SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

Two approaches to constrained stochastic optimal control problems

L. Pfeiffer

SFB-Report No. 2015-007 May 2015

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

Der Wissenschaftsfonds.

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

Two approaches to constrained stochastic optimal control problems

Laurent Pfeiffer*

May 12, 2015

Abstract

In this article, we study and compare two approaches to solving stochastic optimal control problems with an expectation constraint on the final state. The case of a probability constraint is included in this framework. The first approach is based on a dynamic programming principle and the second one uses Lagrange relaxation. These approaches can be used for continuous-time problems; we provide numerical results for an academic example.

Keywords Stochastic optimal control, expectation and probability constraints, dynamic programming, Lagrange relaxation.

AMS subject classification 90C15, 93E20.

1 Introduction

Setting In this article, we study stochastic optimal control problems with an expectation constraint on the final state. This class of problems contains some chance-constrained stochastic optimal control problems. For the sake of clarity, we start by describing the problem under study.

Let $T \in \mathbb{N}^*$, let $(\xi_t)_{t=1,...,T}$ be T mutually independent random values. The assumption of independence is a key assumption, it will enable us to derive dynamic programming principles. To simplify, we assume that the random values ξ_t are identically distributed, with a discrete law. Let $I \in \mathbb{N}^*$ and $(p_i)_{i=1,...,I}$ in $(0,1]^I$ be such that $\sum_{i=1}^I p_i = 1$, we assume that for all $i \in \{1,...,I\}$,

$$\mathbb{P}\big[\xi_t = i\big] = p_i. \tag{1.1}$$

For all $t \in \{0,...,T-1\}$, for all $s \geq t$, we denote by $\mathcal{F}_{t,s}$ the σ-algebra generated by $\{\xi_{t+1},...,\xi_s\}$ and we denote by $\mathbb{F}_t$ the filtration $(\mathcal{F}_{t,s})_{s=t,...,T}$.

Let m, $n \in \mathbb{N}^*$, let S and U be closed subsets of resp. $\mathbb{R}^n$ and $\mathbb{R}^m$. The sets S and U are resp. called the *state* and *control spaces*. For all t, we define the set $\mathcal{U}_t$ of $\mathbb{F}_t$-adapted control processes $(u_s)_{s=t,...,T-1}$ with values in U.

*Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität, Heinrichstraße 36, 8010 Graz, Austria. Mail: laurent.pfeiffer@uni-graz.at
The author thanks J. Frédéric Bonnans for his advice. The research leading to these results has received funding from the Gaspard Monge Program for Optimization and operations research (PGMO).

Let $F : S \times U \times \{1, ..., I\} \to S$. For all $t \in \{0, ..., T\}$, for all $x \in S$, for all $u \in \mathcal{U}_s$, we define the state process $(X_s^{t,x,u})_{s=t,...,T}$ as the unique solution to

$$\begin{cases} X_{s+1} = & F(X_s, u_s, \xi_{s+1}), \quad \forall s = t, ..., T-1, \\ X_t = & x. \end{cases} \tag{1.2}$$

Therefore, when starting at time 0 from the initial state $x_0 \in S$, the decision process is as follows:

$$x_0 \to \text{ decision of } u_0 \ \to \text{ observation of } \xi_1 \ \to x_1 \to \ \cdots \ \to$$
$$x_t \to \text{ decision of } u_t \ \to \text{ observation of } \xi_{t+1} \ \to x_{t+1} \to \ \cdots \to x_T.$$

Note that the state variable $X_s^{t,x,u}$ is $\mathcal{F}_{t,s}$-measurable.

Let $\phi : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$, let $z_0 \in \mathbb{R}$, we consider the following stochastic optimal control problem with an expectation constraint:

$$\min_{u \in \mathcal{U}_0} \mathbb{E}\big[\phi(X_T^{0,x,u})\big], \tag{1.3}$$

$$\text{s.t. } \mathbb{E}\big[g(X_T^{0,x,u})\big] \geq z. \tag{1.4}$$

The case of a probabilistic constraint of the form: $\mathbb{P}\big[h(X_T^{0,x_0,u}) \in K\big] \geq z$ enters into this framework, by setting

$$g(x) = \mathbf{1}_K \circ h(x) := \begin{cases} 1 & \text{if } h(x) \in K, \\ 0 & \text{otherwise.} \end{cases} \tag{1.5}$$

We also define the value function V: for all $t \in \{0, ..., T\}$, $x \in S$, and $z \in \mathbb{R}$,

$$V_t(x, z) \ := \ \min_{u \in \mathcal{U}_t} \mathbb{E}\big[\phi(X_T^{t,x,u})\big], \tag{1.6}$$

$$\text{s.t. } \mathbb{E}\big[g(X_T^{t,x,u})\big] \geq z. \tag{1.7}$$

Context We refer to [16, Section 4] for a general reference on optimization problems with probability constraints. In this article, we focus on two possible approaches to solving this kind of problems. The first one consists in adding a supplementary state variable, denoted by Z. Its initial state is equal to the required level for the constraint, z, it must be a submartingale and must satisfy the following constraint:

$$Z_T \leq g(X_T), \quad \text{almost surely.} \tag{1.8}$$

This enables to replace the expectation constraint (1.4) by the stochastic-target constraint (1.8). Note that when (1.4) is active, the variable Z in the constraint (1.8) can only be the conditional expectation of $g(X_T)$. As a consequence, a dynamic programming principle can be formulated for the value function $V_t(x, z)$ of the problem. This approach was used, for example in [5, Section 3], and in the thesis [6] (in French). It was successfully applied in [7, Chapter 4] to a power management problem for hybrid vehicles. In [4], it is proved that the value function of continuous-time stochastic target problems (therefore including chance-constrained problems) is the viscosity solution to a certain HJB equation (see also [17, Chapter 8]). A reformulation of this HJB equation is also proposed in the PhD thesis of the author [14, Chapter 6] to avoid unbounded controls, using an idea of [3].

The second method is the Lagrange relaxation method [10, Chapter XII], [12], which basically consists in "dualizing" the expectation constraint. Instead of solving the constrained problem, we solve the following problem (which is an unconstrained stochastic optimal control problem), for various values of $\lambda \geq 0$:

$$\mathbb{E}\big[\phi(X_T^{0,x,u}) - \lambda g(X_T^{0,x,u})\big]. \tag{1.9}$$

The starting point of this method is the following: for any $\lambda \geq 0$, the solution is also a solution to the constrained problem for a value of z depending on λ. This technique is used in [1], for example.

If the dynamic of the system is linear and if g is concave, the expectation constraint is a convex constraint, in so far as for two given controls satisfying the expectation constraint, the average control also satisfies the constraint. In general, this property is not satisfied, therefore, the Lagrange relaxation cannot provide an optimal control for all the values of z. Note that the set of controls satisfying the constraint may even be not connected. These issues are studied in [8, 9].

However, a probabilistic relaxation of the problem can be formulated. In a few words, we allow the taken decision to depend on a supplementary continuous random variable. This supplementary random variable and the original ones are mutually independent. The point of view is now the following: instead of optimizing the control strategy, we are allowed to use all the possible control strategies, with a probability measure that has to be optimized. The cost and the expectation constraint become linear with respect to this probability measure. As a consequence, the associated relaxed value function is the convex envelope of the initial value function and is therefore the value function obtained with the Lagrange relaxation method. This idea is usual in the literature, see for example [11, Section 3] and is to some extent similar to the notion of mixed strategies in game theory or to the relaxation of an optimal transportation problem.

The paper is organized as follows. In section 2, we describe the dynamic programming principle satisfied by the value function. We also prove that the boundary of the value function satisfies a dynamic programming principle. In section 3, we describe precisely the probabilistic relaxation, which seems to be new in the framework of chance-constrained problems, and state some new properties of the relaxed value function (theorem 11). In section 4, we compare the two approaches and give new error estimates for the two of them (theorems 16, 17, 19). Note that the two approaches provide an upper and a lower bound of the value function. Section 5 is dedicated to continuous-time problems and numerical schemes for these problems. We give an original proof of the convexity in z of the value function when g is Lipschitz (theorem 22). In section 6, we show some numerical results, for a discretized continuous-time problem. On this simple example, the duality gap is very small and the Lagrange relaxation method is much more efficient.

2 Dynamic programming principles

2.1 Existence of solutions

We make the following assumption on the data.

Assumption 1. *1. For all $i \in \{1, ..., I\}$, $F(\cdot, \cdot, i)$ is continuous on $S \times U$ and ϕ is continuous.*

2. *One of the following statements holds:*

   (a) *g is continuous on S,*

   (b) *for some $p \in \mathbb{N}^*$, g has the following form: $g = \mathbf{1}_K \circ h$, where $h : S \to \mathbb{R}^p$ is continuous and K is a closed subset of $\mathbb{R}^p$.*

The existence of optimal control processes for problem (1.6)-(1.7) is rather easy to establish in this framework. Let us fix t. The elementary events of the σ-algebras of the filtration $\mathbb{F}_t$ can be represented with a tree. At the initial time t, there is the root of the tree. The root has I children, corresponding to the I possible realizations of ξ_{t+1}, each of them having I children corresponding to the I possible realizations of ξ_{t+2}, and so on, up to time T. The tree has $1 + I + I^2 + ... + I^{T-t} = \frac{I^{T-t+1}-1}{I-1}$ nodes and each node, except the leaf nodes, corresponds to a decision. Therefore, there is a clear bijection between $\mathcal{U}_t$ and $U^{\frac{I^{T-t}-1}{I-1}}$ and the stochastic problem (1.6)-(1.7) can be reformulated as a deterministic finite-dimensional problem, as it is well-known. This bijection also defines a topology on $\mathcal{U}_t$.

Lemma 2. *The set of control processes $\mathcal{U}_t$ is compact, and under assumption 1, for any continuous function $\psi : S \to \mathbb{R}$, the mapping: $u \in \mathcal{U}_t \mapsto \mathbb{E}\big[\psi(X_T^{t,x,u})\big]$ is well-defined, continuous, and bounded. Moreover, the mapping $u \mapsto \mathbb{E}\big[g(X_T^{t,x,u})\big]$ is well-defined and bounded, and for all $z \in \mathbb{R}$, the set*

$$\mathcal{U}_t^z := \big\{u \in \mathcal{U}_t \,:\, \mathbb{E}\big[g(X_T^{t,x,u})\big] \geq z\big\} \tag{2.1}$$

is closed in $\mathcal{U}_t$ (and therefore compact).

Proof. The compactness of $\mathcal{U}_t$ is a consequence of the compactness of U, and the continuity of $\mathbb{E}\big[\psi(X_T^{t,x,\cdot})\big]$ a consequence of the continuity of $F(\cdot,\cdot,i)$. The second statement of the lemma follows if g is continuous. Let us consider the case when $g = \mathbf{1}_K \circ h$ with K closed and h continuous. Let $z \in \mathbb{R}$, if $\mathcal{U}_t^z$ is non-empty, let $(u^k)_k$ be a converging sequence in $\mathcal{U}_t^z$. The algebra $\mathcal{F}_{t,T}$ being finite, we can extract a subsequence for which the event $h(X_T^{t,x,u^k}) \in K$ is constant and pass to the limit. □

As a corollary, we obtain the following result about the existence of optimal control processes.

Lemma 3. *Under assumption 1, the problem* (1.6)-(1.7) *has a solution if it is feasible. Moreover, for all t and x, the mapping $z \mapsto V_t(x,z)$ is lower-bounded, left-continuous, and since it is non-decreasing, it is lower semi-continuous.*

2.2 Dynamic programming for the value function

We explain now how to reformulate the problem with an expectation constraint as a problem with a stochastic-target constraint, thanks to the following lemma.

Lemma 4. *Let $z \in \mathbb{R}$, $t \in \{0,...,T\}$, $x \in S$, and $u \in \mathcal{U}_t$. Then, the constraint* (1.7) *holds if and only if there exists an $\mathbb{F}_t$-adapted submartingale $(Z_s)_{s=t,...,T}$ satisfying*

$$Z_t = z \quad \textit{and} \quad Z_T \leq g(X_T^{t,x,u}). \tag{2.2}$$

Moreover, if say $\underline{g} \in \mathbb{R}$ is a lower bound of g and $z \geq \underline{g}$ (resp. $\bar{g}$ is an upper bound and $z \leq \bar{g}$), then, we can impose that $Z \geq \underline{g}$ almost surely (resp. $Z \leq \bar{g}$).

Proof. Let Z be an $\mathbb{F}_t$-adapted submartingale Z satisfying (2.2). Then,

$$\mathbb{E}\big[g(X_T^{t,x,u})\big] \geq \mathbb{E}\big[Z_T\big] \geq Z_t = z,$$

and (1.7) holds. Conversely, if (1.7) holds, set $z_0 = \mathbb{E}\big[g(X_T^{s,x,u})\big]$ and define

$$Z_s = \mathbb{E}\big[g(X_T^{s,x,u}) \,|\, \mathcal{F}_{t,s}\big] - (z_0 - z).$$

Clearly, Z is an adapted martingale and (2.2) holds. If $\underline{g}$ is a lower bound of g, we define the stopping time

$$\tau := \inf\big\{s \,|\, Z_s \leq \underline{g}\big\}$$

and we set $Z'_s = Z_{\min(s,\tau)}$. The process Z' is still a martingale (a *stopped martingale*), and is greater than or equal to $\underline{g}$. Moreover, it is easy to check that $Z'_s \leq \max(\underline{g}, Z_s)$, therefore, (2.2) still holds. The proof is the same if $\bar{g}$ is an upper bound. □

In the sequel, any submartingale Z satisfying (2.2) is called *associated submartingale*. It is easy to check that, as we already mentioned, when the expectation constraint is active, Z is unique and is the conditional expectation of $g(X_T^{t,x,u})$ (thus, is a martingale). In the case of an active probability constraint, Z is the conditional probability that the constraint is satisfied and at the final time, is equal to 1 if the constraint is satisfied, 0 otherwise.

This lemma implies that instead of considering the expectation constraint (1.7), we can view z as a supplementary state variable, whose dynamic must be the one of a submartingale, with a stochastic-target constraint. Using the independence of the random variables, we obtain a dynamic programming principle for the value function.

Proposition 5. *The value function $V_t(x,z)$ satisfies, for all $x \in S$ and all $z \in \mathbb{R}$:*

$$V_T(x,z) = \begin{cases} \phi(x) & \text{if } g(x) \geq z, \\ +\infty & \text{otherwise,} \end{cases} \tag{2.3}$$

$$V_t(x,z) = \inf_{\substack{u \in U,\, (z_i)_i \in \mathbb{R}^I \\ \sum_{i=1}^I p_i z_i \geq z}} \Big\{ \sum_{i=1}^{I} p_i V_{t+1}\big(F(x,u,i), z_i\big) \Big\}, \quad \forall t = T-1, ..., 0. \tag{2.4}$$

Given t and x, for all $u \in \mathcal{U}_t$ and all adapted martingale Z, u is an optimal solution and Z satisfies (2.2) *if and only if for all $t \leq s < T$, (u_s, Z_{s+1}) is almost surely a solution to* (2.4).

2.3 Boundaries of the domain of the value function

In this subsection, we describe the boundaries of the domain of the value function. Given $t \in \{0, ..., T\}$ and $x \in S$, we define the upper boundary of the domain $Z_t^+(x)$ by

$$Z_t^+(x) := \sup_{u \in \mathcal{U}_t} \mathbb{E}\big[g(X_T^{t,x,u})\big]. \tag{2.5}$$

This is the highest reachable value for the constraint and can also be characterized by $Z_t^+(x) = \sup\big\{z \,|\, V_t(x,z) < +\infty\big\}$. In this subsection, we analyse first this boundary and the value

$$V_t^+(x) := V_t(x, Z_t^+(x)) \tag{2.6}$$

of V on the boundary. This function is the maximal cost which is worth paying, considering the expectation constraint.

We also analyse the minimal cost which has to be paid, simply defined by

$$V_t^-(x) := \inf_{u \in \mathcal{U}_t} \mathbb{E}\big[\phi(X_T^{t,x,u})\big]. \tag{2.7}$$

This cost can also be defined as follows: $V_t^-(x) = \lim_{z \to -\infty} V_t(x,z)$. The lower boundary of the domain which is associated is defined by

$$Z_t^-(x) := \sup\big\{ z \,|\, V_t(x,z) = V_t^-(x) \big\}. \tag{2.8}$$

This is a boundary only in the sense that it is the maximum level for the expectation constraint that can be satisfied without paying more than the minimal possible amount.

Note that if assumption 1 holds, then by lemma 2, the values $Z_t^+(x)$ and $V_t^-(x)$ are always finite and there exists an optimal solution to the associated problems. Moreover, the values $V_t^+(x)$ and $Z_t^-(x)$ are then also finite.

Upper boundary The upper boundary $Z_t^+(x)$ is the value function of a stochastic optimal control problem and it satisfies the following dynamic programming principle: for all $x \in S$ and all $z \in \mathbb{R}$,

$$Z_T^+(x) = g(x), \tag{2.9}$$

$$Z_t^+(x) = \sup_{u \in U} \Big\{ \sum_{i=1}^{I} p_i Z_{t+1}^+(F(x,u,i)) \Big\}, \quad \forall t = T-1, ..., 0. \tag{2.10}$$

As we already point out, by lemma 2 and under assumption 1, for all t and x, there exists a control u such that $\mathbb{E}\big[g(X_T^{t,x,u})\big] = Z_t^+(x)$. In the following proposition, we show that there is only one submartingale associated with such a control, which is equal almost surely to the upper boundary, along time.

Proposition 6. *Let $t \in \{0, ..., T\}$, let $x \in S$. Let us set $z = Z_t^+(x)$, let $u \in \mathcal{U}_t$ be such that $\mathbb{E}\big[g(X_T^{t,x,u})\big] = z$ and let Z be an associated submartingale. Then, for all $s \geq t$,*

$$Z_s = Z_s^+(X_s^{t,x,u}), \quad \textit{almost surely.} \tag{2.11}$$

Proof. First, observe that:

$$z = Z_t \leq \mathbb{E}\big[Z_T\big] \leq \mathbb{E}\big[g(X_T^{t,x,u})\big] = z,$$

implying that Z is a martingale which is such that $Z_T = g(X_T^{t,x,u})$ almost surely. By (2.9), $g(X_T^{t,x,u}) = Z_T(X_T^{t,x,u})$, and by the dynamic programming principle (2.10), $Z_s^+(X_s^{t,x,u})$ is a martingale, therefore, for all $s \in \{t, ..., T\}$,

$$Z_s^+(X_s^{t,x,u}) = \mathbb{E}\big[Z_T^+(X_T^{t,x,u}) \,|\, \mathcal{F}_{s,T}\big] = \mathbb{E}\big[Z_T \,|\, \mathcal{F}_{s,T}\big] = Z_s, \quad \text{almost surely.}$$

The proposition is proved. ☐

As a consequence, there is a dynamic programming principle for the value of V on the upper boundary $V_t^+(x)$.

Lemma 7. *Under assumption 1, the mapping $V_t^+(x)$ satisfies the following relations: for all x, for $t = T-1, ..., 0$,*

$$V_T^+(x) = \phi(x), \tag{2.12}$$

$$V_t^+(x) = \inf_{u\in U} \Big\{ \sum_{i=1}^{I} p_i V_{t+1}^+(F(x,u,i)) \,\Big|\, Z_t^+(x) = \sum_{i=1}^{I} p_i Z_{t+1}^+(F(x,u,i)) \Big\}. \tag{2.13}$$

Note that the lemma is false if there does not exist an optimal control for (2.5), this is why we need assumption 1. Relation (2.13) cannot be considered as a "standard" dynamic programming principle, since the mapping Z^+ is itself involved. More precisely the constraint $Z_t^+(x) = \sum_{i=1}^{I} p_i Z_{t+1}^+(F(x,u,i))$ in (2.13) is equivalent to

$$u \in \arg\max_{v\in U} \Big\{ \sum_{i=1}^{I} p_i Z_{t+1}^+(F(x,v,i)) \Big\}. \tag{2.14}$$

Remark 8. *An even small perturbation of the mapping $x \mapsto Z_{t+1}^+(x)$, may generate a large perturbation of the set of maximizers in (2.14), and therefore a large perturbation in $V_t^+(x)$. Therefore, it may be difficult to compute $V_t^+(x)$.*

Lower boundary The minimal cost $V_t^-(x)$ is the value function associated with a stochastic optimal control problem and the associated dynamic programming principle is the following: for all $x \in S$,

$$V_T^-(x) = \phi(x), \tag{2.15}$$

$$V_t^-(x) = \inf_{u\in U} \Big\{ \sum_{i=1}^{I} p_i V_{t+1}^-(F(x,u,i)) \Big\}, \quad \forall t = T-1, ..., 0. \tag{2.16}$$

Under assumption 1, the lower boundary also satisfies a dynamic programming principle: for all $x \in S$, for all $t = T-1$,...,0,

$$Z_T^-(x) = g(x), \tag{2.17}$$

$$Z_t^-(x) = \sup_{u\in U} \Big\{ \sum_{i=1}^{I} p_i Z_{t+1}^-(F(x,u,i)) \,\Big|\, V_t^-(x) = \sum_{i=1}^{I} p_i V_{t+1}^-(F(x,u,i)) \Big\}. \tag{2.18}$$

Similarly to V^+ (see remark 8), a small perturbation of V_{t+1}^- may generate a large perturbation of Z_t^-. Note also that if $u \in \mathcal{U}_t$ is such that $\mathbb{E}\big[g(X_T^{t,x,u})\big] = Z_t^-(x)$ and if Z is an associated martingale (for the level $Z_t^-(x)$), then for all $s \geq t$, $Z_s = Z_s^-(X_s^{t,x,u})$, almost surely (similarly to proposition 6).

3 Lagrange relaxation

We introduce in this section the second approach, based on Lagrange relaxation. We use the ideas recalled in the appendix. We use the notations of section A.1. The Legendre-Fenchel transform, the biconjugate, the subdifferential, the convex envelope will be always considered with respect to z only. For example,

$$V_t^*(x,\lambda) = \sup_{z\in\mathbb{R}} \big\{ \lambda z - V_t(x,z) \big\}. \tag{3.1}$$

3.1 Formulation of the relaxed problem

We describe a probabilistic relaxation of problem (1.6)-(1.7), following the idea of subsection A.3. In the sequel, the superscript "r" will be used for all the relaxed tools. We introduce a supplementary random variable ζ, of uniform law on $[0,1]$ and such that ζ, ξ_1,..., ξ_T are mutually independent. For all $s \geq 0$, we denote by $\mathcal{F}^{\mathrm{r}}_{0,s}$ the σ-algebra generated by ζ, ξ_1,..,ξ_s and we denote by $\mathbb{F}^{\mathrm{r}}_0$ the filtration $(\mathcal{F}^{\mathrm{r}}_{0,s})_{s=0,...,T}$. Finally, we denote by $\mathcal{U}^{\mathrm{r}}_0$ the set of $\mathbb{F}_0$-adapted processes $(u_s)_{s=0,...,T}$ with values in U. For any $u \in \mathcal{U}^{\mathrm{r}}_0$, we denote by $(X^{0,x,u}_s)_{s=0,...,T}$ the solution to the dynamical system (1.2). For all $s \geq 0$, $X^{0,x,u}_s$ is $\mathcal{F}^{\mathrm{r}}_{0,s}$-adapted.

To sum up, the relaxed decision process is now the following:

$$x_0 \to \text{ observation of } \zeta \ \to \ \text{decision of } u_0 \ \to \ \text{observation of } \xi_1 \ \to$$
$$x_1 \to \text{ decision of } u_1 \ \to \ \text{observation of } \xi_2 \ \to x_2 \to \cdots \to x_T.$$

The only difference with the previous process is that we allow the decisions to depend on a supplementary random variable, which is independent of all the others, and has no influence in the dynamics F. Finally, we define the relaxed problem and the associated value function as follows:

$$V^{\mathrm{r}}_0(x,z) \ := \ \min_{u \in \mathcal{U}^{\mathrm{r}}_0} \mathbb{E}\big[\phi(X^{0,x,u}_T)\big], \tag{3.2}$$

$$\text{s.t. } \mathbb{E}\big[g(X^{0,x,u}_T)\big] \geq z. \tag{3.3}$$

For technical reasons, we introduce a new notation. For all $s \geq 0$, for all $\mathcal{F}^{\mathrm{r}}_{0,s}$-measurable random variable Y, for almost all $\alpha \in [0,1]$, there exists a unique random variable Y^α which is $\mathcal{F}_{0,s}$-measurable and such that

$$Y^\zeta = Y, \quad \text{almost surely.} \tag{3.4}$$

We use the same notation for $\mathbb{F}_0$-adapted processes. For example, if u is a relaxed control process, then u^α is the (unrelaxed) control process which is used if ζ takes the value α. This explains in particular why problem (3.2)-(3.3) is the relaxed version of problem (1.6)-(1.7) in the sense of the relaxation process detailed in subsection A.3.

As before, the problem can be reformulated with a stochastic-target constraint, thanks to the following lemma. The proof is the same as the one of lemma 4.

Lemma 9. *Let $z \in \mathbb{R}$, $x \in S$ and $u \in \mathcal{U}^{\mathrm{r}}_0$. Then, the constraint* (3.3) *holds if and only if there exists an $\mathbb{F}^{\mathrm{r}}_0$-adapted submartingale $(Z_s)_{s=0,...,T}$ satisfying:*

$$\mathbb{E}\big[Z_0\big] = z \quad \textit{and} \quad Z_T \leq g(X^{0,x,u}_T). \tag{3.5}$$

Such a submartingale will be still called associated submartingale*, moreover, for almost all $\alpha \in [0,1]$, u^α is an unrelaxed control process, feasible for the level Z^α_0, and $(Z^\alpha_s)_{s=0,...,T}$ is an associated submartingale.*

3.2 Dual approach

The following lemma summarizes some of the main results of subsection A.2 in the framework of stochastic problems.

Lemma 10. *Let $\lambda \geq 0$, $t \in \{0,...,T\}$, $x \in S$. Then,*

$$-V_t^*(x,\lambda) = \inf_{u\in\mathcal{U}_t} \mathbb{E}\big[\phi(X_T^{t,x,u}) - \lambda g(X_T^{t,x,u})\big]. \tag{3.6}$$

Under assumption 1, there exists a solution to this penalized problem and moreover, for all z, the relaxed problem (3.2)-(3.3) has a solution if it is feasible, and $V_0^{**}(x,z) = V_0^{\mathrm{r}}(x,z)$.

Proof. (a) The expression of the Legendre-Fenchel transform is a direct consequence of (A.8). The existence of an optimal solution to problem (3.6) under assumption 1 is a consequence of lemma 2.

(b) Let us prove the existence of a relaxed optimal solution. By lemma 3, the mapping $z \mapsto V_0(x,z)$ is lower-bounded, therefore, by lemma 27, for all z, $V_0^{\mathrm{r}}(x,z) = \mathrm{conv}(V_0)(x,z)$. Let z be such that $V_0^{\mathrm{r}}(x,z) < +\infty$, let $(z_1^k)_k$ and $(z_2^k)_k$ be two sequences in the definition domain $(-\infty, Z_0^+(x)]$ of $V_0^{\mathrm{r}}(x,\cdot)$, let $(\alpha_k)_k$ be a sequence in $[0,1]$ such that

$$\begin{aligned} &\alpha_k z_1^k + (1-\alpha_k)z_2^k = z, \ \forall k, \\ &\mathrm{conv}(V_0)(x,z) = \lim_{k\to\infty} \alpha_k V_0(x,z_1^k) + (1-\alpha_k)V_0(x,z_2^k). \end{aligned}$$

We assume, without loss of generality, that $z_1^k \leq z \leq z_2^k$. This means that $(z_2^k)_k$ is bounded, thus, extracting if necessary a subsequence, it converges to say $\bar{z}_2$. Let us consider two cases. If $\liminf z_1^k = -\infty$, then, extracting if necessary, $\alpha_k \to 0$. We also obtain, since $V_0(x,\cdot)$ is l.s.c., lower bounded, and non-decreasing that $\mathrm{conv}(V_0)(x,z) \geq V_0(x,z_2) \geq V_0(x,z)$ and finally that $\mathrm{conv}(V_0)(x,z) = V_0(x,z)$. In this case, any classical solution is also a relaxed optimal solution. In the other case, when the sequence z_1^k is lower bounded, the sequence is also upper bounded, thus we can extract a subsequence so that $(z_1^k)_k$ converges to say $\bar{z}_1$ and $(\alpha_k)_k$ to say $\bar{\alpha}$. Since $V_0(x,\cdot)$ is l.s.c.,

$$\mathrm{conv}(V_0)(x,z) \geq \bar{\alpha}V_0(x,\bar{z}_1) + (1-\bar{\alpha})V_0(x,\bar{z}_2),$$

and therefore, the previous inequality is an equality. Let u_1 and u_2 be the corresponding optimal solutions (in $\mathcal{U}_0$) to $\bar{z}_1$ and $\bar{z}_2$, we define then $u \in \mathcal{U}_0^{\mathrm{r}}$ such that: $u^\alpha = u_1$, if $\alpha \in (0,\bar{\alpha})$ and $u^\alpha = u_2$ otherwise. This relaxed control is an optimal solution.

(c) Let us prove that $V_0^{**} = \mathrm{conv}(V_0)$. Since $V_0^{\mathrm{r}}(x,\cdot) = \mathrm{conv}(V_0)(x,\cdot)$, it suffices to prove that $\mathrm{conv}(V_0)(x,\cdot)$ is lower semi-continuous. Since it is convex (w.r.t. z), it is continuous on the interior of the domain, which is $(-\infty, Z_0^+(x)]$. Let us check that it is l.s.c. at $\bar{z} := Z_0^+(x)$. Let $z^k \uparrow \bar{z}$, let us assume that $\lim_k \mathrm{conv}(V_0)(x,z^k) < \mathrm{conv}(V_0)(x,\bar{z})$. By (b), there exist for all k real numbers $z_1^k \leq z_2^k \leq \bar{z}$ and $\alpha_k \in [0,1]$ such that

$$\mathrm{conv}(V_0)(x,z_k) = \alpha_k V_0(x,z_1^k) + (1-\alpha_k)V_0(x,z_2^k).$$

We let the reader check that there exists a subsequence $(k_q)_{q\in\mathbb{N}}$ satisfying

$$z_2^{k_q} \to \bar{z} \quad \text{and} \quad \big\{z_1^{k_q} \to \bar{z} \text{ or } \alpha_{k_q} \to 0\big\}.$$

Since $V_0(x,\cdot)$ is l.s.c. and lower bounded, we get that

$$\lim_k \mathrm{conv}(V_0)(x,z^k) \geq V_0(x,\bar{z}) \geq \mathrm{conv}(V_0)(x,\bar{z}),$$

in the two cases ($z_1^{k_q} \to \bar{z}$ or $\alpha_{k_q} \to 0$). We obtain a contradiction, thus $\mathrm{conv}(V_0)(x,\cdot)$ is l.s.c. at $\bar{z}$ and finally, $V_0^{**} = \mathrm{conv}(V_0)$. □

Lemma 10 states that the Legendre-Fenchel transform of the value function is itself the value function of an unconstrained optimal control problem. The associated dynamic programming principle is the following: for all $\lambda \geq 0$, for all $x \in S$,

$$- V_T^*(x,\lambda) = \phi(x) - \lambda g(x), \tag{3.7}$$

$$- V_t^*(x,\lambda) = \inf_{u \in U} \Big\{ \sum_{i=1}^I p_i V_{t+1}^*\big(F(x,u,i),\lambda\big) \Big\}, \quad \forall t = T-1,...,0. \tag{3.8}$$

Note that λ must not be considered as a state variable. Consider now a measurable mapping $(t,x,\lambda) \in \{0,...,T-1\} \times S \times \mathbb{R}_+ \mapsto v(t,x,\lambda) \in U$ which is such that for all (t,x,λ),

$$v(t,x,\lambda) \in \operatorname*{arg\,min}_{u \in U} \Big\{ \sum_{i=1}^I p_i V_{t+1}^*\big(F(x,u,i),\lambda\big) \Big\}. \tag{3.9}$$

This mapping may be non unique. For all $x \in S$, $\lambda \geq 0$, there is a unique control process $u(x,\lambda) \in \mathcal{U}_0$ which is such that for all $s \geq 0$,

$$u_s(x,\lambda) = v(s, X_s^{0,x,u(x,\lambda)}, \lambda), \quad \text{almost surely.} \tag{3.10}$$

This control is an optimal solution to (3.6). Moreover, the associated level for the constraint, that we denote $Z_0(x,\lambda) = \mathbb{E}\big[g(X_T^{0,x,u(x,\lambda)})\big]$, can be computed by solving the following dynamic programming equations: for all $x \in S$,

$$Z_T(x,\lambda) = g(x), \tag{3.11}$$

$$Z_t(x,\lambda) = \sum_{i=1}^I p_i Z_{t+1}(F(x, v(t,x,\lambda), i)), \quad \forall t = T-1,...,0. \tag{3.12}$$

Note that by lemma 27, for all $0 \leq \lambda_1 < \lambda_2$, $Z_t(x,\lambda_1) \leq Z_t(x,\lambda_2)$. Moreover, as in remark 8, a small perturbation of V_{t+1}^* may generate a large perturbation of the mapping $x \mapsto v(t,x,\lambda)$ and therefore a large perturbation of $Z_t(x,\lambda)$.

We finish this section with a nice property satisfied by the subdifferential of the value function.

Theorem 11. *Let $x \in S$, $z \in \mathbb{R}$, let $u \in \mathcal{U}_0^{\mathrm{r}}$ be an optimal solution to the relaxed problem, and let $\lambda \geq 0$ be in $\partial V_0^{\mathrm{r}}(x,z)$. Then, for all associated submartingale Z,*

$$\lambda \in \partial V_s(X_s^{0,x,u}, Z_s), \quad \textit{almost surely.} \tag{3.13}$$

Proof. By lemma 28, $\lambda \in \partial V_0(x, Z_0^\alpha)$ almost surely and for almost all $\alpha \in [0,1]$, u^α is an unrelaxed optimal solution to (1.6)-(1.7) (with the level Z_t^α). By lemma 25, u^α is a solution to the penalized problem (3.6) (with level Z_0^α), and Z^α is an associated submartingale by lemma 9. We consider now two cases.

Let us assume that $Z_0^\alpha = \mathbb{E}\big[g(X_T^{0,x,u^\alpha})\big]$. In this case, $Z_T^\alpha = g(X_T^{0,x,u^\alpha})$ almost surely and Z^α is a martingale. Let us introduce a technical notation. Let $s \geq 0$, let us define for all $(\beta_1,...,\beta_s) \in \{1,...,I\}^s$ the unique control $u^{\alpha,\beta_1,...,\beta_s}$ in $\mathcal{U}_s$ which is such that

$$u^{\alpha,\xi_1,...,\xi_s} = u^\alpha, \quad \text{almost surely.}$$

Then, $u^{\alpha,\xi_1,...,\xi_s}$ is a solution to the penalized problem, starting at time s, point X_s^{t,x,u^α}, by dynamic programming. Since $Z_s^\alpha = \mathbb{E}\big[g(X_T^{t,x,u^\alpha})|\mathcal{F}_s\big]$, we obtain (3.13), with lemma 25. This proves the lemma in this case.

Let us assume now that $Z_0^\alpha < \mathbb{E}[g(X_T^{0,x,u^\alpha})] := z'$, which means that the constraint is not active. Then,

$$V_0^{\mathrm{r}}(x, Z_t^\alpha) = V_0(x, z') \geq V_0^{\mathrm{r}}(x, z').$$

But $V_0^{\mathrm{r}}(x, \cdot)$ is convex and non-decreasing, thus for all $\hat{z} \leq z'$, $V_0^{\mathrm{r}}(x, \hat{z}) = V_0^{\mathrm{r}}(z')$. This means that $\lambda = 0$, therefore u^α is an optimal solution to the minimization problem of $\mathbb{E}[\phi(X_T^{0,x,u})]$, meaning that $V_0(x, Z_0^\alpha) = V_0^-(x, Z_0^\alpha)$, and therefore, $V_s(X_s^{t,x,u^\alpha}), Z_s^\alpha) = V_s^-(X_s^{t,x,u^\alpha})$, thus $Z_s^\alpha \leq Z_s^-(X_s^{t,x,u^\alpha})$ and finally, we obtain that $0 \in \partial V_s(X_s^{0,x,u^\alpha}, Z_s^\alpha)$ with (2.8). This proves the theorem in this second case. □

4 Numerical methods

In this section we analyse the numerical methods associated with the two approaches described in sections 2 and 3. We consider now the following additional assumption.

Assumption 12. *The state space S and the control space U are finite; their cardinal is denoted resp. by $|S|$ and $|U|$. Moreover, the function g is bounded from above and from below. To simplify, we also assume that for all $x \in S$, $0 \leq g(x) \leq 1$.*

Note that $V_t^+(x)$, $V_t^-(x)$, $Z_t^+(x)$, $Z_t^-(x)$ can be computed exactly, with a complexity of $O(T \cdot |S| \cdot |U| \cdot I)$, thanks to the dynamic programming principles derived in section 2.

4.1 Approach based on dynamic programming

We study now the method based on dynamic programming. In particular, we analyse the discretization of the "submartingale" constraint $\sum_{i=1}^I p_i z_i \geq z$ in the dynamic programming equations (2.3)-(2.4).

Description We introduce a parameter $\delta z > 0$, chosen such that $1/\delta z \in \mathbb{N}^*$ and set

$$\mathcal{Z} := \{0, \delta z, 2\delta z, ..., 1\}. \tag{4.1}$$

For all z in $\mathcal{Z}$, we set:

$$z^- := \max \{z' \in \mathcal{Z} \,|\, z' < z\} \tag{4.2}$$

In a few words, the first method that we consider consists in computing an associated sub-martingale taking its values in $\mathcal{Z}$. We simply solve the following dynamic programming equations: for all $x \in S$, for all $z \in \mathcal{Z}$,

$$V_T^{\mathrm{DP1}}(x, z) = V_T(x, z) \tag{4.3}$$

$$V_t^{\mathrm{DP1}}(x, z) = \inf_{\substack{u \in U, \, (z_i)_i \in \mathcal{Z}^I \\ \sum_{i=1}^I p_i z_i \geq z}} \sum_{i=1}^I p_i V_{t+1}^{\mathrm{DP1}}\big(F(x, u, i), z_i\big), \quad \forall t = T-1, ..., 0. \tag{4.4}$$

Complexity and error analysis

Definition 13. *Let W and $\tilde{W} : \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ be two given functions, let $k \in \mathbb{N}^*$. We say that $\tilde{W}$ is a $k\delta z$-approximation of W if for all $z \geq k\delta z$,*

$$\tilde{W}(z - k\delta z) \leq W(z) \leq \tilde{W}(z). \tag{4.5}$$

Lemma 14. *Let $k \in \mathbb{N}^*$, let $W_1 : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ and $\tilde{W}_1 : \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ be two non-decreasing functions such that $\tilde{W}_1$ is a $k\delta z$-approximation of W_1. Let W_0 and $\tilde{W}_0 : \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ be defined by:*

$$W_0(z) := \inf_{\substack{(z_i)_i \in [0,1]^I \\ \sum_{i=1}^I z_i p_i \geq z}} \Big\{ \sum_{i=1}^I p_i W_1(z_i) \Big\}, \quad \tilde{W}_0(z) := \inf_{\substack{(z_i)_i \in \mathcal{Z}^I \\ \sum_{i=1}^I z_i p_i \geq z}} \Big\{ \sum_{i=1}^I p_i \tilde{W}_1(z_i) \Big\}. \tag{4.6}$$

Then, $\tilde{W}_0$ is a $(k+1)\delta z$-approximation of W_0.

Proof. Since $W_1 \leq \tilde{W}_1$ and since the feasible set associated with $\tilde{W}_0$ is included into the feasible set associated with W_0, we easily prove that $W_0 \leq \tilde{W}_0$. It is also easy to check that W_0 and $\tilde{W}_0$ are non-decreasing.

Now, let $z \in \mathcal{Z}$, if $W_0(z) = +\infty$, (4.5) is satisfied, otherwise, let $(z_i)_i$ be ε-optimal for the definition of $W_0(z)$ for a given $\varepsilon > 0$. Then, since $\tilde{W}_1$ and $\tilde{W}_0$ are non-decreasing,

$$\begin{aligned} \varepsilon + W_0(z) &\geq \sum_{i=1}^I p_i W_1(z_i) \geq \sum_{i=1}^I p_i W_1(z_i^-) \\ &\geq \sum_{i=1}^I p_i \tilde{W}_1(z_i^- - k\delta z) \geq \tilde{W}_0(\hat{z}') \geq \tilde{W}_0(z - (k+1)\delta z), \end{aligned}$$

where $\hat{z} = \sum_{i=1}^I p_i (z_i^- - k\delta z) \geq z - (k+1)\delta z$. Passing to the limit when $\varepsilon \to 0$, the result follows. □

The proof of the following lemma is left to the reader.

Lemma 15. *Let k, let W and $\tilde{W} : U \times \mathbb{R} \to \mathbb{R}$ be two given functions such that for all $u \in U$, $\tilde{W}(u, \cdot)$ is a $k\delta z$-approximation of $W(u, \cdot)$. Then, $\inf_{u \in U} \tilde{W}(u, \cdot)$ is a $k\delta z$-approximation of $\inf_{u \in U} W(u, \cdot)$.*

Theorem 16. *The approximation of the value function, $V_t^{DP1}(x, z)$ can be computed in $O(T \cdot |S| \cdot |U| \cdot I|\mathcal{Z}|^I)$ operations. The following estimate holds:*

$$V_t^{DP1}\big(x, z - (T-t)\delta z\big) \leq V_t(x, z) \leq V_t^{DP1}(x, z). \tag{4.7}$$

Proof. It is clear that the complexity is linear with respect to T, $|S|$, and $|U|$. For the minimization w.r.t. $(z_i)_i$ of the r.h.s. of (4.4), one has to examine for a given z all the possible values of z_1, z_2,...,z_{I-1}, and to chose:

$$z_I = \big((1 - p_1 z_1 - ... - p_{I-1} z_{I-1})/p_I\big)^- + \delta z.$$

There are thus $\mathcal{Z}^I$ combinations to take into account and the evaluation of the cost for each of them has a complexity of $O(I)$. The estimate (4.7) can be easily proved by induction with lemmas 14 and 15. □

A variant We describe a variant for the minimization with respect to the variables $(z_i)_i$ in (2.4). For the moment, we simply consider a family of mappings $W_i : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, for $i = 1, ..., I$. Let us define:

$$W(z) := \inf_{\substack{(z_i)_i \in \mathbb{R}^I, \\ \sum_{i=1}^I p_i z_i \geq z}} \sum_{i=1}^I p_i W_i(z_i). \tag{4.8}$$

The associated problem is a non-linear knapsack problem, for which a dynamic programming approach can be used. For all non-empty subset Ω of $\{1, ..., I\}$, we set

$$P_\Omega := \sum_{i\in\Omega} p_i, \quad W_\Omega(z) := \inf_{\substack{(z_i)_{i\in\Omega}\in\mathbb{R}^\Omega,\\ \sum_{i\in\Omega} p_i z_i \geq P_\Omega z}} \frac{1}{P_\Omega} \sum_{i\in\Omega} p_i W_i(z_i). \tag{4.9}$$

Then, for all nonempty subset Ω, for all partition (Ω_1, Ω_2) of Ω,

$$W_\Omega(z) = \inf_{\substack{z_1,z_2\in\mathbb{R},\\ P_{\Omega_1}z_1 + P_{\Omega_2}z_2 \geq P_\Omega z}} \frac{1}{P_\Omega}\big\{P_{\Omega_1}W_{\Omega_1}(z_1) + P_{\Omega_2}W_{\Omega_2}(z_2)\big\}. \tag{4.10}$$

This relation has the form of an inf-convolution and can be considered as a dynamic programming principle with respect to the random events. This means that $W(z)$ ($= W_{\{1,...,I\}}(z)$) can be computed by induction. To this purpose, we need to build a balanced binary tree as follows: each node Ω is a non-empty subset of $\{1, ..., I\}$, and has either cardinal 1 or has two non-empty children Ω_1 and Ω_2 which are such that:

$$(\Omega_1, \Omega_2) \text{ is a partition of } \Omega \quad \text{and} \quad \big||\Omega_1| - |\Omega_2|\big| \leq 1. \tag{4.11}$$

The root is the whole set $\{1, ..., I\}$. The dynamic programming equations for the variant are, $\forall x \in S$, $\forall z \in \mathcal{Z}$,

$$V_T^{\text{DP2}}(x,z) = V_T(x,z), \tag{4.12}$$

$$V_{t,\Omega}^{\text{DP2}}(x,u,z) = \begin{cases} V_{t+1}^{\text{DP2}}(F(x,u,i),z) \text{ if } \Omega = \{i\}; \quad \text{otherwise:} \\ \inf\limits_{z_1,z_2\in\mathcal{Z}} \dfrac{1}{P_\Omega}\big\{P_{\Omega_1}V_{t,\Omega_1}^{\text{DP2}}(x,u,z_1) + P_{\Omega_2}V_{t,\Omega_2}^{\text{DP2}}(x,u,z_2) \,| \\ \qquad\qquad\qquad\qquad\qquad\qquad P_{\Omega_1}z_1 + P_{\Omega_2}z_2 \geq P_\Omega z\big\}, \end{cases} \tag{4.13}$$

$$V_t^{\text{DP2}}(x,z) = \inf_{u\in U} V_{t,\{1,...,I\}}^{\text{DP2}}(x,u,z) \quad \forall t = T-1, ..., 0. \tag{4.14}$$

Theorem 17. *The approximation $V_t^{DP2}(x,z)$ of the value function can be computed in $O(T \cdot |S| \cdot |U| \cdot |\mathcal{Z}|^2 \cdot I)$ operations. The following estimate holds:*

$$V_t^{DP2}\big(x, z - (T-t)\delta z\lceil \ln_2(I)\rceil\big) \leq V_t(x,z) \leq V_t^{DP2}(x,z). \tag{4.15}$$

Proof. For any subset Ω which is not a singleton, for given values of t, x, u, the complexity of computation of $W_{t,\Omega}(x,u,z)$ is of order $\mathcal{Z}^2$ plus the complexity of computation of $W_{t,\Omega_1}(x,u,z)$ and $W_{t,\Omega_2}(x,u,z)$. Since there are I nodes in the tree, the complexity of $W_{t,\{1,...,I\}}(x,u,z)$ is of order $\mathcal{Z}^2 I$.

The error estimate (4.15) is deduced from lemmas 14 and 15. The factor $\lceil \ln_2(I)\rceil$ is the heigth of the tree (which is balanced). □

Remark 18. *Given $\varepsilon > 0$, one has to choose δz such that $\delta z \leq \varepsilon/T$ (resp. $\delta z \leq \varepsilon/(T\lceil \ln_2(I)\rceil)$), in order to ensure an ε-approximation for the first (resp. second) method, the corresponding complexities are given by:*

$$O(T^{I+1}|S| \cdot |U| I\varepsilon^{-I}) \quad (\textit{resp. } O(T^3|S| \cdot |U| I \ln^2(I)\varepsilon^{-2})). \tag{4.16}$$

Obviously, the second method is more efficient.

Taking into account the boundaries The dynamic programming described above can be quicklier computed if one takes into account the boundaries $Z_t^-(x)$ and $Z_t^+(x)$, as well as the corresponding values $V_t^-(x)$ and $V_t^+(x)$. This also allows to exclude the high values of the variables z for which $V_t(x,z) = +\infty$. We first define:

$$Z_{t,\Omega}^+(x,u) = \begin{cases} Z_{t+1}^+(F(x,u,i)) \text{ if } \Omega = \{i\}; \quad \text{otherwise:} \\ \frac{1}{P_\Omega}\big(P_{\Omega_1} Z_{t,\Omega_1}^+(x,u) + P_{\Omega_2} Z_{t,\Omega_2}^+(x,u)\big), \end{cases}$$

$$V_{t,\Omega}^+(x,u) = \begin{cases} V_{t+1}^+(F(x,u,i)) \text{ if } \Omega = \{i\}; \quad \text{otherwise:} \\ \frac{1}{P_\Omega}\big(P_{\Omega_1} V_{t,\Omega_1}^+(x,u) + P_{\Omega_2} V_{t,\Omega_2}^+(x,u)\big). \end{cases}$$

Changing the sign $+$ into the sign $-$, we define similarly $Z_{t,\Omega}^-(x,u)$ and $V_{t,\Omega}^-(x,u)$. If Ω has more than one element, we replace the formula (4.13) by the following one, in order to compute $V_{t,\Omega}^{\text{DP2}}(x,u,z)$:

$$\inf_{z_1 \in \mathcal{Z}_1,\, z_2 \in \mathcal{Z}_2} \frac{1}{P_\Omega}\big\{P_{\Omega_1} W_{t,\Omega_1}(x,u,z_1) + P_{\Omega_2} W_{t,\Omega_2}(x,u,z_2) \,|\, P_{\Omega_1} z_1 + P_{\Omega_2} z_2 \geq P_\Omega z\big\},$$

where, for $i = 1$ or 2:

$$\begin{aligned} &\mathcal{Z}_i = \big(\mathcal{Z} \cap [Z_{t,\Omega_i}^-(x,u), Z_{t,\Omega_i}^+(x,u)]\big) \cup \big\{Z_{t,\Omega_i}^-(x,u), Z_{t,\Omega_i}^+(x,u)\big\}, \\ &V_{t,\Omega_i}^{\text{DP2}}(x,u,Z_{t,\Omega_i}^-(x,u)) = V_{t,\Omega_i}^-(x,u), \\ &V_{t,\Omega_i}^{\text{DP2}}(x,u,Z_{t,\Omega_i}^+(x,u)) = V_{t,\Omega_i}^+(x,u). \end{aligned}$$

This method enables therefore to use the information contained by the boundaries and enables to reduce the cardinal of the sets $\mathcal{Z}_1$ and $\mathcal{Z}_2$, thus enables to save time. Note that theorem 17 still holds if V_t^{DP2} is computing by taking into account the boundaries.

4.2 Approach based on Lagrange relaxation

The approach based on Lagrange relaxation is mainly based on the dynamic programming principle associated with $V_t^*(x,\lambda)$. We discuss in this section the choice of the values of λ for which we compute this conjugate function and how we derive then an approximate relaxed optimal solution, thanks to lemma 26.

Given a value of λ, the conjugate $V_t^*(x,\lambda)$ can be computed with a complexity of $O(T \cdot |S| \cdot |U| \cdot I)$, by (3.7)-(3.8). An associated optimal control provides a solution to the unrelaxed problem with the level $Z_t(x,\lambda)$ (which can be computed with the same complexity), as it is explained in section 3.2.

Let $\Lambda > 0$, let $K \in \mathbb{N}^*$, let us set $\delta\lambda = \Lambda/K$ and let us assume having computed $V_0^*(x,\lambda)$ and $Z_0(x,\lambda)$ for $\lambda = 0, \delta\lambda, ..., \Lambda$. An upper estimate $\hat{V}_0(x,z)$ of $V_0^{\text{r}}(x,z)$ is given by the convex envelope (w.r.t. z) of the following function:

$$W_0(x,z) := \begin{cases} V_0^-(x,0) & \text{if } z = 0 \\ k\delta\lambda - V_0^*(x,\delta\lambda) & \text{if } z = Z_0(x,k\delta\lambda) \\ +\infty & \text{otherwise.} \end{cases} \tag{4.17}$$

For all $x \in S$, $z \in [0,1]$, if $\hat{V}_0(x,z) < +\infty$, then there exists k, k', $\alpha \in [0,1]$ such that $z = \alpha Z_0(x,k\delta\lambda) + (1-\alpha) Z_0(x,k'\delta\lambda)$, by Carathéodory's theorem. Therefore, an approximate relaxed optimal solution is given by any relaxed control, equal to u^k with probability α and to $u^{k'}$ with probability $(1-\alpha)$, where u^k and $u^{k'}$ are the optimal control involved in $Z_0(x, k\delta\lambda)$ and $Z_0(x, k'\delta\lambda)$. A lower estimate of $V_0^{\mathrm{r}}(x,z)$ is given by:

$$\check{V}_0(x,z) = \sup_{k=0,\ldots,K} \{k\delta\lambda z - V_0^*(x, k\delta\lambda)\}. \tag{4.18}$$

Theorem 19. *The following bound on the error holds: for all x and for all $z \leq Z_0(x, K\delta\lambda)$,*

$$\hat{V}_0(x,z) - \check{V}_0(x,z) \leq \frac{\delta\lambda}{4} \tag{4.19}$$

Proof. The theorem is direct consequence of lemma 26. □

It is also possible to solve the problem for fixed values of x and z with a bisection approach. Let us fix x and z. Let $\lambda_0^- < \lambda_0^+$ be such that:

$$Z_0(x, \lambda_0^-) \leq z \leq Z_0(x, \lambda_0^+). \tag{4.20}$$

The procedure builds an increasing sequence $(\lambda_k^-)_k$ and $(\lambda_k^+)_k$ as follows: at iteration k, if $z = Z_k(x, \lambda_k^-)$ or if $z = Z_k(x, \lambda_k^+)$, stop, otherwise set:

$$(\lambda_{k+1}^-, \lambda_{k+1}^+) = \begin{cases} \left(\lambda_k^-, \frac{\lambda_k^- + \lambda_k^+}{2}\right) & \text{if } z \leq Z_0\left(x, \frac{\lambda_k^- + \lambda_k^+}{2}\right), \\ \left(\frac{\lambda_k^- + \lambda_k^+}{2}, \lambda_k^+\right) & \text{otherwise.} \end{cases} \tag{4.21}$$

At any iteration k, an upper and a lower estimate can be computed, with the following bound on the error: $\frac{\lambda_0^+ - \lambda_0^-}{4 \cdot 2^k}$. Let us discuss the choice of λ_0^- and λ_0^+.

If $z \leq Z_0^-(x)$, any solution to the problem without constraint is an optimal solution. Otherwise, we take $\lambda_0^- = 0$. A possible choice for λ_0^+ is the following:

$$\lambda(x,z) := \frac{V_0^+(x,z) - V_0^-(x,z)}{Z_0^+(x,z) - z}, \tag{4.22}$$

if $z < Z_0^+(x)$ (otherwise, if $z > Z_0^+(x)$, the problem is infeasible and if $z = Z_0^+(x)$, the solution can be found by computing $V_0^+(x)$). Let us justify that $Z_0(x, \lambda(x,z)) \geq z$. Let $\lambda \geq 0$, then

$$V_0^*(x,\lambda) = \lambda Z_0(x,\lambda) - V_0^{\mathrm{r}}\big(x, Z_0(x,\lambda)\big) \leq \lambda Z_0(x,\lambda) - V_0^-(x). \tag{4.23}$$

Moreover,

$$V_0^+(x) = V_0(x, Z_0^+(x)) \geq \lambda Z_0^+(x) - V_0^*(x,\lambda), \tag{4.24}$$

therefore, combined with (4.23), we obtain that:

$$V_0^+(x) - V_0^-(x) \geq \lambda(Z_0^+(x) - Z_0(x,\lambda)) \tag{4.25}$$

and finally that $Z_0(x,\lambda) \geq Z_0^+(x) - \frac{V_0^+(x) - V_0^-(x)}{\lambda}$, proving that the suggested value is suitable.

We finish this section with a remark comparing the two approaches.

Remark 20. *1. The error estimate for the approach based on Lagrange relaxation does not increase with the number of time steps.*

2. The approach based on Lagrange relaxation allows to focus on a given value of z (or actually even a given subinterval of $[0,1]$), for a given initial point x, which is not possible with the dynamic-programming approach. Moreover, the computation of $V_0^*(x,\lambda)$ for different values of λ can be parallelized.

5 Continuous-time problems

In this section, we formulate a continuous-time stochastic optimal control problem with an expectation constraint. In theorem 22, we prove that if the function g is Lipschitz continuous, the value function of the unrelaxed problem is already convex with respect to z.

5.1 Formulation and convexity of the value function

Let n, m, d in $\mathbb{N}^*$, let $T > 0$, let W be a d-dimensional Brownian motion. We denote by $\mathbb{F}_t = (\mathcal{F}_s)_{s\in[t,T]}$ the filtration generated by $(W_s)_s$ on the interval (t, T). Let U be a compact subset of $\mathbb{R}^m$, let $\mathcal{U}$ the set of $\mathbb{F}$-adapted processes with value in U. For $t \in [0, T]$, $x \in \mathbb{R}^n$, $u \in \mathcal{U}$, consider the following stochastic differential equation on $[0, T]$:

$$\begin{cases} \mathrm{d}X_s^{t,x,u} = & f(X_s^{t,x,u}, u_s)\mathrm{d}s + \sigma(X_s^{t,x,u}, u_s)\mathrm{d}W_s, \\ X_t^{t,x,u} = & x. \end{cases} \tag{5.1}$$

We aim at solving:

$$V_t(x, z) = \inf_{u\in\mathcal{U}_t} \mathbb{E}\big[\phi(X_T^{t,x,u})\big], \quad \text{s.t. } \mathbb{E}\big[g(X_T^{t,x,u}) \geq 0\big] \geq z, \tag{5.2}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a given mapping.

Assumption 21. *The mappings $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^k$, $\phi : \mathbb{R}^n \to \mathbb{R}$ satisfy the classical regularity assumptions: there exists $L > 0$ such that for all $x, y \in \mathbb{R}^n$, for all $u \in U$,*

$$|f(x, u)| + |\sigma(x, u)| + |\phi(x)| \leq L(1 + |x|), \tag{5.3}$$

$$|f(x, u) - f(y, u)| + |\sigma(x, u) - \sigma(y, u)| + |\phi(x) - \phi(y)| \leq L|y - x|, \tag{5.4}$$

so that the stochastic differential equation is well-posed [13, Section 5]. We assume that for all $x \in \mathbb{R}^n$, $|g(x)| \leq L(1 + |x|)$.

In the next theorem, we prove that the value function is convex with respect to z, if g is Lipschitz. This property is strongly linked to the relaxation technique studied in section 3. Let us give the main idea. Let u_1 and u_2 be two controls ensuring the levels $z_1 < z_2$. In a relaxed framework, we would prove the convexity by using the two controls u_1 and u_2 with probability 1/2. In an unrelaxed framework, the idea consists in observing the Brownian motion during a very short time, in order to decide which control should be used.

Theorem 22. *If g is uniformly Lipschitz, then for all $(t, x) \in [0, T]$, the mapping $z \mapsto V(t, x, z)$ is convex.*

Proof. Let $t \in [0, T)$, let $x \in \mathbb{R}^n$, let z_1 and z_2 be such that $V(t, x, z_1) < +\infty$, $V(t, x, z_2) < +\infty$, let $z = (z_1 + z_2)/2$. Let u^1 and u^2 in $\mathcal{U}_t$ be such that

$$\mathbb{E}\big[g(X_T^{t,x,u^1})\big] \geq z_1 \quad \text{and} \quad \mathbb{E}\big[g(X_T^{t,x,u^2})\big] \geq z_2.$$

Let $\varepsilon > 0$. Let $\tilde{u}^1 \in \mathcal{U}_{t+\varepsilon}$ be such that the two following processes:

$$(u_{t+s}^1)_{s\in[0,T-(t+\varepsilon)]} \quad \text{and} \quad (\tilde{u}_{t+\varepsilon+s}^1)_{s\in[0,T-(t+\varepsilon)]}$$

have the same law. In other words, $\tilde{u}^1$ is obtained by delaying u^1 of ε. We similarly define $\tilde{u}^2$. Let $p \in [0,1]$, let $w(p)$ be such that $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{w(p)} e^{-\theta^2/2} \mathrm{d}\theta = p$, let A_p be the following event: $W^1_{t+\varepsilon} - W^1_t \leq \sqrt{\varepsilon} w(p)$ so that $\mathbb{P}\big[A_p\big] = p$. Finally, let $u^0 \in U$ and let $u(p) \in \mathcal{U}_t$ be such that

$$u_t(p) = u^0, \quad \forall t \in (t, t+\varepsilon), \quad u_t(p) = \begin{cases} \tilde{u}^1_t & \text{if } A(p) \text{ is realized} \\ \tilde{u}^2_t & \text{otherwise.} \end{cases}$$

In short, we use control u^1 with probability p and control u^2 with probability $(1-p)$, after a small delay during which we observe the Brownian motion, in order to decide which control to use. We claim that there exists a constant $C > 0$ independent of p which is such that:

$$\mathbb{E}\big[g(X_T^{t,x,u(p)})\big] \geq p z_1 + (1-p) z_2 - C\sqrt{\varepsilon}, \tag{5.5}$$

$$\mathbb{E}\big[\phi(X_T^{t,x,u(p)})\big] \leq p\mathbb{E}[\phi(X_T^{t,x,u^1})] + (1-p)\mathbb{E}[\phi(X_T^{t,x,u^2})] + C\sqrt{\varepsilon}. \tag{5.6}$$

We will prove these estimates later. Let us set $p_\varepsilon = \frac{1}{2} - C\sqrt{\varepsilon}/(z_2 - z_1)$. Then,

$$\begin{aligned} \mathbb{E}\big[g(X_T^{t,x,u(p_\varepsilon)})\big] &\geq \frac{1}{2}(z_1 + z_2), \\ \mathbb{E}\big[\phi(X_T^{t,x,u(p_\varepsilon)})\big] &\leq \frac{1}{2}\big(\mathbb{E}[\phi(X_T^{t,x,u^1})] + \mathbb{E}[\phi(X_T^{t,x,u^2})]\big) + O(\sqrt{\varepsilon}). \end{aligned}$$

To the limit when $\varepsilon \downarrow 0$, we obtain that $V(t,x,z) \leq \frac{1}{2}\big(\mathbb{E}[\phi(X_T^{t,x,u^1})] + \mathbb{E}[\phi(X_T^{t,x,u^2})]\big)$, and the result follows by minimizing the r.h.s. of the previous inequality.

Now, let us prove estimate (5.5), the proof of (5.6) being similar. First,

$$\mathbb{E}\big[g\big(X_T^{t,x,u(p)}\big)\big] = p\mathbb{E}\Big[g\Big(X_T^{t+\varepsilon, X_{t+\varepsilon}^{t,x,u^0}, \tilde{u}^1}\Big) \,\Big|\, A_p\Big] + (1-p)\mathbb{E}\Big[g\Big(X_T^{t+\varepsilon, X_{t+\varepsilon}^{t,x,u^0}, \tilde{u}^2}\Big) \,\Big|\, \bar{A}_p\Big].$$

Then, since g is Lipschitz,

$$\begin{aligned} \mathbb{E}\Big[g\Big(X_T^{t+\varepsilon, X_{t+\varepsilon}^{t,x,u^0}, \tilde{u}^1}\Big) \,\Big|\, A_p\Big] &= \mathbb{E}\big[g\big(X_T^{t+\varepsilon,x,\tilde{u}^1}\big)\big] + O\big(\mathbb{E}\big[X_T^{t+\varepsilon, X_{t+\varepsilon}^{t,x,u^0}, \tilde{u}^1} - X_T^{t+\varepsilon,x,\tilde{u}^1}\big]\big) \\ &= \mathbb{E}\big[g\big(X_T^{t+\varepsilon,x,\tilde{u}^1}\big) \,|\, A_p\big] + O\big(\mathbb{E}\big[|X_{t+\varepsilon}^{t,x,u^0} - x|\big]\big) \\ &= \mathbb{E}\big[g\big(X_{T-\varepsilon}^{t,x,u^1}\big)\big] + O(\sqrt{\varepsilon}) = \mathbb{E}\big[g\big(X_T^{t,x,u^1}\big)\big] + O(\sqrt{\varepsilon}). \end{aligned}$$

These last estimates follow from classical a priori estimates for stochastic differential equations, themselves consequences of Gronwall's lemma and Itô's isometry. A similar estimate holds for u^2 and (5.5) follows. □

Remark 23. *1. This result does not cover the case of chance-constrained problems when $g = \mathbf{1}_{\mathbb{R}_+} \circ h$. Proving the convexity (or the non-convexity) in this case is still an open question to us.*

2. The notion of relaxed controls can naturally be defined for continuous-time problems. If g is Lipschitz, relaxing does not modify the value function.

5.2 Numerical scheme

Numerical schemes for unconstrained stochastic optimal control problems consist basically in a consistent discretization of the stochastic differential equation, so that the resulting scheme is stable, monotonic, and consistent in the sense of [2]. We use the same approach here.

For all $m \in \mathbb{N}$, for all family $(x_i)_{i=1,\dots,m+1}$ in $\mathbb{R}^n$, we say that the convex envelope $\mathcal{T}$ of the set $\{x_1,...,x_{m+1}\}$ is a non-degenerate m-simplex if the family $(x_2-x_1,...,x_{m+1}-x_1)$ is linearly independent. The points x_1,...,x_{m+1} are then called vertices (of the simplex $\mathcal{T}$).

Let us fix $N_T \in \mathbb{N}^*$, let $\delta t = T/N_T$. Let $N_X \in \mathbb{N}^*$, let $S = (x_i)_{i=1,\dots,N_X}$ be a family in $\mathbb{R}^n$, let $X = \text{conv}(S)$. Let $\mathcal{T} = (\mathcal{T}_i)_{i=1,\dots,N}$ be a family of n-simplices with vertices in $\{x_1,...,x_{N_X}\}$, this family is called triangulation if $X = \cup_{i=1}^N \mathcal{T}_i$, and for all $1 \leq i < j \leq N$, if $\mathcal{T}_i \cap \mathcal{T}_j$ is non-empty, then is a non-degenerate m-simplex with $m < n$ and with all its vertices in the set of vertices of X_i and X_j.

For all $x \in X$, there exist a unique $K(x) \in \mathbb{N}^*$ and two unique families

$$(i_1(x),...,i_{K(x)}(x)) \quad \text{and} \quad (\alpha_1(x),...,\alpha_{K(x)}(x)) \tag{5.7}$$

in resp. $\{1,...,N_X\}$ and $(0,1]$, which are such that

$$x = \sum_{j=1}^{I(x)} \alpha_j(x) x_{i_j(x)}, \quad \sum_{j=1}^{I(x)} \alpha_j(x) = 1, \quad i_1 < ... < i_{I(x)}, \tag{5.8}$$

and such that there exists l for which $\{x_{i_1(x)},...,x_{i_{I(x)}(x)}\} \subset \mathcal{T}_l$. To alleviate the notations, we denote: $y_j(x) = x_{i_j(x)}$.

Finally, we discretize the SDE with a controlled Markov chain which is very close to the one studied in the previous sections, except that now, the set of random events I is a function of x and u, and the corresponding probabilities are also functions of x and u:

$$\begin{cases} \tilde{F}(x,u,i) = \begin{cases} P_X(x + f(x,u)\delta t + \sigma_i(x,u)\sqrt{\delta t}), & \text{for } i = 1,...,d, \\ P_X(x + f(x,u)\delta t - \sigma_{i-d}(x,u)\sqrt{\delta t}), & \text{for } i = d+1,...,2d, \end{cases} \\ I(x,u) = \{(i,j) \,|\, i = 1,...,2d,\ j = 1,...,K(x_i)\}, \\ F(x,u,i,j) = y_j(\tilde{F}(x,u,i)), \\ p(x,u,i,j) = \alpha_j(\tilde{F}(x,u,i))/(2d), \end{cases} \tag{5.9}$$

where P_X is the projection on X. Observe that $|I(x,u)| \leq 2d(n+1)$.

Remark 24. *1. This approach suffers from the curse of dimensionality, since N_X increases exponentially with n for a good discretization of $\mathbb{R}^n$.*

2. The theoretical results proved in the previous sections remain true, even if the dynamics F and the probabilities p depend on x and u.

3. The control space must be discretized, in order to solve the dynamic programing principles.

6 Numerical experiments

In this section, we present some numerical results for the following problem:

$$V_0(x,z) = \frac{1}{T} \inf_{u\in\mathcal{U}_0} \mathbb{E}\Big[\int_t^T u_s^2 \mathrm{d}s\Big], \quad \text{subject to: } \mathbb{P}[X_T^{0,x,u} \geq 0] \geq z, \tag{6.1}$$

with the dynamics $\mathrm{d}X_t = u_t \mathrm{d}t + \mathrm{d}W_t$, $u_t \in U := [0,1]$. The problem is discretized with the following parameters:

- $T = 10$, $\delta t = 0,5$, $N_T = 20$,
- $U = \{0; 0,1; 0,2; ...; 1\}$,
- $S = \{-20; -19; ...; 20\}$, $X = [-20, 20]$, $(\mathcal{T}_i)_{i=1,...,40} = \{[-20, -19], ..., \}$.

The value function of the discretized problem is denoted by $\bar{V}_t(x,z)$, for $t = 0$; $0,5$; ...; 10, $x \in S$, $z \in [0,1]$.

Boundaries The boundaries of the discretized problem $\bar{Z}_t^-(x)$ and $\bar{Z}_t^+(x)$, as well as the corresponding costs can be easily computed, since they are reached for the constant control strategies resp. equal to 0 and 1. The boundaries of the continuous-time problem are given by the following formulas:

$$Z_t^+(x) = \sup\nolimits_{u\in\mathcal{U}_t} \mathbb{P}\big[X_T^{t,x,u} \geq 0\big] = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x+T-t}{\sqrt{T-t}}} e^{-\theta^2/2} \mathrm{d}\theta, \quad V_t^+(x) = (T-t),$$
$$Z_t^-(x) = \sup\{z \,|\, V_t^-(x,z) = 0\} = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x}{\sqrt{T-t}}} e^{-\theta^2/2} \mathrm{d}\theta, \quad V_t^-(x) = 0.$$

Figure 1 represents the boundaries $\bar{Z}_t^-(x)$, $Z_t^-(x)$, $\bar{Z}_t^+(x)$ and $Z_t^+(x)$ (resp. in dark blue, green, black, and blue).

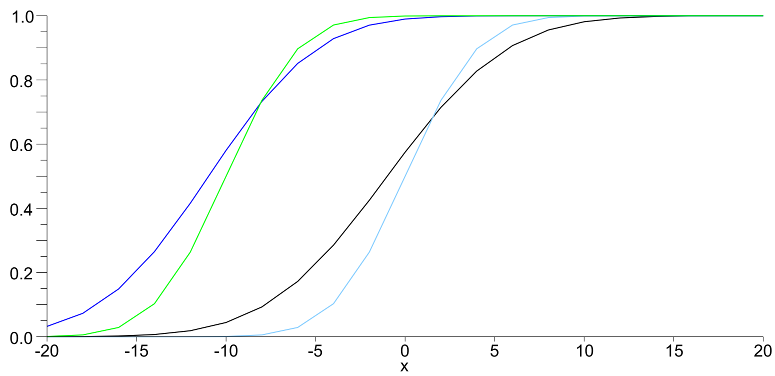

Figure 1: Boundaries (at time 0).

Comparison of the two approaches A representation of $\bar{V}_0(\cdot,\cdot)$ is shown on figure 2. The bold lines correspond to the boundaries $\bar{Z}_0^-$ and $\bar{Z}_0^+$. The problem is infeasible in the dark area. Observe that for $x < -5$, the problem is infeasible, except for small values of z. To the contrary, when $x > 5$, the value of the problem is 0, except for high values of z.

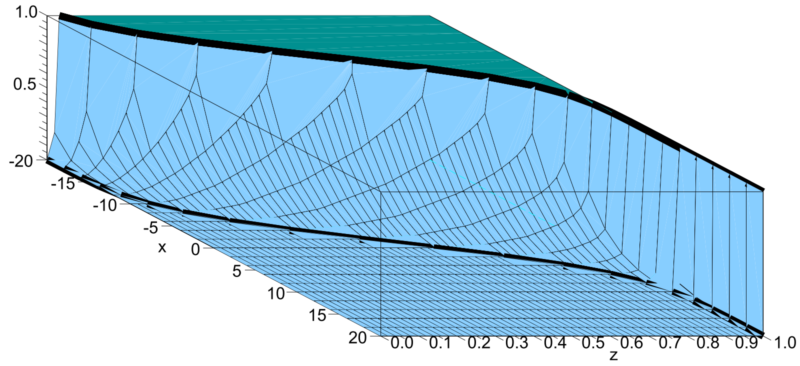

Figure 2: Value function (at time 0).

Dynamic programming				Lagrange relaxation		
$1/\delta z$	Error	Time		K	Error	Time
50	0,63	2,2		10	1,58	8,19
100	0,361	3,9		25	1,07	17,2
250	$1{,}72.10^{-1}$	9,3		50	$5{,}06.10^{-1}$	33,8
500	$8{,}91.10^{-2}$	20		100	$1{,}54.10^{-1}$	60,5
1000	$4{,}74.10^{-2}$	50		250	$2{,}72.10^{-2}$	151
2 500	$1{,}96.10^{-2}$	193		500	$7{,}33.10^{-3}$	295
5 000	$9{,}81.10^{-3}$	637		1 000	$1{,}75.10^{-3}$	584
10 000	$5{,}4.10^{-3}$	2 130		2 500	$2{,}92.10^{-4}$	1 460
25 000	$2{,}32.10^{-3}$	11 500		5 000	$5{,}66.10^{-5}$	2 900
50 000	$1{,}25.10^{-3}$	43 500		10 000	0	5 820

Table 1: Error and computation time for the two approaches.

The variant of the approach by dynamich programming has been tested; our problem has 4 random events, at each time step ($2d(n+1)$, with $d = n = 1$), which can be represented in a tree of heigth 2. The original method (without the tree representation) has been discarded, considering its complexity ($O(\delta z^{-4})$) with respect to δz. The problem has been solved for various values of δz.

The lower estimate $\check{V}_0(x,z)$ of the relaxed value function has been computed for the fixed value $\Lambda = 100$ and various values of K. We have observed that for the values: $\delta z^{-1} = 5.10^4$, and $K = 10^4$, the error, measured as an L^1-norm (on the domain of the functions) is very small:

$$\|V_0^{\mathrm{DP2}}(\cdot,\cdot) - \check{V}_0(\cdot,\cdot)\|_1 \approx 1,25.10^{-3}. \tag{6.2}$$

Remember that $V_0^{\mathrm{DP2}}(\cdot,\cdot) \geq \bar{V}_0(\cdot,\cdot) \geq \check{V}_0(\cdot,\cdot)$. This means that for this problem, the duality gap is very small. We have therefore used $\check{V}_0(\cdot,\cdot)$ as a reference value function for the measure of the error in table 1. The time is measured in seconds, the error is the L^1-norm at time 0. The time required for the dynamic programming approach (resp. the approach based on Lagrange relaxation) are approximately:

$$2.10^{-5}\delta z^{-2} \quad (\text{resp. } K(4, 4.10^{-4} + 1, 2.10^{-5}\delta z^{-1})).$$

The constant $4, 4.10^{-4}$ corresponds to the time needed to compute once the conjugate value function. This time is negligible for $\delta z < 0, 05$.

Conclusion On this academic example, both methods are tractable, but the approach based on Lagrange relaxation is obviously more efficient. As we already mentioned, this approach has a certain number advantages:

- it is much easier to program (and can be vectorized), and many different approaches (not necessarily based on the discretization of the SDE) from the litterature can be used
- the strategy for finding the suitable values of λ can be improved
- the computation of the conjugate value function can be parallelized, for different values of λ.

As a future work, this study coud be extended to the case with several constraints, and some error estimates for continuous-time problems could be derived, at least when g is regular enough.

A Appendix: generalities on Lagrange relaxation

A.1 First definitions

This section is a short introduction to some notions of convex analysis [15] and Lagrange relaxation [10, Chapter XII]. The notations that we use here are independent of the article. Let $V : z \in \mathbb{R} \mapsto \mathbb{R} \cup \{+\infty\}$. We denote by V^* its *Legendre-Fenchel transform*, defined for all $\lambda \in \mathbb{R}$ by

$$V^*(\lambda) := \sup_{z \in \mathbb{R}} \big\{ \lambda z - V(z) \big\}. \tag{A.1}$$

The function V^* is convex and lower semi-continuous and if $z \mapsto V(z)$ is nondecreasing, which is the case for the considered functions in the article, then $V^*(\lambda) = +\infty$, for all $\lambda < 0$. By definition, for all z and λ,

$$V(z) + V^*(\lambda) \geq z\lambda. \tag{A.2}$$

The *subdifferential* ∂V is the multimapping, possibly empty, defined for all $z \in \mathbb{R}$ by

$$\partial V(z) := \{\lambda \in \mathbb{R} \, : \, V(z') - V(z) \geq \lambda(z' - z), \, \forall z' \in \mathbb{R}\}. \tag{A.3}$$

It is easy to check that for all z and λ,

$$V(z) + V^*(\lambda) = z\lambda \Longleftrightarrow \lambda \in \partial V(z) \Longleftrightarrow z \in \partial V^*(\lambda) \Longrightarrow V(z) = V^{**}(z), \tag{A.4}$$

where V^{**} is the Legendre-Fenchel transform of V^*, also called *biconjugate* of V. Finally, we denote by $\text{conv}(V)$ the *convex envelope* of V, defined as the greatest convex function smaller than or equal to V. We denote by $\overline{\text{conv}}(V)$ the *l.s.c. convex envelope* of V, defined as the greatest lower semi-continuous convex function smaller or equal than V. It is easy to check that $\text{conv}(V)^* = \overline{\text{conv}}(V)^* = V^*$. By the Fenchel-Moreau-Rockafellar Theorem, if $\overline{\text{conv}}(V)$ is proper (that is to say, if for all z, $\overline{\text{conv}}(V)(z) > -\infty$ and if for at least one z, $\overline{\text{conv}}(V)(z) < +\infty$), then

$$\overline{\text{conv}}(V) = V^{**}. \tag{A.5}$$

A.2 Duality

Let us consider now an abstract family of optimization problems:

$$V(z) := \inf_{x \in S} f(x) \quad \text{s.t. } g(x) \geq z, \tag{A.6}$$

where $S \subset \mathbb{R}^k$ is a given non-empty set, with $k \in \mathbb{N}^*$ and where the functions $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ are given. The function V is called *value function*. Equivalently,

$$V(z) = \inf_{x \in S} \sup_{\lambda \geq 0} \big\{ f(x) - \lambda(g(x) - z) \big\}. \tag{A.7}$$

Let us compute the Legendre-Fenchel transform of V. Let $\lambda \geq 0$, then

$$\begin{aligned} V^*(\lambda) &= \sup_{z \in \mathbb{R}} \big\{ \lambda z - \inf_{x,\, g(x) \geq z} f(x) \big\} = \sup_{\substack{x,z: \\ z \leq g(x)}} \big\{ \lambda z - f(x) \big\} = \sup_{x} \big\{ \lambda g(x) - f(x) \big\} \\ &= -\inf_{x} \big\{ f(x) - \lambda g(x) \big\}. \end{aligned} \tag{A.8}$$

Since V is nondecreasing, for all $\lambda < 0$, $V^*(\lambda) = +\infty$. We call *penalized problem* the minimization problem in (A.8). By (A.8), V^{**} is obtained by inverting the operators inf and sup in (A.7):

$$V^{**}(z) = \sup_{\lambda \geq 0} \inf_{x \in S} \big\{ f(x) - \lambda(g(x) - z) \big\}. \tag{A.9}$$

Lemma 25. *Let $\lambda \geq 0$, let $x \in S$, let $z = g(x)$. Then,*

$$\big\{ x \text{ is a solution to (A.6) and } \lambda \in \partial V(z) \big\} \iff \big\{ x \text{ is a solution to (A.8)} \big\}. \tag{A.10}$$

In this case, $\lambda \in \partial V(z)$, $z \in \partial V^*(\lambda)$, and $V(z) = V^{**}(z)$.

Proof. Let us assume that x is a solution to (A.6) and that $\lambda \in \partial V(z)$. Then, by (A.4),

$$V^*(\lambda) = -(V(z) - \lambda z) = -(f(x) - \lambda g(x)),$$

therefore, x is a solution to (A.8). Let us assume now that x is a solution to (A.8). Let x' be such that $g(x') \geq z$. Then,

$$f(x') \geq f(x) + \lambda(g(x) - g(x')) \geq f(x),$$

which proves that x is a solution to (A.6). The last statement of the lemma is a direct consequence of (A.4). □

The following lemma and figure 3 explain how to derive a lower and an upper estimate of V^{**} by solving the penalized problem. We give a bound for the maximal error generated by these estimates.

Lemma 26. *Let $0 \leq \lambda_1 < \lambda_2$, let x_1 and x_2 in S, let $z_1 = g(x_1)$ and $z_2 = g(x_2)$. We assume that x_1 and x_2 are solutions to (A.8) for resp. $\lambda = \lambda_1$ and $\lambda = \lambda_2$. Then, $z_1 \leq z_2$. Moreover, as illustrated on figure 3 (on the left),*

$$V^{**}(z) \geq \check{V}(z) := \max_{i=1,2} \big\{ \lambda_i (z - z_i) + V(z_i) \big\}, \quad \forall z \in \mathbb{R}, \tag{A.11}$$

$$V^{**}(z) \leq \hat{V}(z) := \frac{(z - z_1)V(z_2) + (z_2 - z)V(z_1)}{z_2 - z_1}, \quad \forall z \in [z_1, z_2]. \tag{A.12}$$

Finally, an upper bound of $\hat{V} - \check{V}$ on $[z_1, z_2]$ is:

$$\Big(\lambda_2 - \frac{V(z_2) - V(z_1)}{z_2 - z_1} \Big) \Big(\frac{V(z_2) - V(z_1)}{z_2 - z_1} - \lambda_1 \Big) \frac{z_2 - z_1}{\lambda_2 - \lambda_1} \leq \frac{1}{4} (\lambda_2 - \lambda_1)(z_2 - z_1). \tag{A.13}$$

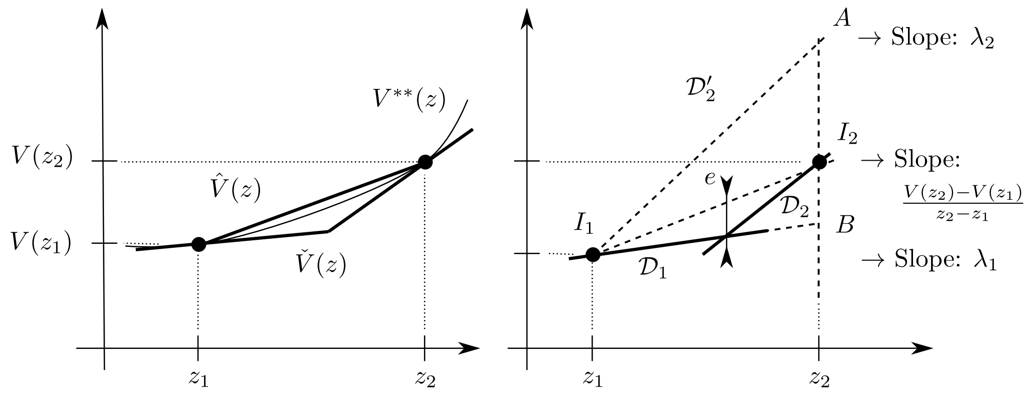

Figure 3: Lower and upper estimates of V^{**}

Proof. Observe first that

$$f(x_1) - \lambda_1 g(x_1) \le f(x_2) - \lambda_1 g(x_2), \quad f(x_2) - \lambda_2 g(x_2) \le f(x_1) - \lambda_2 g(x_1).$$

Summing these two inequalities, we obtain that

$$(\lambda_2 - \lambda_1) g(x_1) \le (\lambda_2 - \lambda_1) g(x_2),$$

which proves that $z_1 \le z_2$. By lemma 25, x_1 and x_2 are solutions to (A.6) with resp. $z = z_1$ and z_2 and $\lambda_1 \in \partial V^*(z_1)$, $\lambda_2 \in \partial V^*(z_2)$ and $V(z_1) = V^{**}(z_1)$ and $V(z_2) = V^{**}(z_2)$. The upper estimate on V^{**} follows, since V^{**} is convex. By definition and by lemma 25,

$$V^{**}(z) \ge \lambda_1 z - V_1^*(\lambda_1) = \lambda_1 (z - z_1) + V(z_1), \tag{A.14}$$

and the same holds with λ_2 and z_2; this proves the lower estimate.

We give a geometrical proof of (A.13), illustrated by figure 3 (on the right). On this figure, the coordinates of points I_1 and I_2 are resp. $(z_1, V(z_1))$ and $(z_2, V(z_2))$. The bold lines $\mathcal{D}_1$ and $\mathcal{D}_2$ correspond to the lower estimate of V^{**}, their slopes are resp. λ_1 and λ_2. The upper estimate is the segment $[I_1, I_2]$. The greatest gap between the lower and the upper estimate is given by e.

The line $\mathcal{D}_2'$ is the parallel to $\mathcal{D}_2$ passing by I_1. Finally, the points A and B are the intersections of resp. $\mathcal{D}_2'$ and $\mathcal{D}_1$ with the vertical axis passing by I_2. Note that $AB = (\lambda_2 - \lambda_1)(z_2 - z_1)$. Let us set $\alpha = \frac{BI_2}{AB}$. We have $\alpha \in [0,1]$ and $e = \alpha(1-\alpha)AB$. The value of α is given by

$$\alpha = \frac{1}{\lambda_2 - \lambda_1} \Big(\frac{V(z_2) - V(z_1)}{z_2 - z_1} - \lambda_1 \Big). \tag{A.15}$$

The first inequality in (A.13) follows, and the second one follows from the inequality: $\alpha(1-\alpha) \le 1/4$ for all $\alpha \in [0,1]$. □

A.3 Relaxation

We finish this section with two equivalent relaxed formulations of problems (A.6). We assume now that

$$V(-\infty) := \lim_{z \to -\infty} V(z) > -\infty. \tag{A.16}$$

This is equivalent to assume that the unconstrained problem $\inf_{x \in S} f(x)$ has a finite value. The value function V being non decreasing, it is thus bounded from below, and so is its convex envelope. Moreover, since V is non decreasing, its domain $\{z \,|\, V(z) < +\infty\}$ is an interval of the form $(-\infty, \bar{z})$ or $(-\infty, \bar{z}]$, therefore the convex envelope has the same domain.

Let us fix a random variable ξ, of uniform law on $[0, 1]$. We denote by $\mathcal{M}(S, \zeta)$ the space of random variables with value in S and measurable with respect to ζ. We introduce a new family of optimization problems, that we call *relaxed problems*, with value $V^{\mathrm{r}}(z)$:

$$V^{\mathrm{r}}(z) := \inf_{X \in \mathcal{M}(S,\zeta)} \mathbb{E}\big[f(X)\big] \quad \text{s.t. } \mathbb{E}\big[g(X)\big] \geq z. \tag{A.17}$$

This family is linked to the family of problems (A.6). Instead of taking one decision x, we are allowed to make depend the decision on ζ, and the constraint must only be satisfied "in expectation".

Lemma 27. *The relaxed value function V^{r} is the convex envelope of V:*

$$V^{\mathrm{r}} = \mathrm{conv}(V). \tag{A.18}$$

Proof. Let z be such that $\mathrm{conv}(V)(z) < +\infty$. By Carathéodory's theorem,

$$\mathrm{conv}(V)(z) = \inf_{\substack{z_1, z_2, \alpha \in [0,1] \\ \alpha z_1 + (1-\alpha) z_2 = z}} \alpha V(z_1) + (1-\alpha) V(z_2).$$

Let z, let $\varepsilon > 0$, let $z_1 \in \mathbb{R}$, $z_2 \in \mathbb{R}$, $\alpha \in [0, 1]$, $x_1 \in S$, $x_2 \in S$ be such that

$$\begin{aligned}
& z = \alpha z_1 + (1-\alpha) z_2, \\
& \alpha V(z_1) + (1-\alpha) V(z_2) \leq \mathrm{conv}(V)(z) + \varepsilon/2, \\
& f(x_1) \leq V(z_1) + \varepsilon/2, \;\; g(x_1) \geq z_1, \\
& f(x_2) \leq V(z_2) + \varepsilon/2, \;\; g(x_2) \geq z_2.
\end{aligned}$$

We set: $X = x_1$ if $\zeta \in (0, \alpha)$, x_2 otherwise. The variable X is feasible for (A.17) and therefore, $V^{\mathrm{r}}(z) \leq \mathrm{conv}(V)(z) + \varepsilon$. Passing to the limit, we obtain that $V^{\mathrm{r}} \leq \mathrm{conv}(z)$.

Let $z \in \mathbb{R}$ be such that $V^{\mathrm{r}}(z) < +\infty$, let $\varepsilon > 0$, let X be an ε-solution of (A.17). Let Z be the real random variable defined by $Z = g(X)$. Then,

$$\mathrm{conv}(V)(z) \leq \mathbb{E}\big[V(Z)\big] \leq \mathbb{E}\big[f(X)\big] \leq V^{\mathrm{r}}(z) + \varepsilon.$$

Passing to the limit, we obtain that $\mathrm{conv}(V) \leq V^{\mathrm{r}}$ and finally the equality. ☐

Lemma 28. *Let $z \in \mathbb{R}$, let X be a relaxed optimal solution to problem* (A.17)*. Let Z be a real random variable, adapted to ξ and such that*

$$\mathbb{E}\big[Z\big] \geq z \quad \textit{and} \quad Z \leq g(X) \;\; \textit{almost surely.} \tag{A.19}$$

Then, X is a solution to the problem (A.6) *with the level $z = Z$ almost surely. Moreover, if $\lambda \in \partial V^{\mathrm{r}}(z)$, then $\lambda \in \partial V(Z)$ and X is a solution to the penalized problem* (A.8) *almost surely (that is to say, $V^*(\lambda) = -(f(X) - \lambda g(X))$ almost surely).*

Proof. Let us prove the first part of the lemma. First, by lemma 27,

$$\mathbb{E}\big[f(X)\big] = V^{\mathrm{r}}(z) = \mathrm{conv}(V)(z) = \mathrm{conv}(V)(\mathbb{E}\big[Z\big]) \leq \mathbb{E}\big[V(Z)\big]. \tag{A.20}$$

Moreover, $V(Z) \leq f(X)$ almost surely, since $Z \leq g(X)$ almost surely. Combined with (A.20), it follows that $f(X) = V(Z)$ almost surely. Note that it follows also that $V^{\mathrm{r}}(z) = \mathbb{E}\big[V(Z)\big]$.

Let us prove the second part of the lemma. Let $\lambda \in \partial V^{\mathrm{r}}(z)$, using the fact that $V^{\mathrm{r},*} = V^*$, we obtain that

$$\mathbb{E}\big[\lambda Z - V^*(\lambda)\big] \geq \lambda z - V^{\mathrm{r},*}(\lambda) = V^{\mathrm{r}}(z) = \mathbb{E}\big[V(Z)].$$

Since $\lambda Z - V^*(\lambda) \leq V(Z)$, we obtain that $\lambda Z - V^*(\lambda) = V(Z)$ almost surely and therefore that $\lambda \in \partial V(Z)$ almost surely. Finally, X is a solution to the penalized problem almost surely by lemma 25. □

Let us denote now by $\mathcal{P}(S)$ the set of probability measures on S. With a similar proof to the one of lemma 27, we can show that

$$V^{\mathrm{r}}(z) = \inf_{\mu \in \mathcal{M}(S)} \int_S f(x)\mathrm{d}\mu(x) \quad \text{s.t.} \quad \int_S g(x)\mathrm{d}\mu(x) \geq z. \tag{A.21}$$

Instead of taking one decision, we are allowed to take several decisions simultaneously. The cost function and the constraint are now linear with respect to the new optimization variable μ.

When V^{r} is lower semi-continuous and such that $V^{\mathrm{r}}(z) > -\infty$, for all z, then it is equal to V^{**}. This means that it can be computed with a dual approach, motivated by lemma 25, and that error estimates can be derived with lemma 26. Note that we need the existence of optimal solutions to the penalized problems.

References

[1] L. Andrieu, G. Cohen, and F. J. Vázquez-Abad. Gradient-based simulation optimization under probability constraints. *European J. Oper. Res.*, 212(2):345–351, 2011.

[2] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[3] O. Bokanowski, B. Bruder, S. Maroso, and Z. Zidani. Numerical approximation for a superreplication problem under gamma constraints. *SIAM. Num. Analysis.*, 47(3):2289–2320, 2009.

[4] B. Bouchard, R. Elie, and C. Imbert. Optimal control under stochastic target constraints. *SIAM Journal on Control and Optimization*, 48(5):3501–3531, 2010.

[5] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2010.

[6] P. Girardeau. *Résolution de grands problèmes en optimisation stochastique dynamique et synthèse de lois de commande.* PhD thesis, Université Paris-Est, 2010.

[7] G. Granato. *Optimisation de lois de gestion energétique des véhicules hybrides.* PhD thesis, Ecole Polytechnique, 2012.

[8] R. Henrion. On the connectedness of probabilistic constraint sets. *Journal of Optimization Theory and Applications*, 112(3):657–663, 2002.

[9] R. Henrion and C. Strugarek. Convexity of chance constraints with independent random variables. *Comput. Optim. Appl.*, 41(2):263–276, 2008.

[10] J. B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms: Part 2: Advanced Theory and Bundle Methods.* Convex Analysis and Minimization Algorithms. Springer, 1993.

[11] H. J. Kushner. Numerical methods for stochastic control problems in continuous time. *SIAM J. Control Optim.*, 28(5):999–1048, 1990.

[12] C. Lemaréchal. Lagrangian relaxation. In Michael Junger and Denis Naddef, editors, *Computational Combinatorial Optimization*, volume 2241 of *Lecture Notes in Computer Science*, pages 112–156. Springer Berlin Heidelberg, 2001.

[13] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Hochschultext / Universitext. Springer, 2003.

[14] L. Pfeiffer. *Analyse de sensibilité pour des problèmes de commande optimale. Commande optimale stochastique sous contrainte en probabilité.* PhD thesis, Ecole Polytechnique, 2013.

[15] R. T. Rockafellar. *Convex Analysis.* Princeton mathematical series. Princeton University Press, 1970.

[16] A. Shapiro, D. Dentcheva, and A. P. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory.* MPS-SIAM series on optimization. Society for Industrial and Applied Mathematics, 2009.

[17] N. Touzi. *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE.* Fields Institute Monographs. Springer, 2012.