descent numerical method following [13].

The article is organized as follows: In Section 2 we present some preliminary results which are essential for the next sections. In Section 3 we state the main assumptions that we make in the entire paper, we define the uncertain minimax optimal control problem, the sample average approach and the associated discrete time problem, showing the existence of minimizers for all of them. In Section 4, we prove the epi-convergence of the objective functionals of the sample average problem to the original one and also a minimizers convergence result. In addition, we provide convergence results for the values and minimizers of the discrete time problems. In Section 5 we analyze the directional differentiability of the cost functionals and we introduce optimality conditions for all the problems. In Section 6, we propose an algorithm to solve the discrete time problem and present a convergence result. Finally, in Section 7 we give some preliminary numerical results.

2 Preliminaries

In this section, for the convenience of the reader, we recall some notions and results that are essential for the next sections.

Definition 1 (Carathéodory functions) Let $(\Omega, \mathcal{A})$ be a measurable space, X, Y metric spaces and consider the function $f : X \times \Omega \to Y$. The function f is called *Carathéodory function* if

1. $x \mapsto f(x, \omega)$ is continuous for each $\omega \in \Omega$,
2. $\omega \mapsto f(x, \omega)$ is measurable for each $x \in X$.

In particular, it is well known that Carathéodory functions are jointly measurable.

Now we give the notion of *epi-convergence*, which is a fundamental convergence concept for sequences of lower semicontinuous functions in optimization theory, and variational analysis. It was studied for the first time in the 60's, in the initial works [25, 26] and [16]. Then, in the 80's it was used in the study of approximation of nonlinear programming problems such as [3–5]. For a detailed analysis and historical remarks on this topic we recommend [21, Chapter 7]. As in [19], we also introduce the concept of *random lower semicontinuity*.

Definition 2 Let (X, d) be a separable complete metric space. Consider the sequence of lower semicontinuous functions $f_M : X \to \mathbb{R}$, $M \in \mathbb{N}$. We say that f_M epi-converges to f, denoted $f_M \overset{epi}{\to} f$, if and only if

1. $\liminf f_M(x_M) \geq f(x)$ whenever $x_M \to x$,
2. $\lim f_M(x_M) = f(x)$ for at least one sequence $x_M \to x$.

Definition 3 Let (X, d) be a separable complete metric space with B the Borel sigma-field. Let $\mathbb{P}$ be a probability measure on the measurable space $(\Omega, \mathcal{A})$ such that $\mathcal{A}$ is $\mathbb{P}$-complete. A function $f : X \times \Omega \to \mathbb{R}$ is a random lower semicontinuous function if and only if

1. for all $\omega \in \Omega$, the function $x \to f(x, \omega)$ is lower semicontinuous,
2. $(x, \omega) \mapsto f(x, \omega)$ is $B \otimes \mathcal{A}$ measurable.

As we said in the introduction, following the lines stated in [19], one of the aims in this paper is to approximate an uncertain minimax optimal control problem by a sample average approach. In order to justify this approximation, we need the following result.

Theorem 4 [5, Theorem 2.3] *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space such that $\mathcal{A}$ is $\mathbb{P}$-complete. Let (X, d) be a separable complete metric space. Suppose that the function $f : X \times \Omega \to \mathbb{R}$ is a random lower semicontinuous function and there exists an integrable function $a_0 : \Omega \to \mathbb{R}$ such that $f(x, \omega) \geq a_0(\omega)$ almost surely. Let $\{\omega_1, \ldots, \omega_M\}$ be an independent $\mathbb{P}$-distributed random draw, and define*

$$\hat{f}(\cdot, \omega_1, \ldots, \omega_M) := \frac{1}{M} \sum_{i=1}^{M} f(\cdot, \omega_i). \tag{1}$$

Then, as $M \to \infty$, $\hat{f}(\cdot, \omega_1, \ldots, \omega_M)$ epi-converges almost surely to $\mathbb{E}^{\mathbb{P}} f(\cdot, \omega)$.

The last result that we include is about the convergence of minimizers.

Theorem 5 [2, Theorem 2.5] *Let (X, d) be a separable complete metric space. Consider a sequence of lower semicontinuous functions $f_M : X \to \mathbb{R}$ such that f_M epi-converges to f. If $\{x_M\}_{M \in \mathbb{N}} \subset X$ is a sequence of global minimizers of f_M, and $\bar{x}$ is any accumulation point of this sequence (along a subsequence indexed by a set $K \subset \mathbb{N}$), then $\bar{x}$ is a global minimizer of f and* $\lim_{M \in K} \inf_{x \in X} f_M(x_M) = \inf_{x \in X} f(x)$.

Further analysis on this topic can be found in [14].

3 Minimax Optimal Control Problems

In this section we present the uncertain minimax optimal control problem (P) that we want to solve, the first approximation via a sample average (P_M), and finally a discrete time problem that approximates (P_M). We state the main assumptions that we make in the entire paper, we prove the well-posedness of all the problems and also the existence of minimizers. The relationship between the values and optimal solutions of the problems will be studied in Section 4.

3.1 Uncertain Minimax Optimal Control Problem

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space. We consider the following uncertain system

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} y(t,\omega) = g(t, y(t,\omega), u(t), \omega), & t \in [0,T] \\ y(0,\omega) = x + \phi(\omega), & x \in \mathbb{R}^r, \end{cases} \tag{2}$$

for $\omega \in \Omega$, where $g : [0,T] \times \mathbb{R}^r \times \mathbb{R}^m \times \Omega \to \mathbb{R}^r$ is a given function. In the notation above $y_u(t,\omega) \in \mathbb{R}^r$ denotes the state function and $u(t) \in \mathbb{R}^m$ the control. We define the set of controls as

$$\mathcal{U} = \{u : [0,T] \to U \subset \mathbb{R}^m; u(\cdot) \text{ measurable}\}, \tag{3}$$

where U is a compact and convex set, and the function to be minimized $J : \mathcal{U} \to \mathbb{R}$ is defined as

$$J(u) = \mathbb{E}^{\mathbb{P}} \left[\sup_{t \in [0,T]} f(y_u(t,\omega), \omega) \right], \tag{4}$$

with $f : \mathbb{R}^r \times \Omega \to \mathbb{R}$. Then the uncertain minimax optimal control problem that we consider is

$$\min J(u); \quad u \in \mathcal{U}. \tag{P}$$

In order to prove the well-posedness of the above definitions, we state the main assumptions that we consider in this work.

(A.1) The function g is linear in y and u, i.e.

$$g(t, y, u, \omega) = A(t,\omega)y + B(t,\omega)u + C(t,\omega), \tag{5}$$

where $A : [0,T] \times \Omega \to \mathbb{R}^{r \times r}$, $B : [0,T] \times \Omega \to \mathbb{R}^{r \times m}$ and $C : [0,T] \times \Omega \to \mathbb{R}^r$ are Carathéodory functions, bounded on $[0,T] \times \Omega$. For $\Psi = A, B, C$ we denote M_Ψ their bounds. In addition, there exist measurable functions $L_\Psi : \Omega \to \mathbb{R}$, such that for all $t, s \in [0,T]$ and for each $\omega \in \Omega$,

$$|\Psi(t,\omega) - \Psi(s,\omega)| \leq L_\Psi(\omega)|t - s|. \tag{6}$$

Also, the function ϕ is measurable and bounded on Ω, with bound M_ϕ .

(A.2) The function f is Carathéodory, and the map $y \mapsto f(y,\omega)$ is convex and continuously differentiable for each $\omega \in \Omega$. Also, there exists a function $C_f : \Omega \to \mathbb{R}$ that belongs to $L^1(\Omega)$, such that for all $y \in \mathbb{R}^r$, and $\omega \in \Omega$,

$$|\nabla_y f(y,\omega)| \leq C_f(\omega)\,[|y| + 1]\,. \tag{7}$$

In addition, there exists $\bar{x} \in \mathbb{R}^r$ such that $f(\bar{x}, \cdot)$ belongs to $L^1(\Omega)$.

Remark 6 The linearity of the dynamics and the convexity of the function f prove that the functional J is convex. On one hand, this implies that local minimizers are global minimizers, then Theorem 5 assures the convergence of the solutions of the sample average approximation problems to optimal controls for the continuous problem (see Theorem 14). Weaker convexity assumptions would require a deeper consistency analysis. On the other hand, convexity leads to necessary and sufficient optimality conditions. Otherwise, the lack of convexity may require stronger differentiability hypotheses in order to obtain at least necessary optimality conditions. Also, the linearity of the dynamics will be useful to derive the variation of the state with respect to control (see Theorem 20).

Lemma 7 *Under the assumption* $(A.1)$, *for any* $u \in \mathcal{U}$ *and* $\omega \in \Omega$ *there exists a unique solution* y_u *of the system* (2), *and the map* $(t,\omega) \mapsto y_u(t,\omega)$ *is a Carathéodory function. In addition, there exists* $M_y > 0$ *such that for each* $\omega \in \Omega$,

$$\sup_{t\in[0,T]} |y_u(t,\omega)| \le M_y. \tag{8}$$

Proof It is clear that for all $u \in \mathcal{U}$ there exists a unique solution y_u of (2), and y_u is a Carathéodory function. By Hölder inequality, we have

$$\begin{aligned} |y_u(t,\omega)| &\le \textstyle\int_0^t |A(s,\omega)||y_u(s,\omega)|\mathrm{d}s + \int_0^t |B(s,\omega)||u(s)|\mathrm{d}s \\ &\quad + \textstyle\int_0^t |C(s,\omega)|\mathrm{d}s + |x| + |\phi(\omega)| \\ &\le \textstyle\int_0^t |A(s,\omega)||y_u(s,\omega)|\mathrm{d}s + T^{\frac{1}{2}} M_B ||u||_{L^2} + T M_C + |x| + |\phi(\omega)|. \end{aligned} \tag{9}$$

Then, by Grönwall's Lemma we obtain

$$\sup_{t\in[0,T]} |y_u(t,\omega)| \le M(u,\omega), \tag{10}$$

where $M(u,\omega) := [T^{\frac{1}{2}} M_B ||u||_{L^2} + T M_C + |x| + |\phi(\omega)|]e^{TM_A}$. Since the set U is compact and the function ϕ is bounded, the result follows. □

Remark 8 Let $K \subset \mathbb{R}^r$ be a bounded set, then for any $y, x \in K$, by assumption $(A.2)$ we have

$$\begin{aligned} |f(x,\omega) - f(y,\omega)| &\le \textstyle\int_0^1 |\nabla_y f(y + \xi(x-y),\omega)||x-y|\mathrm{d}\xi \\ &\le C_f(\omega)\,[|y| + |x-y| + 1]\,|x-y| \\ &\le \tilde{C}_f(\omega)|x-y|. \end{aligned} \tag{11}$$

where $\tilde{C}_f \in L^1(\Omega)$ since K is bounded. In particular, by the previous lemma, there exists $L_f \in L^1(\Omega)$ such that for all $u, v \in \mathcal{U}$ and $\omega \in \Omega$ we obtain

$$\sup_{t\in[0,T]} |f(y_u(t,\omega),\omega) - f(y_v(t,\omega),\omega)| \le L_f(\omega) \sup_{t\in[0,T]} |y_u(t,\omega) - y_v(t,\omega)|. \tag{12}$$

In the remainder we use the last inequality several times.

Lemma 9 *If assumption* $(A.1)$ *holds, then there exists a constant* $C > 0$ *such that for all* $u, v \in \mathcal{U}$ *and* $\omega \in \Omega$ *we have*

$$\sup_{t\in[0,T]} |y_u(t,\omega) - y_v(t,\omega)| \le C||u-v||_{L^2}. \tag{13}$$

Proof By Hölder inequality we obtain,

$$|y_u(t,\omega) - y_v(t,\omega)| \le \int_0^t |A(s,\omega)||y_u(s,\omega) - y_v(s,\omega)|\mathrm{d}s + T^{\frac{1}{2}} M_B ||u-v||_{L^2}. \tag{14}$$

The result follows by the Grönwall's Lemma, defining $C := T^{\frac{1}{2}} M_B e^{TM_A}$. □

Now, to prove that the functional J is well defined we need the following result. For notational convenience we define the function $F : \mathcal{U} \times \Omega \to \mathbb{R}$, as

$$F(u, \omega) := \max_{t \in [0,T]} f(y_u(t, \omega), \omega) \tag{15}$$

Proposition 10 *Assume* $(A.1)$ *and* $(A.2)$ *hold. The function* F *is well defined, it is a Carathéodory function and there exists an integrable function* $L_F : \Omega \to \mathbb{R}$ *such that for all* $u, v \in \mathcal{U}$,

$$|F(u, \omega) - F(v, \omega)| \leq L_F(\omega)||u - v||_{L^2}. \tag{16}$$

Proof By Lemma 7, we know that the map $t \mapsto y_u(t, \omega)$ is continuous, and by $(A.2)$ so is the function f on the first variable. Then $\sup_{t \in [0,T]} f(y_u(t, \omega), \omega)$ is actually a maximum, so F is well defined.

Now, by Remark 8 for every $\omega \in \Omega$ we have,

$$\begin{aligned} |F(u, \omega) - F(v, \omega)| &\leq \max_{t \in [0,T]} |f(y_u(t, \omega)) - f(y_v(t, \omega))| \\ &\leq L_f(\omega) \max_{t \in [0,T]} |y_u(t, \omega) - y_v(t, \omega)| \\ &\leq L_f(\omega) C||u - v||_{L^2}, \end{aligned} \tag{17}$$

where C is given by Lemma 9, then $u \mapsto F(u, \omega)$ is continuous for all $\omega \in \Omega$ and (16) holds with $L_F(\omega) := L_f(\omega)C$. By $(A.2)$ we deduce that L_F belongs to $L^1(\Omega)$.

Now we have to prove that for all $u \in \mathcal{U}$ the application $\omega \mapsto F(u, \omega)$ is measurable. We fix $u \in \mathcal{U}$, for all $t \in [0, T]$ we have $\omega \mapsto y_u(t, \omega)$ is measurable by Lemma 7, since f is a Carateódory function and so it is jointly measurable, we deduce that $\omega \mapsto f(y_u(t, \omega), \omega)$ is measurable. Then the map $\psi : \Omega \to \mathbb{R}$ defined as

$$\psi(\omega) = \sup_{t \in [0,T] \cap \mathbb{Q}} f(y_u(t, \omega), \omega) \tag{18}$$

is measurable. Since f is continuous in the first variable and the map $t \mapsto y_u(t, \omega)$ is continuous for each $\omega \in \Omega$, by the density of rational numbers we obtain

$$\psi(\omega) = F(u, \omega). \tag{19}$$

We conclude that $F(u, \cdot)$ is measurable. □

Corollary 11 *If* $(A.1)$ *and* $(A.2)$ *hold, the functional* J *is well defined and the problem* (P) *has a solution.*

Proof Since the function $F(u, \cdot)$ is measurable, it is clear that the functional J is well defined. Now in order to prove the existence of minimizer we start proving that J is a proper function. For any $u \in \mathcal{U}$, we have

$$F(u, \omega) = \max_{t \in [0,T]} f(y_u(t, \omega), \omega) \geq f(x + \phi(\omega), \omega). \tag{20}$$

Also, by $(A.2)$ we obtain that $\omega \mapsto f(x + \phi(\omega), \omega)$ is integrable, in fact by Remark 8 we have

$$|f(x + \phi(\omega), \omega)| \leq |f(\bar{x}, \omega)| + L_f(\omega)|x + \phi(\omega) - \bar{x}|, \tag{21}$$

and the right hand side (r.h.s.) belongs to $L^1(\Omega)$. We can conclude that $J(u) > -\infty$ for all $u \in \mathcal{U}$. Now to see that J is not identically $+\infty$, by Lemma 7 we obtain that for any $u \in \mathcal{U}$ and for all $t \in [0, T]$, and $\omega \in \Omega$,

$$\begin{aligned} |f(y_u(t,\omega),\omega)| &\leq |f(x+\phi(\omega),\omega)| + L_f(\omega)|y_u(t,\omega) - y_u(0,\omega)| \\ &\leq |f(x+\phi(\omega),\omega)| + L_f(\omega)2M_y. \end{aligned} \tag{22}$$

Since the r.h.s. is clearly integrable, we deduce that $J(u)$ is finite.

By the previous proposition, for $u, v \in \mathcal{U}$ we obtain

$$|J(u) - J(v)| \leq \mathbb{E}^{\mathbb{P}}[L_F]||u - v||_{L^2}, \tag{23}$$

then J is Lipschitz continuous on $\mathcal{U}$. Since for each $\omega \in \Omega$, the function $f(\cdot,\omega)$ is convex, and the state equation is linear, we can conclude that the function $F(\cdot,\omega)$ is convex on $\mathcal{U}$. By the linearity of the expectation, J is a convex function on $\mathcal{U}$. Therefore, J is weakly lower semicontinuous on $\mathcal{U}$. By the compactness of the set U, the set of controls $\mathcal{U}$ is closed and bounded in $L^2[0, T]$. Thus, there exists a minimizer of J in $\mathcal{U}$. □

3.2 Sample Average Approximation

The idea of this section is to approximate the problem presented in the above section using a sample average approximation. Under the same assumptions, let $\{\omega_1, \ldots, \omega_M\}$ be an independent $\mathbb{P}$-distributed random draw, we consider the state given by

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} y(t,\omega_i) = g(t, y(t,\omega_i), u(t), \omega_i), & t \in [0, T] \\ y(0,\omega_i) = x + \phi(\omega_i), \end{cases} \quad \forall i = 1, \ldots, M, \tag{24}$$

and the cost functional $J_M : \mathcal{U} \to \mathbb{R}$ defined as

$$J_M(u) := \frac{1}{M} \sum_{i=1}^{M} F(u, \omega_i). \tag{25}$$

Then we obtain the following optimal control problem

$$\min J_M(u); \quad u \in \mathcal{U}. \tag{P_M}$$

The problem (P_M) is a deterministic optimal control problem, and by the assumptions made it is clear that it is well-posed. In this case we also have the existence of minimizers.

Proposition 12 *Under assumptions $(A.1)$ and $(A.2)$, for all $M \in \mathbb{N}$ there exists a solution of (P_M).*

Proof For each ω_i, $i = 1, \ldots, M$, the function $u \mapsto F(u, \omega_i)$ is continuous and convex, then also is J_M, and that implies that J_M is a weakly lower semicontinuous function of u. Since the set $\mathcal{U}$ is closed and bounded in $L^2[0, T]$, there exists a minimizer of J_M in $\mathcal{U}$. □

3.3 Discrete Time Approximation

In the next section we will justify how we can approximate (P) by (P_M) when $M \to \infty$. But in order to solve (P_M) we define a discrete time approximate problem, which will be suitable solved by a descent numerical method.

We fix $M \in \mathbb{N}$ and $\{\omega_1, \ldots, \omega_M\}$ an independent $\mathbb{P}$-distributed random draw. We divide the interval $[0, T]$ into N subintervals with common length $h = T/N$ and we restrict the controls to be sectionally constant. So, the set of discrete controls is

$$\mathcal{U}^h = \{u \in \mathcal{U} : u \text{ is constant in } [kh, (k+1)h),\ k = 0, ..., N-1\}. \tag{26}$$

A discrete policy u is identified as $\{u^n\}_{n=0}^{N-1}$, $u^n \in U \subset \mathbb{R}^m$, so $\mathcal{U}^h$ can be identified as $U^N \subset \mathbb{R}^{m\times N}$.

We introduce an approximated discrete time system. For $u \in \mathcal{U}^h$ we define the response y_u of the discrete time system by the recursive formula

$$\begin{cases} y_u^{n+1}(\omega_i) = y_u^n(\omega_i) + hg(t_n, y_u^n(\omega_i), u^n, \omega_i), & n = 0, ..., N-1, \\ y_u^0(\omega_i) = x + \phi(\omega_i), \end{cases} \tag{27}$$

for $i = 1, \ldots, M$. Defining the functional $J_M^h : \mathcal{U}^h \to \mathbb{R}$, as

$$J_M^h(u) := \frac{1}{M}\sum_{i=1}^{M} \max_{n=0,\ldots,N} f(y_u^n(\omega_i), \omega_i),$$

the discrete time optimal control problem is

$$\min J_M^h(u); \quad u \in \mathcal{U}^h. \tag{P_M^h}$$

Clearly, the minimization problem has a solution since J_M^h is continuous over the compact set $\mathcal{U}^h$.

4 Convergence

In this section we analyze the relationship between the previous defined problems. We start showing that the objective functions of (P_M) epi-converge to the objective function of (P), and we also show that the accumulation points of minimizers of (P_M) are minimizers of (P) and the value of (P_M) converges to the value of (P). After that, we study the discretization in time, and we demonstrate that the value of the problem (P_M^h) converges to the value of (P_M) for all M. We finish the section proving that there exists a sequence of discrete problems $(P_M^{h_M})$ where the accumulation points of the sequence of minimizers are optimal solutions of (P).

By the Strong Law of Large Numbers, we know that $J_M(u) \to J(u)$ as $M \to \infty$, for almost all $u \in \mathcal{U}$. Following the ideas of [19] we can show a strong convergence result.

Theorem 13 *Assume that $(A.1)$ and $(A.2)$ hold. Let $\{\omega_1, \ldots, \omega_M\}$ be an independent $\mathbb{P}$-distributed random draw, then $J_M \overset{epi}{\to} J$ as $M \to \infty$, almost surely in $\mathcal{U}$.*

Proof By Lemma 10, the function F is Carathéodory and that implies that F is $B \otimes \mathcal{A}$ measurable and continuous in u for each $\omega \in \Omega$. So, we can conclude that F is a random lower semicontinuous function. We also have

$$F(u, \omega) = \max_{t\in[0,T]} f(y_u(t, \omega), \omega) \geq f(y_u(0, \omega), \omega) = f(x + \phi(\omega), \omega). \tag{28}$$

Thus, as in the proof of Corollary 11 we can see that the map $\omega \mapsto f(x+\phi(\omega), \omega)$ belongs to $L^1(\Omega)$. Then, by Theorem 4 we can conclude that $J_M \overset{epi}{\to} J$ a.s. in $\mathcal{U}$. $\square$

Under the same assumptions, we have a convergence result about the minimizers and the value of (P_M) and (P).

Theorem 14 *Let $\{u_M\}_{M\in\mathbb{N}}$ be a sequence of optimal controls for (P_M) and $\bar{u}$ an accumulation point, i.e. $\bar{u} = \lim_{M\in K} u_M$ where K is an infinite subset of $\mathbb{N}$. Then, $\bar{u}$ is a minimizer of (P) and in addition $\lim_{M\in K} J_M(u_M) = J(\bar{u})$.*

Proof Since the functions J_M and J are convex, all the minimizers are global minimizers. We can apply Theorem 5 because J_M is continuous for all $M \in \mathbb{N}$ and by Theorem 13, $J_M \overset{epi}{\to} J$ a.s. in $\mathcal{U}$. □

In order to analyze the relationship between (P_M^h) and (P), we start by comparing the continuous and the discrete states associated to a given discrete control $u \in \mathcal{U}^h$. We note that $\mathcal{U}^h \subset \mathcal{U}$. In what follows we assume $M \in \mathbb{N}$ and the independent $\mathbb{P}$-distributed random draw $\{\omega_1, \ldots, \omega_M\}$ are fixed.

Lemma 15 *Let $u \in \mathcal{U}^h$ be a given control. Let y_u be the solution of (2) and (y_u^n) the solution of (27), both associated to u. Then, for all ω_i, $i = 1, \ldots, M$ there exists $C_y(\omega_i)$ such that*

$$\max_{n=0,\ldots,N} |y_u(t_n, \omega_i) - y_u^n(\omega_i)| \le C_y(\omega_i)h. \tag{29}$$

Proof On one hand, denoting M_U the bound of the compact set $U \subset \mathbb{R}^m$ and defining $L_y := M_A M_y + M_B M_U + M_C$, where M_y is given by Lemma 7, we obtain

$$\sup_{t_n \le t < t_{n+1}} |y_u(t, \omega) - y_u(t_n, \omega)| \le L_y h. \tag{30}$$

On the other hand, a straightforward calculation gives, for all $n = 0, \ldots, N-1$,

$$\begin{aligned} |y_u(t_{n+1}) - y_u^{n+1}| \le\ & |y_u(t_n) - y_u^n| + \textstyle\int_{t_n}^{t_{n+1}} L_A h|y_u(s)| + M_A|y_u(s) - y_u^n|\mathrm{d}s \\ & + \textstyle\int_{t_n}^{t_{n+1}} L_B h|u^n|\mathrm{d}s + \int_{t_n}^{t_{n+1}} L_C h\mathrm{d}s, \end{aligned} \tag{31}$$

where the argument ω_i is omitted for notational convenience. Thus,

$$\begin{aligned} |y_u(t_{n+1}) - y_u^{n+1}| \le\ & |y_u(t_n) - y_u^n| + h^2 M_y L_A + hM_A[L_y h + |y_u(t_n) - y_u^n|] \\ & + h^2 L_B M_U + h^2 L_C \\ \le\ & (1 + hM_A)|y_u(t_0) - y_u^0| + \textstyle\sum_{k=0}^{n-1}(1 + hM_A)^k h^2 C_2, \end{aligned} \tag{32}$$

where $C_2(\omega_i) := M_y L_A(\omega_i) + M_A L_y + L_B(\omega_i)M_U + L_C(\omega_i)$. Therefore, the result follows by taking $C_y(\omega_i) := Te^{TM_A}C_2(\omega_i)$. □

Here one of the main result of this section, we state that the value of the discrete time problem (P_M^h) converges to the value of (P_M).

Theorem 16 *Let $\bar{u}^h$ be an optimal control for (P_M^h), and $\bar{u}$ an optimal solution for (P_M), then*

$$\lim_{h\downarrow 0} J_M^h(\bar{u}^h) = J_M(\bar{u}). \tag{33}$$

Proof Let N_h be a natural number and $h = \frac{T}{N_h}$. Since $\mathcal{U}^h \subset \mathcal{U}$, any $u^h \in \mathcal{U}^h$ is an admissible control for (P_M), then

$$\begin{aligned} |J_M(u^h) - J_M^h(u^h)| &\le \sum_{i=1}^{M} |\max_{t\in[0,T]} f(y_{u^h}(t,\omega_i),\omega_i) - \max_{n=0,\dots,N_h} f(y_{u^h}^n(\omega_i),\omega_i)| \\ &\le \sum_{i=1}^{M} |\max_{t\in[0,T]} f(y_{u^h}(t,\omega_i),\omega_i) - \max_{n=0,\dots,N_h} f(y_{u^h}(t_n,\omega_i),\omega_i)| \\ &\quad + \sum_{i=1}^{M} |\max_{n=0,\dots,N_h} f(y_{u^h}(t_n,\omega_i),\omega_i) - \max_{n=0,\dots,N_h} f(y_{u^h}^n(\omega_i),\omega_i)| \\ &\le \sum_{i=1}^{M} |L_f(\omega_i)L_y|h + \sum_{i=1}^{M} |L_f(\omega_i)C_y(\omega_i)|h. \end{aligned} \tag{34}$$

Where the last inequality holds by (29), (30), and Remark 8. Since the points $\{\omega_1,\dots,\omega_M\}$ are fixed, we can conclude that there exists $\bar{C}_M > 0$, such that

$$|J_M(u^h) - J_M^h(u^h)| \le \bar{C}_M h. \tag{35}$$

Now, by the optimality of $\bar{u}$ we have

$$J_M(\bar{u}) \le J_M(\bar{u}^h) \le J_M(\bar{u}^h) - J_M^h(\bar{u}^h) + J_M^h(\bar{u}^h) \le J_M^h(\bar{u}^h) + \bar{C}_M h. \tag{36}$$

On the other hand, for all $\varepsilon > 0$, by the density of uniformly continuous functions in $L^2[0,T]$ (which in turn can be approximated by uniform step functions), there exist N_ε such that for all $N_h > N_\varepsilon$ there exists $u^h \in \mathcal{U}^h$ such that

$$||\bar{u} - u^h||_{L^2} \le \varepsilon. \tag{37}$$

Since $F(\cdot,\omega_i)$ is Lipschitz, so is J_M and being $\bar{u}^h$ optimal for (P_M^h), we obtain

$$\begin{aligned} J_M^h(\bar{u}^h) &\le J_M^h(u^h) - J_M(u^h) + J_M(u^h) \\ &\le \bar{C}_M h + J_M(u^h) - J_M(\bar{u}) + J_M(\bar{u}) \\ &\le \bar{C}_M h + L_{J_M}\varepsilon + J_M(\bar{u}). \end{aligned} \tag{38}$$

The result follows by (36) and (38). □

Theorem 17 *For each $M \in \mathbb{N}$ there exists $h_M > 0$ with $h_M \to 0$ as $M \to \infty$ such that if $\bar{u}_M^{h_M} \in \mathcal{U}^{h_M}$ is an optimal control for $(P_M^{h_M})$, then any accumulation point of the sequence $\{\bar{u}_M^{h_M}\}$ is an optimal control for (P).*

Proof By Theorem 16, for each $M \in \mathbb{N}$ we can choose $h_M > 0$ such that

$$J_M^{h_M}(\bar{u}_M^{h_M}) \le J_M(\bar{u}_M) + \frac{1}{M}. \tag{39}$$

We can also choose $h_M \to 0$ such that $\bar{C}_M h_M \to 0$ as $M \to \infty$, where $\bar{C}_M$ is given by (35).

Now let $\bar{u}$ be an accumulation point of $\{\bar{u}_M^{h_M}\}$, i.e. there exists a subsequence of $\{\bar{u}_M^{h_M}\}$, still denoted $\{\bar{u}_M^{h_M}\}$ such that $\bar{u}_M^{h_M} \to \bar{u}$. Then, by the epi-convergence of J_M to J given by Theorem 13, (39) and (35) we obtain

$$\begin{aligned} J(\bar{u}) &\le \liminf J_M(\bar{u}_M^{h_M}) - J_M^h(\bar{u}_M^{h_M}) + J_M^h(\bar{u}_M^{h_M}) \\ &\le \liminf \bar{C}_M h_M + J_M(\bar{u}_M) + \tfrac{1}{M} \\ &= \inf_{u\in\mathcal{U}} J(u) \end{aligned} \tag{40}$$

where the equality holds by Theorem 14. Thus, $\bar{u}$ is an optimal control. □

Remark 18 By the nature of the proofs of Theorem 13 and Theorem 16, we are not able to deduce an error estimate bound for the approximations that we consider. Nevertheless, if we only consider the discretization in time, i.e., we focus in the approximation of (P_M) via (P_M^h), it is possible to obtain an error estimate of order $\sqrt{h}$ for the value function, following [12].

5 Optimality Conditions

In this section, based on [13] we state a set of optimality conditions for all the problems.

5.1 Uncertain Minimax Optimal Control Problem

We start analyzing the directional differentiability of the functional J. We recall that $T_{\mathcal{U}}(u)$ is the tangent cone to $\mathcal{U}$ in u (see [9]).

Proposition 19 *Under assumptions* $(A.1)$ *and* $(A.2)$*, the function* J *is directionally differentiable at any* $u \in \mathcal{U}$ *and the directional derivative in a direction* $v \in T_{\mathcal{U}}(u)$ *is given by*

$$J'(u; v) = \mathbb{E}^{\mathbb{P}} \sup_{t \in C_{u,\omega}} \langle \nabla f(y_u(t,\omega),\omega), z_v(t,\omega) \rangle \,, \tag{41}$$

where $C_{u,\omega}$ *is the set of critical times*

$$C_{u,\omega} = \underset{t \in [0,T]}{\mathit{argmax}}\, f(y_u(t,\omega),\omega), \tag{42}$$

and z_v *solves the differential equation*

$$\begin{cases} \dfrac{\mathrm{d}z}{\mathrm{d}t}(t,\omega) = A(t,\omega)z(t,\omega) + B(t,\omega)v(t), & t \in [0,T] \\ z(0,\omega) = 0, \end{cases} \tag{43}$$

for all $\omega \in \Omega$.

Proof Fix $\omega \in \Omega$, by [13, Proposition 2.1], $F(\cdot,\omega)$ is directionally differentiable in $\mathcal{U}$ and the directional derivative in a direction $v \in T_{\mathcal{U}}(u)$ is given by

$$F'(u,\omega; v) = \sup_{t \in C_{u,\omega}} \langle \nabla f(y_u(t,\omega),\omega), z_v(t,\omega) \rangle \,, \tag{44}$$

where $C_{u,\omega}$ is given by (42) and $z_v(\cdot,\omega)$ solves (43).

For any $h > 0$, by Lemma 10 we have

$$\left| \frac{F(u+hv,\omega) - F(u,\omega)}{h} \right| \le \frac{L_F(\omega)h||v||_{L^2}}{h} \tag{45}$$

Since $L_F \in L^1(\Omega)$ we can apply the Dominated Convergence Theorem and conclude that

$$J'(u; v) = \mathbb{E}^{\mathbb{P}} \sup_{t \in C_{u,\omega}} \langle \nabla f(y_u(t,\omega),\omega), z_v(t,\omega) \rangle \,. \tag{46}$$

□

Following [13], for each $\omega \in \Omega$ the solution of (43) is given by

$$z_v(t,\omega) = \int_0^t S_{ts}(\omega)B(s,\omega)v(s)\mathrm{d}s, \tag{47}$$

where the matrix S_{ts} is the solution of the system

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} S_{ts}(\omega) = A(t,\omega)S_{ts}(\omega), & t \in [s,T] \\ S_{ss}(\omega) = I. \end{cases} \tag{48}$$

Now, the directional derivative can be written as

$$J'(u;v) = \mathbb{E}^{\mathbb{P}} \sup_{t \in C_{u,\omega}} \left\langle \nabla f(y_u(t,\omega)), \int_0^t S_{ts}(\omega)B(s,\omega)v(s)\mathrm{d}s \right\rangle. \tag{49}$$

Defining for each $u \in \mathcal{U}$, $t \in [0,T]$ and $\omega \in \Omega$, the element of $L^2[0,T]$

$$q_{u,t}(s,\omega) := I_t(s)B^\top(s,\omega)S_{ts}(\omega)^\top \nabla f(y_u(t,\omega),\omega), \quad \forall s \in [0,T],$$

where $I_t(s)$ is equal to 1 if $s \le t$ and 0 otherwise, we can rewrite (49) as

$$J'(u;v) = \mathbb{E}^{\mathbb{P}} \sup_{t \in C_{u,\omega}} \langle q_{u,t}(\omega), v \rangle. \tag{50}$$

Therefore, we have a first order optimality condition based on the fact that if u is an optimizer then every directional derivative is non-negative for every direction in $T_{\mathcal{U}}(u)$, which is also a sufficient condition since the function is convex ([9]). The following result is the analogous of [13, Theorem 2.1].

Theorem 20 *Assume* $(A.1)$ *and* $(A.2)$ *hold. Let* $u \in \mathcal{U}$*, then* u *is optimal if and only if*

$$\min_{v \in T_{\mathcal{U}}(u)} \mathbb{E}^{\mathbb{P}}[\sup_{t \in C_{u,\omega}} \langle q_{u,t}(\omega), v \rangle] = 0. \tag{51}$$

We also have necessary optimality conditions which do not involve the computation of the set of critical times, see [13, Theorem 2.2]

Theorem 21 *Condition* (51) *implies*

$$\inf_{v \in T_{\mathcal{U}}(u)} \mathbb{E}^{\mathbb{P}}[\sup_{t \in [0,T]} \langle q_{u,t}(\omega), v \rangle] = 0, \tag{52}$$

$$\inf_{v \in T_{\mathcal{U}}(u)} \mathbb{E}^{\mathbb{P}}[\sup_{t \in [0,T]} \{ f(y_u(t,\omega),\omega) - F(u,\omega) + \langle q_{u,t}(\omega), v \rangle \}] = 0, \tag{53}$$

and for any $\rho > 0$*,*

$$\inf_{v \in T_{\mathcal{U}}(u)} \mathbb{E}^{\mathbb{P}}[\sup_{t \in [0,T]} \{ f(y_u(t,\omega),\omega) - F(u,\omega) + \langle q_{u,t}(\omega), v \rangle \}] + \frac{\rho}{2} \|v\|^2 = 0. \tag{54}$$

5.2 Sample Average Approximation

Along this section, we assume that $M \in \mathbb{N}$ and the independent $\mathbb{P}$-distributed random draw $\{\omega_1, \ldots, \omega_M\}$ are fixed. Following the same ideas of the above section, we present optimality conditions for problem (P_M).

The following result is straightforward from [13, Theorem 2.1] and (44).

Theorem 22 *Under assumptions* $(A.1)$ *and* $(A.2)$, $u \in \mathcal{U}$ *is an optimal control for* (P_M) *if and only if*

$$\inf_{v \in T_{\mathcal{U}}(u)} \frac{1}{M} \sum_{i=1}^{M} \sup_{t \in C_{u,\omega_i}} \langle q_{u,t}(\omega_i), v \rangle = 0. \tag{55}$$

We also have the analogous necessary optimality conditions.

Theorem 23 *If* $u \in \mathcal{U}$ *is an optimal control for* (P_M) *then,*

$$\inf_{v \in T_{\mathcal{U}}(u)} \frac{1}{M} \sum_{i=1}^{M} \sup_{t \in [0,T]} \langle q_{u,t}(\omega_i), v \rangle = 0, \tag{56}$$

$$\inf_{v \in T_{\mathcal{U}}(u)} \frac{1}{M} \sum_{i=1}^{M} \sup_{t \in [0,T]} \left\{ f(y_u(t,\omega_i),\omega_i) - F(u,\omega_i) + \langle q_{u,t}(\omega_i), v \rangle \right\} = 0, \tag{57}$$

and for any $\rho > 0$,

$$\inf_{v \in T_{\mathcal{U}}(u)} \frac{1}{M} \sum_{i=1}^{M} \sup_{t \in [0,T]} \left\{ f(y_u(t,\omega_i),\omega_i) - F(u,\omega_i) + \langle q_{u,t}(\omega_i), v \rangle \right\} + \frac{\rho}{2} \|v\|^2 = 0. \tag{58}$$

Proof Let u be optimal. Since $v = 0$ is an admissible direction, by Theorem 22 we have

$$0 = \inf_{v \in T_{\mathcal{U}}(u)} \frac{1}{M} \sum_{i=1}^{M} \sup_{t \in C_{u,\omega_i}} \langle q_{u,t}(\omega_i), v \rangle \leq \inf_{v \in T_{\mathcal{U}}(u)} \frac{1}{M} \sum_{i=1}^{M} \sup_{t \in [0,T]} \langle q_{u,t}(\omega_i), v \rangle \leq 0. \tag{59}$$

So (55) implies (56). From the definition of C_{u,ω_i}, for all $i = 1, \dots, M$,

$$\begin{aligned} \sup_{t \in C_{u,\omega_i}} \langle q_{u,t}(\omega_i), v \rangle &\leq \sup_{t \in [0,T]} \left\{ f(y_u(t,\omega_i),\omega_i) - F(u,\omega_i) + \langle q_{u,t}(\omega_i), v \rangle \right\} \\ &\leq \sup_{t \in [0,T]} \langle q_{u,t}(\omega_i), v \rangle . \end{aligned} \tag{60}$$

By (55) and (56), we obtain (57). Analogously, (58) follows. □

5.3 Discrete Time Approximation

For the sake of completeness, we adapt some results from [13, Chapter 3].

Proposition 24 *Given a discrete policy* $u = \{u^n\}_{n=0}^{N-1}$, *the functional* J_M^h *is directionally differentiable at* u *and for any* $v \in T_{\mathcal{U}^h}(u)$ *we have*

$$J_M^{h\,\prime}(u,v) = \frac{1}{M} \sum_{i=1}^{M} \max_{n \in C_{u,\omega_i}} \langle \nabla f(y_u^n(\omega_i),\omega_i), z_v^n(\omega_i) \rangle \tag{61}$$

where $C_{u,\omega_i} = \operatorname{argmax} \{ f(y_u^n(\omega_i),\omega_i) : 0 \leq n \leq N \}$ *is the set of critical times, and* z_v *solves, for* $i = 1, \dots, M$, *the following system of difference equations*

$$\begin{cases} z^{n+1}(\omega_i) = [I + hA(t_n,\omega_i)] z^n(\omega_i) + hB(t_n,\omega_i) v^n, \quad n = 0, \dots, N-1, \\ z^0(\omega_i) = 0. \end{cases} \tag{62}$$

The solution of (62) can be written as a function of v, in fact

$$z_v^n(\omega_i) = \sum_{j=0}^{n-1} S_{n-1,j}(\omega_i) v^j \tag{63}$$

where S satisfies $S_{n+1,j}(\omega_i) = [I + hA(t_{n+1}, \omega_i)]S_{n,j},\ 0 \le j \le n$ and $S_{jj}(\omega_i) = hB(t_j, \omega_i),\ \forall j \ge 0,\ \ i = 1, \dots, M$. Defining,

$$q_{u,n}^j(\omega_i) := \begin{cases} 0 & \forall j \ge n, \\ S_{n-1,j}^\top(\omega_i) \nabla f(y_u^n(\omega_i), \omega_i) & \forall j < n, \end{cases} \tag{64}$$

for all $i = 1, \dots, M$, we can conclude

$$J_M^{h\,\prime}(u; v) = \frac{1}{M} \sum_{i=1}^{M} \max_{n \in C_{u,\omega_i}} \sum_{j=0}^{n-1} \left\langle q_{u,n}^j(\omega_i), v^j \right\rangle = \frac{1}{M} \sum_{i=1}^{M} \max_{n \in C_{u,\omega_i}} \left\langle q_{u,n}(\omega_i), v \right\rangle,$$

where $q_{u,n}(\omega_i)$ is the matrix with columns $q_{u,n}^j(\omega_i)$, v is identified with the matrix of columns v^j and the last product is defined as $\langle q_{u,n}(\omega_i), v\rangle := tr(q_{u,n}^\top(\omega_i) v)$.

Now we easily obtain the first optimality condition for the discrete problem.

Theorem 25 *Let $u \in \mathcal{U}^h$ and define $\mathcal{U}_u^h := \mathcal{U}^h - u$. Then u is an optimal control for (P_M^h) if and only if*

$$\min_{v \in \mathcal{U}_u^h} \frac{1}{M} \sum_{i=1}^{M} \max_{n \in C_{u,\omega_i}} \left\langle q_{u,n}(\omega_i), v \right\rangle = 0. \tag{65}$$

Note that the minimization over $\mathcal{U}_u^h$ in (65) is equivalent to the minimization over $T_{\mathcal{U}^h}(u)$. In fact, the set of controls $\mathcal{U}^h$ is convex and since the sample $\{\omega_1, \dots, \omega_M\}$ is fixed, the functionals $q_{u,n}(\omega_i)$ are bounded in $L^2[0, T]$.

In order to develop a convergent numerical method, we propose an analogous version of the optimality condition (58) which was only a necessary one in the continuous-time framework. However, in the discrete case, condition (66) is not only necessary, but also sufficient.

Theorem 26 *Condition (65) is equivalent to*

$$\min_{v \in \mathcal{U}_u^h} \frac{1}{M} \sum_{i=1}^{M} \max_{n=0,..,N} \left\{ f(y_u^n(\omega_i), \omega_i) - F^h(u, \omega_i) + \left\langle q_{u,n}(\omega_i), v \right\rangle \right\} + \frac{\rho}{2} ||v||^2 = 0, \tag{66}$$

for any $\rho > 0$ where $F^h : \mathcal{U}^h \times \Omega \to \mathbb{R}$ is defined as

$$F^h(u, \omega) := \max_{n=0,\dots,N} f(y_u^n(\omega), \omega). \tag{67}$$

Proof As in Theorem 23 we can see that (65) implies (66). Now assume that (66) holds while (65) does not hold. Then, there exists $v \in \mathcal{U}_u^h$ such that

$$\frac{1}{M} \sum_{i=1}^{M} \max_{n \in C_{u,\omega_i}} \left\langle q_{u,n}(\omega_i), v \right\rangle < 0. \tag{68}$$

Let $\bar{\lambda} > 0$ be small enough such that

$$\frac{1}{M}\sum_{i=1}^{M}\max_{n\in C_{u,\omega_i}}\left\langle q_{u,n}(\omega_i), v\right\rangle + \frac{\rho}{2}\bar{\lambda}||v||^2 < 0. \tag{69}$$

Then, for all $0 < \lambda < \bar{\lambda}$ we have

$$\frac{1}{M}\sum_{i=1}^{M}\max_{n\in C_{u,\omega_i}}\left\langle q_{u,n}(\omega_i), \lambda v\right\rangle + \frac{\rho}{2}||\lambda v||^2 < 0. \tag{70}$$

Now, for all $i = 1,\ldots,M$ we define the functions $a_i, b_i : \mathcal{U}_u^h \to \mathbb{R}$ by

$$\begin{aligned} a_i(v) &:= \max_{n\in C_{u,\omega_i}}\left\langle q_{u,n}(\omega_i), v\right\rangle + \tfrac{\rho}{2}||v||^2,\\ b_i(v) &:= \max_{n\notin C_{u,\omega_i}}\left\{f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \left\langle q_{u,n}(\omega_i), v\right\rangle\right\} + \tfrac{\rho}{2}||v||^2. \end{aligned} \tag{71}$$

It is straightforward that a_i and b_i are continuous for all $i = 1,\ldots,M$. Also,

$$\lim_{\lambda\downarrow 0} a_i(\lambda v) = 0, \tag{72}$$

$$\lim_{\lambda\downarrow 0} b_i(\lambda v) = \max_{n\notin C_{u,\omega_i}}\left\{f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i)\right\} = -\delta_i < 0. \tag{73}$$

We define $\delta := \min\{\delta_1,\ldots,\delta_M\} > 0$. Then there exists λ^b such that for all $0 < \lambda \le \lambda^b$ we have

$$b_i(\lambda v) < -\frac{\delta}{2}, \tag{74}$$

for all $i = 1,\ldots,M$. By (72) there exists λ^a such that for all $0 < \lambda \le \lambda^a$,

$$-\frac{\delta}{2} < a_i(\lambda v) < \frac{\delta}{2}, \tag{75}$$

for all $i = 1,\ldots,M$. Thus, if $0 < \lambda < \min\{\lambda^a,\lambda^b,\bar{\lambda}\}$ then $b_i(\lambda v) < a_i(\lambda v)$ for all $i = 1,\ldots,M$. Therefore,

$$\max_{n=0,..,N}\left\{f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \left\langle q_{u,n}(\omega_i), \lambda v\right\rangle\right\} + \frac{\rho}{2}||\lambda v||^2 = a_i(\lambda v), \tag{76}$$

for all $i = 1,\ldots,M$. So by the definition of a_i and (70) we have

$$\begin{aligned} \tfrac{1}{M}\sum_{i=1}^{M}\max_{n=0,..,N}\left\{f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \left\langle q_{u,n}(\omega_i), \lambda v\right\rangle\right\} + \tfrac{\rho}{2}||\lambda v||^2\\ = \tfrac{1}{M}\textstyle\sum_{i=1}^{M} a_i(\lambda v) < 0, \end{aligned} \tag{77}$$

which contradicts (66). So we conclude that (65) holds. □

Remark 27 Note that condition (65) is also equivalent to

$$\min_{v\in\mathcal{U}_u^h}\frac{1}{M}\sum_{i=1}^{M}\max_{n=0,..,N}\left\{f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \left\langle q_{u,n}(\omega_i), v\right\rangle\right\} = 0. \tag{78}$$

Indeed, for all $v \in \mathcal{U}_u^h$ and ω_i, $i = 1,\ldots,M$ we have

$$\begin{aligned} &\max_{n=0,..,N}\left\{f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \left\langle q_{u,n}(\omega_i), v\right\rangle\right\}\\ &\quad\le \max_{n=0,..,N}\left\{f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \left\langle q_{u,n}(\omega_i), v\right\rangle\right\} + \tfrac{\rho}{2}||\lambda v||^2 \le 0. \end{aligned} \tag{79}$$

Then (78) implies (66) which is equivalent to (65). The converse follows as in Theorem 23.

6 Algorithm

In this section, following the lines of [13] we present a numerical method to solve the discrete time problem (P_M^h), which is based on the optimality condition (66). Let define $\theta : U^N \to \mathbb{R}$ and $\eta : U^N \to \mathbb{R}^{m\times N}$ (where we identify $\mathcal{U}^h \equiv U^N$, $\mathcal{U}_u^h \equiv U_u^N$) as

$$\theta(u) := \min_{v\in U_u^N} \frac{1}{M}\sum_{i=1}^{M} \max_{n=0,..,N} \left\{ f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \langle q_{u,n}(\omega_i), v\rangle \right\} + \frac{\rho}{2}||v||^2 \quad (80)$$

$$\eta(u) := \operatorname*{argmin}_{v\in U_u^N} \sum_{i=1}^{M} \max_{n=0,...,N} \left\{ f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \langle q_{u,n}(\omega_i), v\rangle \right\} + \frac{\rho}{2}\|v\|^2 . \quad (81)$$

An admissible control u satisfying (66) is optimal; otherwise, the minimizer in (66) gives a descent direction of the functional J_M^h, in fact

$$J_M^{h\,\prime}(u;\eta(u)) = \frac{1}{M}\sum_{i=1}^{M} \max_{n\in C_{u,\omega_i}} \langle q_{u,n}(\omega_i), \eta(u)\rangle \le \theta(u) < 0. \quad (82)$$

Taking that into account, we introduce an algorithm that computes at each step a descent direction solving (66) and performs an Armijo line search. Using condition (66) has two main advantages. On the one hand, the supremum is computed over the whole set of times. The application $u \mapsto C_u$ is not always continuous as a set-valued function, which is a drawback in the aim to obtain convergence properties. On the other hand, the quadratic term regularizes the operator to be minimized, which turns to be strongly convex so it has unique solution. Therefore, the functions θ and η are continuous (see [20, Chapter 5]).

Algorithm 1

Step 1: Choose the parameters $\alpha, \beta \in (0,1)$ and $\rho > 0$. Set $k = 1$ and choose the initial point $u_1 \in U^N$.

Step 2: Compute:

$$y_{u_k}^n(\omega_i),\ f(y_{u_k}^n(\omega_i),\omega_i),\quad n = 0,\dots,N,\quad i = 1,\dots,M$$

$$F^h(u_k,\omega_i) = \max_{n=0,..,N} f(y_{u_k}^n(\omega_i),\omega_i),\ i = 1,\dots,M.$$

Step 3: Compute $\theta(u_k)$ and $\eta(u_k)$ given by (80) and (81), respectively.

Step 4: If $\theta(u_k) = 0$, Stop (u_k satisfies the optimality condition). Else, find the maximum $\lambda_k = \beta^j,\ \ j \in \mathbb{N}_0$, such that

$$J_M^h(u_k + \lambda_k\,\eta(u_k)) < J_M^h(u_k) + \alpha\lambda_k\,\theta(u_k).$$

Step 5: Set $u_{k+1} = u_k + \lambda_k\,\eta(u_k)$, $k = k+1$ and restart Step 2.

In practice, problems (80) and (81) are solved by introducing auxiliary variables $\xi_i \in \mathbb{R},\ i = 1,\dots,M$, and considering the quadratic programs

$$\begin{aligned} &\min \tfrac{\rho}{2}\|v\|^2 + \tfrac{1}{M}\sum_{i=1}^{M}\xi_i \\ \text{s.t.}\ &\xi_i \ge f(y_u^n(\omega_i),\omega_i) - F^h(u,\omega_i) + \langle q_{u,n}(\omega_i), v\rangle, \\ &v \in U_u^N, \xi_i \in \mathbb{R},\quad n = 0,...,N,\quad i = 1,\dots,M \end{aligned} \quad (83)$$

which can be solved efficiently by standard algorithms.

Following the arguments of [13, Theorem 4.1] we obtain the next convergence result.

Theorem 28 *Let $\{u_k\}$ be the sequence generated by Algorithm 1. Then, either $\{u_k\}$ finishes at a minimizer or it is infinite and every accumulation point of $\{u_k\}$ is optimal.*

7 Numerical Results

In this section, we illustrate the implementation of Algorithm 1 on two simple examples. The first one is a toy example where it is easy to obtain the optimal control analytically, while the second one consists in controlling the amplitude of a harmonic oscillator on a given time interval.

We coded Algorithm 1 with Scilab 5.5.2 (see www.scilab.org). The implemented stopping test was $\|\theta(u_k)\| < 10^{-6}$ and the quadratic programs (83) were solved with `quapro` Scilab routine. The Armijo parameters were $\alpha = 0.1$ and $\beta = 0.9$. Tests were run on a 3.40 GHz, 8GB RAM, Intel Core i7 processor PC. The preliminary results seem to be promissory, though the addressed problems are small academic examples. Despite the fact that the presented approach uses only time discretization, an acceptable accuracy may require large sample sizes and small time step-sizes, so the approximation scheme may lead to large scale problems (even more so with high dimensional systems). This drawback is already mentioned in the literature (see [15, 19, 22]) and further study is needed on this topic, but that is beyond the scope of the present work.

7.1 A Simple Example with Analytic Solution

Consider the problem (P) with the uncertain parameter uniformly distributed in $\Omega = [0, 1]$. The trajectory evolves in $\mathbb{R}$ satisfying $\dot{y}(t, \omega) = B(\omega)u(t)$, for $t \in [0, 1]$ and $y(0, \omega) = 0$, where $B(\omega) = 1$ if $\omega \in [0, a]$ and $B(\omega) = -1$ if $\omega \in (a, 1]$, for some fixed $a \in (0, 1)$. The control constraint set is $U = [0, 1]$ and the cost function is given by $f(y, \omega) = y$.

Hence, since $u \geq 0$, it is straightforward that

$$F(u, \omega) = \begin{cases} \max\limits_{t\in[0,1]} \int_0^t u(s)\mathrm{d}s = \int_0^1 u(s)\mathrm{d}s, & \text{if } \omega \in [0, a], \\ -\min\limits_{t\in[0,1]} \int_0^t u(s)\mathrm{d}s = 0, & \text{if } \omega \in (a, 1]. \end{cases}$$

Therefore, $J(u) = a\int_0^1 u(s)\mathrm{d}s$, which attains its unique minimum at $u \equiv 0$. Suppose that an independent uniform distributed sample $\{\omega_1, \omega_2, ..., \omega_M\}$ is taken from $[0, 1]$. If $\omega_i \in [0, a]$ for at least one $i \in \{1, ..., M\}$, the problem (P_M) has a unique minimizer $u \equiv 0$. Otherwise, if $\omega_i \in (a, 1]$ for all i, then $J \equiv 0$ and any control in $\mathcal{U}$ is a minimizer. So, if a is small, one can expect to need large sample sizes M in order to obtain meaningful approximations, in the sense that the functional J evaluated at an optimal control of (P_M) is near the optimum of (P). Precisely, the probability of $\omega_i \in [0, a]$ for at least one $i \in \{1, ..., M\}$ is $1-(1-a)^M$, so for obtaining a meaningful problem (P_M) with probability larger than $p \in (0, 1)$, it is necessary that $\frac{\ln(1-p)}{\ln(1-a)} < M$. This behavior is illustrated in Table 1, where results for decreasing values of a and increasing values of M are showed.

For all the trials, the step-size was $h = 0.05$ and the initial control was $u_1 = (1, ..., 1) \in U^{20}$ with $\|u_1\| \approx 4.47$. We report the number of sample elements ω_i belonging to $[0, a]$, the obtained objective function value and the norm of the optimal control, as well as the

Table 1 Numerical results for different values of M. If no ω_i belongs to $[0, a]$, the initial control is optimal, otherwise, the optimal control $u \equiv 0$ is obtained

M	$\omega_i \leq a$	J	$\|u\|$	it.	$\omega_i \leq a$	J	$\|u\|$	it.
	$a = .1$				$a = .05$			
10	0	0.00	4.47	1	0	0.00	4.47	1
15	3	0.00	0.00	2	0	0.00	4.47	1
20	1	0.00	0.00	5	3	0.00	0.00	4
30	4	0.00	0.00	3	5	0.00	0.00	3
35	3	0.00	0.00	4	1	0.00	0.00	8
40	8	0.00	0.00	2	2	0.00	0.00	5
45	6	0.00	0.00	3	4	0.00	0.00	4
50	3	0.00	0.00	5	4	0.00	0.00	4
	$a = .025$				$a = .0125$			
10	0	0.00	4.47	1	0	0.00	4.47	1
15	0	0.00	4.47	1	0	0.00	4.47	1
20	3	0.00	0.00	3	1	0.00	0.00	5
25	0	0.00	4.47	1	1	0.00	0.00	6
30	0	0.00	4.47	1	0	0.00	4.47	1
35	0	0.00	4.47	1	1	0.00	0.00	8
40	1	0.00	0.00	9	0	0.00	4.47	1
45	1	0.00	0.00	10	2	0.00	0.00	6
50	2	0.00	0.00	6	1	0.00	0.00	11

required number of iterations. Note that, when no sample elements belong to $[0, a]$, only one iteration is required because any control is optimal (in particular the initial control u_1). Otherwise, the algorithm obtains the optimal control $u \equiv 0$, and tends to be more efficient as more sample elements belong to $[0, a]$. For the larger problem with just one sample element in $[0, a]$, 11 iterations were necessary to achieve the tolerance, but it took less than 10 seconds.

7.2 Harmonic Oscillator

We consider a harmonic oscillator with natural frequency ω uniformly distributed on $[0, 1]$, starting from an initial point with non-zero velocity. The aim is to design a control that minimize the amplitude on a given time interval $[0, 2]$. This is a variation of the problem studied in [19, Section 7.1], where the system starts at an extremum point with zero velocity and the aim is to stabilize the oscillator at the final time.

In particular, we performed numerical experiments for the dynamical system given by $\dot{y} = A_\omega y + u$ with the initial condition $y(0) = (0, 1)^\top$, where $y(t) = (y_1(t), y_2(t))^\top : [0, 2] \to \mathbb{R}^2$, $u(t) = (u_1(t), u_2(t))^\top : [0, 2] \to [-3, 3]^2$, and

$$A_\omega = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

In this case, the functional cost is given by $J(u) = \mathbb{E}^{\mathbb{P}} \max_{t \in [0,2]} y_1^2(t)$.

We analyze the behavior of the system in three different cases: the uncontrolled system, the system with a control designed for a specific single parameter $\bar{\omega}$ and the system with the control obtained by the sample average approach. Specifically, considering a sample size M, we take a sample $\{\omega_1, ..., \omega_M\}$ in order to define the problem (P_M) and compute, for step-sizes h, the objective function values $J_M^h(u)$ for the controls $u \equiv 0$, $u = \bar{u}$ and $u = u^*$, where $\bar{u}$ solves (P_1) with the fixed parameter $\bar{\omega}$ and u^* solves (P_M^h).

First, we consider the single parameter approach, i.e. problem (P_1) with some fixed parameter $\hat{\omega} \in [0, 1]$. It is straightforward that this problem has the exact minimum $J(\hat{u}) = 0$ at the control $\hat{u} \equiv (\hat{\omega}, 0)$. Now, by introducing this control in problem (P), the obtained trajectories for any $\omega \neq 0$ are

$$\begin{cases} y_{\hat{\omega},1}(t,\omega) &= \left(\frac{\hat{\omega}}{\omega} - 1\right)\sin(\omega t), \\ y_{\hat{\omega},2}(t,\omega) &= \left(1 - \frac{\hat{\omega}}{\omega}\right)\cos(\omega t) + \frac{\hat{\omega}}{\omega}. \end{cases}$$

Table 2 Convergence table for P_M^h as h goes to zero. Comparison between the values of the cost function J_M^h at the three different controls 0, $\bar{u}$ and u^*

N	it.	QP it.	Time (s)	$J_M^h(u^*)$	$\|\theta(u^*)\|$	$J_M^h(0)$	$J_M^h(\bar{u})$
$M = 10$							
10	244	16	14.44	0.007788	8.401005e-07	0.800874	0.215130
20	125	17	23.01	0.005969	6.543774e-07	0.725481	0.195252
40	99	15	161.55	0.005190	9.733017e-07	0.690280	0.186078
80	42	14	176.20	0.005024	7.751443e-07	0.673507	0.181744
160	13	14	206.21	0.005068	1.039236e-07	0.665313	0.179636
$M = 20$							
10	468	15	53.00	0.006660	8.285572e-08	0.766726	0.215746
20	25	16	8.95	0.005152	2.079706e-07	0.696047	0.197660
40	30	15	66.89	0.004592	8.617495e-07	0.663169	0.189361
80	31	16	246.04	0.004378	9.072021e-07	0.647514	0.185447
160	22	15	729.86	0.004338	1.866180e-07	0.639851	0.183539
$M = 40$							
10	400	17	90.54	0.014411	9.826213e-07	0.749512	0.216604
20	53	17	39.07	0.009176	8.677528e-07	0.681199	0.199407
40	25	20	88.54	0.007388	4.323203e-07	0.649630	0.191585
80	9	22	168.57	0.006662	9.191242e-07	0.634515	0.187871
160	14	20	1171.86	0.006308	3.047304e-07	0.627107	0.186059
$M = 80$							
10	221	19	110.51	0.009605	7.934991e-07	0.740930	0.217189
20	57	18	97.71	0.006882	1.144276e-07	0.673885	0.200459
40	17	17	175.53	0.006152	6.315244e-07	0.642891	0.192849
80	22	19	984.30	0.005836	8.265107e-07	0.628013	0.189226
160	15	17	3284.38	0.005746	2.460912e-07	0.620732	0.187460

Hence,

$$\sup_{t\in[0,2]} \left(y_{\hat{\omega},1}(t,\omega)\right)^2 = \begin{cases} \left(\frac{\hat{\omega}}{\omega}-1\right)^2 \sin^2(2\omega), & \omega \in (0,\pi/4), \\ \left(\frac{\hat{\omega}}{\omega}-1\right)^2, & \omega \in [\pi/4, 1], \end{cases}$$

and therefore

$$J(\hat{\omega}) = \int_0^{\pi/4} \left(\frac{\hat{\omega}}{\omega}-1\right)^2 \sin^2(2\omega)\mathrm{d}\omega + \int_{\pi/4}^1 \left(\frac{\hat{\omega}}{\omega}-1\right)^2 \mathrm{d}\omega,$$

which attains its minimum at $\hat{\omega} = \bar{\omega}$, where

$$\bar{\omega} = \frac{\int_0^{\pi/4} \frac{\sin^2(2\omega)}{\omega}\mathrm{d}\omega - \ln(\pi/4)}{\int_0^{\pi/4} \frac{\sin^2(2\omega)}{\omega^2}\mathrm{d}\omega - 1 + 4/\pi} \approx .394139.$$

Thus, $\bar{u} \equiv (\bar{\omega}, 0)$ is the best control that can be obtained by the single parameter approach and this is the control that we use in the numerical trials.

In Table 2, we show the results for the sample sizes $M \in \{10, 20, 40, 80\}$. For each case, fixing the problem (P_M), we consider the approximations (P_M^h) with step-sizes $h = 2/N$, for $N \in \{10, 20, 40, 80, 160\}$. For each M, when $N = 10$ the algorithm was initialized at $u = \bar{u}$ and for the successive approximations the initial point was the linear interpolation of the previous solution. For each problem (P_M^h), we report the number of iterations, the average number of iterations required to solve the quadratic programs (83) at each iteration of the algorithm, the corresponding total computational time and the obtained optimal value $J_M^h(u^*)$, as well as the values of J_M^h at $u \equiv 0$ and $u = \bar{u}$.

As expected, the optimal value tends to decrease as the step-size h tends to zero. However, the optimal values are slightly smaller for $M = 10$ and $M = 20$ than for $M = 40$ and $M = 80$, with the same behavior for the single parameter control. For the different values of M, the number of iterations is large for $N = 10$ and quite small for the successive

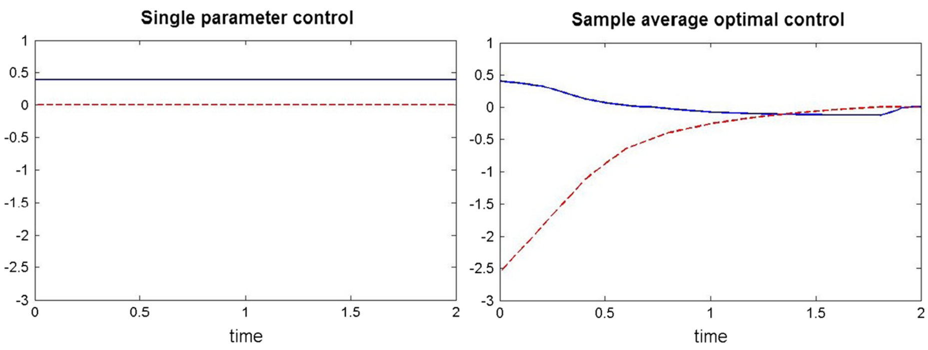


Fig. 1 Obtained optimal controls for the single parameter approach and the sample average approach

values of N, due to the choice of the initial points. This reduction is important in the aim of decrease the total execution time, since there is an appreciable growth of computational time per iteration as M and N increase.

Note that our approach significantly improves the obtained results comparing with the uncontrolled case and the single parameter approach. This behavior is also illustrated in the graphics which we present below, for the particular case $M = 80$ and $N = 160$. Figure 1 shows both components (the first one in solid line and the second one in dashed line) of the control $\bar{u}$ obtained by the single parameter approach and the optimal control u^* obtained by the sample average approach.

In Fig. 2, for the three cases described above, we show on the left the first component of the trajectories, i.e. $y_1(\cdot, \omega_i)$ for all $i = 1, \ldots, M$, and on the right the second one $y_2(\cdot, \omega_i)$. It can be observed that the sample average approach notably improves the results, decreasing

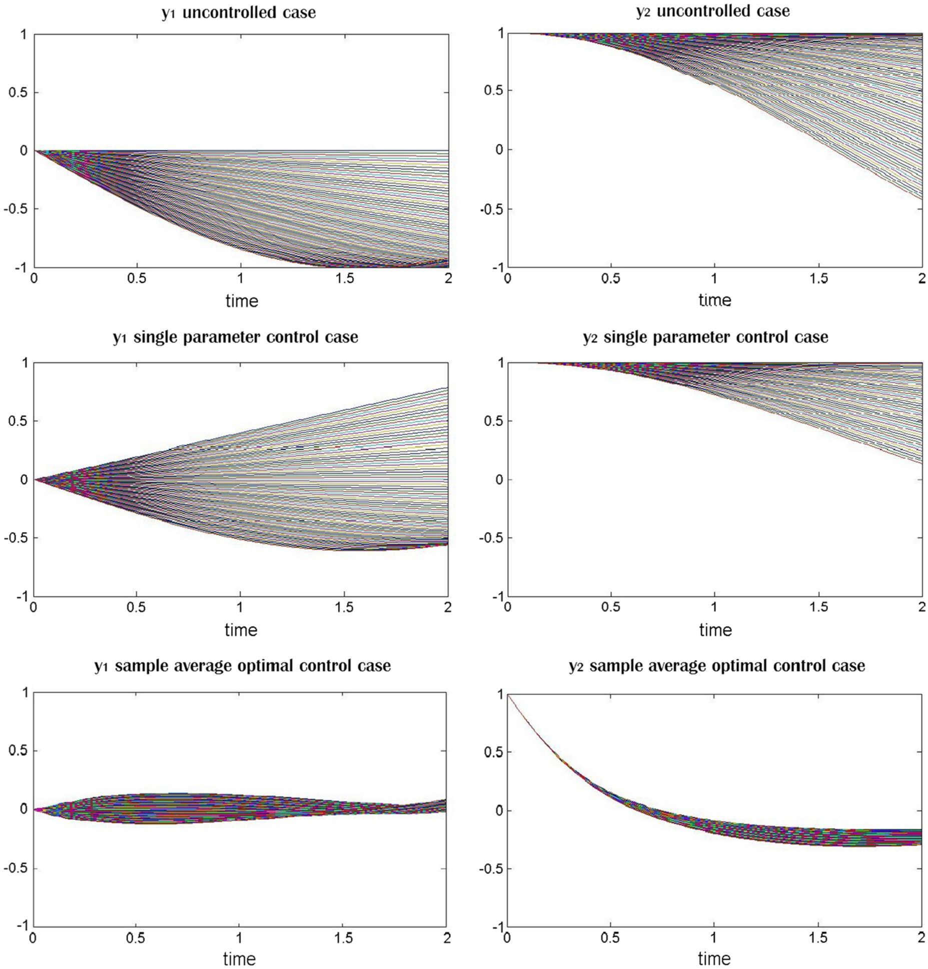


Fig. 2 Obtained dynamics for the uncontrolled case, the single parameter case and the sample average case

not only the range of amplitudes in the considered time interval, but also the the range of final velocities.

Acknowledgements We thank the anonymous referees for their useful comments and suggestions.

References

1. Aragone, L.S., Di Marco, S., González, R.L.V.: Numerical analysis of a minimax optimal control problem with an additive final cost. Math. Models Methods Appl. Sci. **12**(2), 183–203 (2002). https://doi.org/10.1142/S021820250200160X
2. Artstein, Z., Wets, R.J.: Consistency of minimizers and the slln for stochastic programs. Journal of Convex Analysis **2**(1–2), 1–17 (1995). http://eudml.org/doc/226566
3. Attouch, H.: Variational Convergence for Functions and Operators. Applicable Math- Ematics Series. Pitman, London (1984)
4. Attouch, H., Wets, R.J.B.: Approximation and convergence in nonlinear optimization. Nonlinear Programming **4**, 367–394 (1981)
5. Attouch, H., Wets, R.J.B.: A convergence theory for saddle functions. Trans. Am. Math. Soc. **280**(1), 1–41 (1983)
6. Barron, E.N.: The Bellman equation for control of the running max of a diffusion and applications to look-back options. Appl. Anal. **48**, 205–222 (1993). https://doi.org/10.1080/00036819308840158
7. Barron, E.N., Ishi, H.: The Bellman equation for minimizing the maximum cost. Nonlinear Anal. Theory Methods Appl. **13**(9), 1067–1090 (1989). https://doi.org/10.1016/0362-546X(89)90096-5
8. Bokanowski, O., Picarelli, A., Zidani, H.: Dynamic programming and error estimates for stochastic control problems with maximum cost. Appl. Math. Optim. **71**(1), 125–163 (2015). https://doi.org/10.1007/s00245-014-9255-3
9. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
10. Bulirsch, R., Montrone, F., Pesch, H.J.: Abort landing in the presence of windshear as a minimax optimal control problem, part 1: Necessary conditions. J. Optim. Theory Appl. **70**(1), 1–23 (1991). https://doi.org/10.1007/BF00940502
11. Bulirsch, R., Montrone, F., Pesch, H.J.: Abort landing in the presence of windshear as a minimax optimal control problem, part 2: Multiple shooting and homotopy. J. Optim. Theory Appl. **70**(2), 223–254 (1991). https://doi.org/10.1007/BF00940625
12. Di Marco, S., González, R.L.V.: Minimax optimal control problems. Numerical analysis of the finite horizon case ESAIM: Mathematical Modelling and Numerical Analysis **1**(33), 23–54 (1999). https://doi.org/10.1051/m2an:1999103
13. Gianatti, J., Aragone, L.S., Lotito, P.A., Parente, L.A.: Solving minimax control problems via nonsmooth optimization. Oper. Res. Lett. **44**(5), 680–686 (2016). https://doi.org/10.1016/j.orl.2016.08.001. http://www.sciencedirect.com/science/article/pii/S016763771630075X
14. King, A.J., Wets, R.J.: Epi-consistency of convex stochastic programs. Stochastics and Stochastic Reports **34**(1-2), 83–92 (1991). https://doi.org/10.1080/17442509108833676
15. Kleywegt, A.J., Shapiro, A., Homem-de Mello, T.: The sample average approximation method for stochastic discrete optimization. SIAM J. Optim. **12**(2), 479–502 (2002). https://doi.org/10.1137/S1052623499363220
16. Mosco, U.: Convergence of convex sets and of solutions of variational inequalities. Adv. Math. **3**(4), 510–585 (1969)
17. Oberle, H.J.: Numerical solution of minimax optimal control problems by multiple shooting technique. J. Optim. Theory Appl. **50**(2), 331–357 (1986). https://doi.org/10.1007/BF00939277
18. Petersen, I.R., James, M.R., Dupuis, P.: Minimax optimal control of stochastic uncertain systems with relative entropy constraints. IEEE Trans. Autom. Control **45**(3), 398–412 (2000). https://doi.org/10.1109/9.847720
19. Phelps, C., Royset, J.O., Gong, Q.: Optimal control of uncertain systems using sample average approximations. SIAM J. Control. Optim. **54**(1), 1–29 (2016). https://doi.org/10.1137/140983161
20. Polak, E.: Computational Methods in Optimization. a Unified Approach. Academic Press, New York (1986)
21. Rockafellar, R., Wets, R.: Variational Analysis. Grundlehren der mathematischen Wissenschaften. Springer, Berlin (2009)

22. Royset, J.O., Polak, E.: Implementable algorithm for stochastic optimization using sample average approximations. J. Optim. Theory Appl. **122**(1), 157–184 (2004). https://doi.org/10.1023/B:JOTA.0000041734.06199.71
23. Ugrinovskii, V., Petersen, I.: Absolute stabilization and minimax optimal control of uncertain systems with stochastic uncertainty. SIAM J. Control. Optim. **37**(4), 1089–1122 (1999). https://doi.org/10.1137/S0363012996309964
24. Vinter, R.B.: Minimax optimal control. SIAM J. Control. Optim. **44**(3), 939–968 (2005). https://doi.org/10.1137/S0363012902415244
25. Wijsman, R.A.: Convergence of sequences of convex sets, cones and functions. Bull. Am. Math. Soc. **70**(1), 186–188 (1964)
26. Wijsman, R.A.: Convergence of sequences of convex sets, cones and functions. Trans. Am. Math. Soc. **II**(123), 32–45 (1966)
27. Wu, C., Teo, K.L., Wang, X.: Minimax optimal control of linear system with input-dependent uncertainty. J. Frankl. Inst. **351**(5), 2742–2754 (2014). https://doi.org/10.1016/j.jfranklin.2014.01.012. http://www.sciencedirect.com/science/article/pii/S0016003214000143