



HAL
open science

Epiconvergence of relaxed stochastic optimization problem

Vincent Leclère

► **To cite this version:**

| Vincent Leclère. Epiconvergence of relaxed stochastic optimization problem. 2013. hal-00848275v2

HAL Id: hal-00848275

<https://hal.science/hal-00848275v2>

Preprint submitted on 26 Sep 2013 (v2), last revised 2 Nov 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

Epiconvergence of relaxed stochastic optimization problems

Vincent Leclère

Université Paris-Est, CERMICS (ENPC),
 6-8 avenue Blaise Pascal, Cité Descartes,
 F-77455 Marne-la-Vallée
vincent.leclere@cermics.enpc.fr

In this paper we consider approximations of dynamic stochastic optimization problems and their convergence. We focus on the relaxation of an almost sure constraint in a weaker conditional expectation constraint. We show an epiconvergence result relying on the Kudo convergence of σ -algebras and continuity of the objective and constraint operators. We also present some classical constraints and objective functions in stochastic optimization and give some conditions insuring their continuity. We are motivated by a decomposition algorithm that uses such a relaxation.

Key words:

MSC2000 subject classification:

OR/MS subject classification: Primary: ; secondary:

History:

1. Introduction Stochastic optimization problems often consist in minimizing a cost over a set of random variables. If the set of events is infinite, the minimization is done over an infinite dimensional space. Consequently there is a need for approximation. We are interested in the approximation of almost sure constraints, say $\theta(\mathbf{U}) = 0$ almost surely (a.s.), by a conditional expectation constraint like $\mathbb{E}(\theta(\mathbf{U}) \mid \mathcal{F}_n) \geq 0$ a.s.

Consider the following problem,

$$\min_{\mathbf{U} \in \mathcal{U}} J(\mathbf{U}), \tag{1a}$$

$$s.t. \quad \theta(\mathbf{U}) = 0 \quad \text{a.s.}, \tag{1b}$$

where the set of controls \mathcal{U} is a set of random variables over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If Ω is not finite, \mathcal{U} may be of infinite dimension. Moreover the constraint (1b) is a functional constraint that can roughly be seen as an infinity of constraints. For tractability purposes we consider approximations of this problem. In order to give theoretical results for the approximations of Problem (1) the right notion of convergence is epi-convergence. Indeed, under some additional technical conditions, the epi-convergence insures the convergence of both the optimal value and the optimal solutions.

One way of approximating Problem (1) consists in approximating the probability \mathbb{P} . Roughly speaking the Sample Average Approximation procedure consist in simulating a set of scenarios under the real probability \mathbb{P} . Then we solve Problem (1) under the empirical probability on the

set of simulated scenarios. In this literature (see [5], [8]) the authors are interested in problems where the controls are deterministic. However other epi-convergence results have been shown for more general spaces of controls, including spaces of random variables or random processes (see [24] and references therein, as well as [11], [13], [12]). More generally, the idea of discretizing or quantizing the set Ω , for example by use of finite scenario trees has been largely studied in the field of Stochastic Programming (see [22] for a thorough presentation).

Instead of approximating the probability space we propose a way to approximate constraints, especially almost sure constraints. The main idea is to replace a constraint by its conditional expectation with respect to (w.r.t.) a σ -algebra \mathcal{B} . This is in some sense an aggregation of constraints. This approximation appears when considering duality schemes for dynamic stochastic optimization problem.

More precisely, we relax the almost sure constraint (1b) by replacing it by its conditional expectation, i.e.

$$\mathbb{E}(\theta(\mathbf{U}) \mid \mathcal{B}) = 0. \quad (2)$$

If λ is an integrable optimal multiplier for Constraint (1b), then $\lambda_{\mathcal{B}} = \mathbb{E}(\lambda \mid \mathcal{B})$ is an optimal multiplier for Constraint (2). This leads to look for \mathcal{B} -measurable multiplier, which may authorize decomposition-coordination methods where the sub-problems are easily solvable. This is presented in the last part of our paper. In this case, the approximation can be seen as a decision rule approach in the dual, where we choose to restrict the multiplier in the class of \mathcal{B} -measurable random variables. Works using a decision rule approach on the dual problem are found in [10]. We refer the reader to [23] for more comments on the decision rule approach, especially for linear decision rules.

The paper is organized as follows. Section 2 presents the general form of the problem considered and its approximation. Section 3 shows, after a few recalls on convergence notions of random variables, functions and σ -algebras, conditions on the sequence of approximate problems guaranteeing its convergence toward the initial problem. The main assumptions are the Kudo's convergence of σ -algebra, and the continuity - as operators - of the constraint function Θ and objective function J . Section 4 gives some examples of continuous objective and constraint functions that represent usual stochastic optimization problems. Finally Section 5 presents a decomposition-coordination algorithm using this type of relaxation.

2. Problem statement We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a topological spaces of controls \mathcal{U} . Let \mathcal{V} be the spaces of random variables with value in a Banach \mathbb{V} with finite moment of order $p \in [1, \infty)$, denoted $\mathcal{V} = L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$.

We consider now a stochastic optimization problem

$$\min_{\mathbf{U} \in \mathcal{U}} J(\mathbf{U}), \quad (3a)$$

$$s.t. \quad \Theta(\mathbf{U}) \in -C, \quad (3b)$$

with J mapping \mathcal{U} into $\mathbb{R} \cup \{+\infty\}$, and Θ mapping \mathcal{U} into \mathcal{V} . We assume that $C \subset \mathcal{V}$ is a closed convex cone of \mathcal{V} , and that \mathbb{V} is a separable Banach space with separable dual (the fact that C is a cone is not essential for our results).

To give an example of cost operator, assume that $\mathcal{U} \subset L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{U})$, where \mathbb{U} is a Banach. The usual choice for the criterion is the expected cost $J(\mathbf{U}) := \mathbb{E}(j(\mathbf{U}))$, for a suitable cost function $j: \mathbb{U} \rightarrow \mathbb{R}$. Other choices could be risk measures (see [1] for example) like Conditional-Value-at-Risk (see [17] for a definition), worst-case or robust approaches. The constraint operator Θ cover various cases, for example

- almost sure constraint: $\Theta(\mathbf{U})(\omega) := \theta(\mathbf{U}(\omega))$, where θ maps \mathbb{U} into \mathbb{V} and $\theta(\mathbf{U}) \in -C$ is realized almost surely;

- measurability constraint: $\Theta(\mathbf{U}) := \mathbb{E}(\mathbf{U} \mid \mathcal{B}) - \mathbf{U}$, with $C = \{0\}$, expresses that \mathbf{U} is measurable with respect to the σ -algebra \mathcal{B} , that is, $\mathbb{E}(\mathbf{U} \mid \mathcal{B}) = \mathbf{U}$;
- risk constraint: $\Theta(\mathbf{U}) := \rho(\mathbf{U}) - a$, where ρ is a conditional risk measure, and C is the cone of positive random variables.

We introduce a stability assumption of the set C that will be made throughout this paper.

DEFINITION 1. We consider a sequence $(\mathcal{F}_n)_{n \in \mathbb{N}}$ of sub-fields of \mathcal{F} . The closed convex cone C is said to be *stable w.r.t. $(\mathcal{F}_n)_{n \in \mathbb{N}}$* , if for all $n \in \mathbb{N}$ we have

$$\forall \mathbf{V} \in C, \quad \mathbb{E}(\mathbf{V} \mid \mathcal{F}_n) \in C.$$

A first widely used example would be $C = \{0\}$, or more generally any deterministic closed convex cone, another example would be the set of almost surely positive random variables.

We now consider the following relaxation of Problem (3)

$$\min_{\mathbf{U} \in \mathcal{U}} J(\mathbf{U}), \tag{4a}$$

$$s.t. \quad \mathbb{E}(\Theta(\mathbf{U}) \mid \mathcal{F}_n) \in -C, \tag{4b}$$

where C is assumed to be stable w.r.t the sequence $(\mathcal{F}_n)_{n \in \mathbb{N}}$.

We denote the set of admissible controls of Problem (3)

$$\mathcal{U}^{ad} := \{\mathbf{U} \in \mathcal{U} \mid \Theta(\mathbf{U}) \in -C\}, \tag{5}$$

and the corresponding set of admissible controls of Problem (4)

$$\mathcal{U}_n^{ad} := \{\mathbf{U} \in \mathcal{U} \mid \mathbb{E}(\Theta(\mathbf{U}) \mid \mathcal{F}_n) \in -C\}. \tag{6}$$

Problems (3) and (4) can also be written¹ as

$$\min_{\mathbf{U} \in \mathcal{U}} \underbrace{J(\mathbf{U}) + \chi_{\mathcal{U}^{ad}}(\mathbf{U})}_{:= \bar{J}(\mathbf{U})}, \tag{7}$$

and

$$\min_{\mathbf{U} \in \mathcal{U}} \underbrace{J(\mathbf{U}) + \chi_{\mathcal{U}_n^{ad}}(\mathbf{U})}_{:= \bar{J}_n(\mathbf{U})}. \tag{8}$$

Note that we have $\mathcal{F}_n \subset \mathcal{F}$, and that C is stable w.r.t $(\mathcal{F}_n)_{n \in \mathbb{N}}$, thus $\mathcal{U}^{ad} \subset \mathcal{U}_n^{ad}$: Problem (4) is a relaxation of the original Problem (3) as it has the same objective function but a wider set of admissible controls.

Replacing an almost sure constraint by a conditional expectation constraint is similar to an aggregation of constraints. For example consider a finite set $\Omega = \{\omega_i\}_{i \in \llbracket 1, N \rrbracket}$ ², with a probability \mathbb{P} such that, for all $i \in \llbracket 1, N \rrbracket$, we have $\mathbb{P}(\omega_i) = p_i > 0$. Consider a partition $\mathcal{B} = \{B_l\}_{l \in \llbracket 1, |\mathcal{B}| \rrbracket}$ of Ω , and the σ -algebra $\mathcal{F}_{\mathcal{B}}$ generated by the partition \mathcal{B} . Assume that $C = \{0\}$, then the relaxation presented consists in replacing the constraint

$$\theta(\mathbf{U}) = 0 \quad \mathbb{P} - \text{a.s.}$$

¹ We use the notation χ_A for the characteristic function of A , that is the function such that $\chi_A(x) = 0$ if $x \in A$, and $\chi_A(x) = +\infty$ elsewhere.

² We denote by $\llbracket a, b \rrbracket$ the set of all integers between a and b .

which is equivalent to N constraints

$$\theta(\mathbf{U}(\omega_i)) = 0 \quad \forall i \in \llbracket 1, N \rrbracket,$$

by the collection of $|\mathcal{B}| \leq N$ (where $|\mathcal{B}|$ is the number of sets in the partition \mathcal{B}) constraints

$$\sum_{i \in B_l} p_i \theta(\mathbf{U}(\omega_i)) = 0 \quad \forall l \in \llbracket 1, |\mathcal{B}| \rrbracket.$$

3. Epiconvergence result In this section we show the epiconvergence of the sequence of approximated cost functions $(\tilde{J}_n)_{n \in \mathbb{N}}$ (defined in (8)) towards J (defined in (7)). In a first place we need some recalls on convergence of random variables, epiconvergence of functions and convergence of σ -algebras. Moreover a technical result is required.

3.1. Preliminaries Assume that $p \in [1, +\infty)$ and denote $q \in (1, +\infty]$ such that $1/q + 1/p = 1$. Recall that \mathbb{V} is a separable Banach space with separable dual \mathbb{V}^* .

Convergence of random variables A sequence $(\mathbf{X}_n)_{n \in \mathbb{N}}$ of $L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$ is said to converges strongly toward $\mathbf{X} \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$, and denoted $\mathbf{X}_n \rightarrow_{L^p} \mathbf{X}$ if

$$\lim_{n \rightarrow \infty} \mathbb{E}(\|\mathbf{X}_n - \mathbf{X}\|_{\mathbb{V}}^p) = 0.$$

A sequence $(\mathbf{X}_n)_{n \in \mathbb{N}}$ of $L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$ is said to weakly converges toward $\mathbf{X} \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$, and denoted $\mathbf{X}_n \rightharpoonup_{L^p} \mathbf{X}$ if

$$\forall \mathbf{X}' \in L^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V}^*), \quad \lim_{n \rightarrow \infty} \mathbb{E}(\langle \mathbf{X}_n - \mathbf{X}, \mathbf{X}' \rangle_{\mathbb{V}, \mathbb{V}^*}) = 0.$$

For more details we refer the reader to [19].

Epiconvergence of functions We first recall the definition of the Painlevé-Kuratowski convergence of sets. Let E be a topological space and consider a sequence $(A_n)_{n \in \mathbb{N}}$ of subsets of E . Then the inner limit of $(A_n)_{n \in \mathbb{N}}$ is the set of accumulation points of any sequence $(x_n)_{n \in \mathbb{N}}$ such that $x_n \in A_n$, i.e,

$$\underline{\lim}_n A_n = \{x \in E \mid \forall n \in \mathbb{N}, \quad x_n \in A_n, \quad \lim_{k \rightarrow \infty} x_n = x\}, \quad (9)$$

and the outer limit of $(A_n)_{n \in \mathbb{N}}$ is the set of accumulation points of any sub-sequence $(x_{n_k})_{k \in \mathbb{N}}$ of a sequence $(x_n)_{n \in \mathbb{N}}$ such that $x_n \in A_n$, i.e,

$$\overline{\lim}_n A_n = \{x \in E \mid \exists (n_k)_{k \in \mathbb{N}}, \quad \forall k \in \mathbb{N}, \quad x_{n_k} \in A_{n_k}, \quad \lim_{k \rightarrow \infty} x_{n_k} = x\}. \quad (10)$$

We say that $(A_n)_{n \in \mathbb{N}}$ converges toward A in the Painlevé-Kuratowski sense if

$$A = \overline{\lim}_n A_n = \underline{\lim}_n A_n.$$

A sequence $(J_n)_{n \in \mathbb{N}}$ of functions taking value into $\mathbb{R} \cup \{+\infty\}$ is said to epi-converge toward a function J if the sequence of epigraphs of J_n converges toward the epigraph of J , in the Painlevé-Kuratowski sense. For more details and properties of epi-convergence, see Rockafellar-Wets [16] in finite dimension, and Attouch [2] for infinite dimension.

Convergences of σ -algebras Let \mathcal{F} be a σ -algebra and $(\mathcal{F}_n)_{n \in \mathbb{N}}$ a sequence of sub-fields of \mathcal{F} . It is said that the sequence $(\mathcal{F}_n)_{n \in \mathbb{N}}$ Kudo-converges toward the σ -algebra \mathcal{F}_∞ , and denoted $\mathcal{F}_n \rightarrow \mathcal{F}_\infty$, if for each set $F \in \mathcal{F}$, $\left(\mathbb{E}(1_F \mid \mathcal{F}_n)\right)_{n \in \mathbb{N}}$ converges in probability toward $\mathbb{E}(1_F \mid \mathcal{F}_\infty)$.

In [9], Kudo shows that $\mathcal{F}_n \rightarrow \mathcal{F}_\infty$ if and only if for each integrable random variable \mathbf{X} , $\mathbb{E}(\mathbf{X} \mid \mathcal{F}_n)$ converges in L^1 toward $\mathbb{E}(\mathbf{X} \mid \mathcal{F}_\infty)$. In [14], Piccinini extends this result to the convergence in L^p in the strong or weak sense with the following lemma.

LEMMA 1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be a sequence of sub- σ -algebras of \mathcal{F} . The following statements are equivalent:

1. $\mathcal{F}_n \rightarrow \mathcal{F}_\infty$,
2. $\forall \mathbf{X} \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$, $\mathbb{E}(\mathbf{X} \mid \mathcal{F}_n) \rightarrow_{L^p} \mathbb{E}(\mathbf{X} \mid \mathcal{F}_\infty)$,
3. $\forall \mathbf{X} \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$, $\mathbb{E}(\mathbf{X} \mid \mathcal{F}_n) \rightharpoonup_{L^p} \mathbb{E}(\mathbf{X} \mid \mathcal{F}_\infty)$.

We have the following useful proposition where both the random variable and the σ -algebra are parametrized by n .

PROPOSITION 1. Assume that $\mathcal{F}_n \rightarrow \mathcal{F}_\infty$, and $\mathbf{X}_n \rightarrow_{L^p} \mathbf{X}$ (resp. $\mathbf{X}_n \rightharpoonup_{L^p} \mathbf{X}$) then $\mathbb{E}(\mathbf{X}_n \mid \mathcal{F}_n) \rightarrow_{L^p} \mathbb{E}(\mathbf{X} \mid \mathcal{F}_\infty)$ (resp. $\mathbb{E}(\mathbf{X}_n \mid \mathcal{F}_n) \rightharpoonup_{L^p} \mathbb{E}(\mathbf{X} \mid \mathcal{F}_\infty)$).

Proof The weak-limit case is detailed in [14]. We show the strong convergence case. If $\mathbf{X}_n \rightarrow_{L^p} \mathbf{X}$, then

$$\begin{aligned} \|\mathbb{E}(\mathbf{X}_n \mid \mathcal{F}_n) - \mathbb{E}(\mathbf{X} \mid \mathcal{F})\|_{L^p} &\leq \|\mathbb{E}(\mathbf{X}_n \mid \mathcal{F}_n) - \mathbb{E}(\mathbf{X} \mid \mathcal{F}_n)\|_{L^p} \\ &\quad + \|\mathbb{E}(\mathbf{X} \mid \mathcal{F}_n) - \mathbb{E}(\mathbf{X} \mid \mathcal{F})\|_{L^p} \end{aligned}$$

As the conditional expectation is a contraction operator, we have

$$\|\mathbb{E}(\mathbf{X}_n \mid \mathcal{F}_n) - \mathbb{E}(\mathbf{X} \mid \mathcal{F}_n)\|_{L^p} \leq \|\mathbf{X}_n - \mathbf{X}\|_{L^p} \rightarrow 0.$$

Moreover we have

$$\|\mathbb{E}(\mathbf{X} \mid \mathcal{F}_n) - \mathbb{E}(\mathbf{X} \mid \mathcal{F})\|_{L^p} \rightarrow 0$$

by Lemma 1, hence the result.

We finish by a few properties on the Kudo-convergence of σ -algebras (for more details we refer to [9] and [4]):

1. the topology associated with the Kudo-convergence is metrizable;
2. the set of σ -fields generated by the partitions of Ω is dense in the set of all σ -algebras;
3. if a sequence of random variables $(\mathbf{X}_n)_{n \in \mathbb{N}}$ converges in probability toward \mathbf{X} and for all $n \in \mathbb{N}$ we have $\sigma(\mathbf{X}_n) \subset \sigma(\mathbf{X})$, then we have the Kudo-convergence of $(\sigma(\mathbf{X}_n))_{n \in \mathbb{N}}$ toward $\sigma(\mathbf{X})$.

3.2. Main result Recall that \mathcal{U} is endowed with a topology τ , and that $\mathcal{V} = L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V})$, with $p \in [1, \infty)$.

THEOREM 2. Let \mathcal{V} be endowed with the strong or weak topology. Assume that C is stable w.r.t $(\mathcal{F}_n)_{n \in \mathbb{N}}$. If the two mappings Θ and J are continuous, and if $(\mathcal{F}_n)_{n \in \mathbb{N}}$ Kudo-converges toward \mathcal{F} , then $(\tilde{J}_n)_{n \in \mathbb{N}}$ (defined in (7)) epi-converges toward \tilde{J} (defined in (8)).

Note that $(\mathcal{F}_n)_{n \in \mathbb{N}}$ is not assumed to be a filtration, and that \mathcal{F}_n is not assumed to be finite.

Proof To prove the epi-convergence of $(\tilde{J}_n)_{n \in \mathbb{N}}$ toward \tilde{J} it is sufficient to show that \mathcal{U}_n^{ad} (defined in (6)) converges toward \mathcal{U}^{ad} (defined in (5)) in the Painlevé-Kuratowski sense. Indeed it implies the epi-convergence of $(\chi_{\mathcal{U}_n^{ad}})_{n \in \mathbb{N}}$ toward $\chi_{\mathcal{U}^{ad}}$, and adding a continuous function preserves the epi-convergence (Attouch [2, Th 2.15]).

By stability of C w.r.t. $(\mathcal{F}_n)_{n \in \mathbb{N}}$ we have that, for all $n \in \mathbb{N}$, $\mathcal{U}^{ad} \subset \mathcal{U}_n^{ad}$ and thus $\mathcal{U}^{ad} \subset \underline{\lim}_n \mathcal{U}_n^{ad}$ (for any $x \in \mathcal{U}^{ad}$ take the constant sequence equal to x).

We now show that $\mathcal{U}^{ad} \supset \overline{\lim}_n \mathcal{U}_n^{ad}$. Let \mathbf{U} be an element of $\overline{\lim}_n \mathcal{U}_n^{ad}$. By Definition (10), there is a sequence $(\mathbf{U}_{n_k})_{k \in \mathbb{N}}$ that τ -converges to \mathbf{U} , such that for all $k \in \mathbb{N}$, $\mathbb{E}(\Theta(\mathbf{U}_{n_k}) | \mathcal{F}_{n_k}) \in -C$. As Θ is continuous, we have $\Theta(\mathbf{U}_{n_k}) \rightarrow \Theta(\mathbf{U})$ strongly (resp. weakly) in L^p . Moreover we have that $\mathcal{F}_{n_k} \rightarrow \mathcal{F}$, and consequently by Corollary 1,

$$\mathbb{E}(\Theta(\mathbf{U}_{n_k}) | \mathcal{F}_{n_k}) \rightarrow_{L^p} \mathbb{E}(\Theta(\mathbf{U}) | \mathcal{F}) = \Theta(\mathbf{U}).$$

Thus $\Theta(\mathbf{U})$ is the limit of a sequence in $-C$. By closedness of C (weak and strong as C is convex³), we have that $\Theta(\mathbf{U}) \in -C$ and thus $\mathbf{U} \in \mathcal{U}^{ad}$.

The practical consequences for the convergence of the approximation (4) toward the original Problem 3 is given in the following Corollary.

COROLLARY 3. Assume that $\mathcal{F}_n \rightarrow \mathcal{F}$, and that J and Θ are continuous. Then the sequence of Problems (4) approximates Problem (3) in the following sense. If $(\mathbf{U}_n)_{n \in \mathbb{N}}$ is a sequence of control such that for all $n \in \mathbb{N}$,

$$\tilde{J}_n(\mathbf{U}_n) < \inf_{\mathbf{U} \in \mathcal{U}} \tilde{J}_n(\mathbf{U}) + \varepsilon_n, \text{ where } \lim_n \varepsilon_n = 0,$$

then, for every converging sub-sequence $(\mathbf{U}_{n_k})_{k \in \mathbb{N}}$, we have

$$\tilde{J}(\lim_k \mathbf{U}_{n_k}) = \min_{\mathbf{U} \in \mathcal{U}} \tilde{J}(\mathbf{U}) = \lim_k \tilde{J}_{n_k}(\mathbf{U}_{n_k}).$$

Moreover if $(\mathcal{F}_n)_{n \in \mathbb{N}}$ is a filtration, then the convergences are monotonous in the sense that the optimal value is non-decreasing in n .

Proof The convergence result is a direct application of Attouch [2, Th. 1.10, p. 27]. Monotonicity is given by the fact that, if $(\mathcal{F}_n)_{n \in \mathbb{N}}$ is a filtration, then for $n > m$ then $\mathcal{U}_n^{ad} \subset \mathcal{U}_m^{ad}$.

3.3. Dynamic Problem We extend Problem (3) into the following dynamic problem

$$\begin{aligned} \min_{\mathbf{U} \in \mathcal{U}} \quad & J(\mathbf{U}), \\ \text{s.t.} \quad & \Theta_t(\mathbf{U}_t) \in -C_t \quad \forall t \in \llbracket 1, T \rrbracket, \\ & \mathbf{U}_t \preceq \mathcal{F}_t, \end{aligned} \tag{11}$$

where $\mathbf{U}_t \preceq \mathcal{F}_t$ stands for “ \mathbf{U}_t is \mathcal{F}_t -measurable”. Here \mathbf{U} is a stochastic process of control $(\mathbf{U}_t)_{t \in \llbracket 1, T \rrbracket}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with value in \mathbb{U} . We have T constraints operators Θ_t taking values in $L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{V}_t)$, where $(\mathcal{F}_t)_{t \in \llbracket 1, T \rrbracket}$ is a sequence of σ -algebra. Note that $(\mathcal{F}_t)_{t \in \llbracket 1, T \rrbracket}$ is not necessarily a filtration. Then, for each $t \in \llbracket 1, T \rrbracket$ we define a sequence of approximating σ -algebra $(\mathcal{F}_{n,t})_{n \in \mathbb{N}}$. For all $t \in \llbracket 1, T \rrbracket$, C_t is a closed convex cone stable w.r.t $(\mathcal{F}_{n,t})_{n \in \mathbb{N}}$.

Finally we consider the sequence of approximated problem

$$\begin{aligned} \min \quad & J(\mathbf{U}), \\ \text{s.t.} \quad & \mathbb{E}(\Theta(\mathbf{U}_t) | \mathcal{F}_{n,t}) \in -C_t \quad \forall t \in \llbracket 1, T \rrbracket. \end{aligned} \tag{12}$$

Furthermore we denote

$$\mathcal{U}_t^{ad} := \{\mathbf{U}_t \in \mathcal{U}_t \mid \Theta(\mathbf{U}_t) \in -C_t\},$$

and

$$\mathcal{U}_{n,t}^{ad} := \{\mathbf{U}_t \in \mathcal{U}_t \mid \mathbb{E}(\Theta(\mathbf{U}_t) | \mathcal{F}_{n,t}) \in -C_t\}.$$

³ if C is non-convex we need it to be sequentially closed.

We define the set of admissible controls for the original problem

$$\mathcal{U}^{ad} = \mathcal{U}_0^{ad} \times \cdots \times \mathcal{U}_T^{ad},$$

and accordingly for the relaxed problem

$$\mathcal{U}_n^{ad} = \mathcal{U}_{n,0}^{ad} \times \cdots \times \mathcal{U}_{n,T}^{ad}.$$

In order to show the convergence of the approximation proposed here, we consider the functions

$$\tilde{J}(\mathbf{U}) = J(\mathbf{U}) + \chi_{\mathcal{U}^{ad}}(\mathbf{U}),$$

and

$$\tilde{J}_n(\mathbf{U}) = J(\mathbf{U}) + \chi_{\mathcal{U}_n^{ad}}(\mathbf{U}),$$

and show the epi-convergence of \tilde{J}_n to \tilde{J} .

THEOREM 4. If Θ and J are continuous, and if for all $t \in \llbracket 1, T \rrbracket$, $(\mathcal{F}_{t,n})_{n \in \mathbb{N}}$ Kudo-converges to \mathcal{F}_t , then $(\tilde{J}_n)_{n \in \mathbb{N}}$ epi-converges to \tilde{J} .

Proof The proof is deduced from the one of Theorem 2. By following the same steps we obtain the Painlevé-Kuratowski convergence of $\mathcal{U}_{n,t}^{ad}$ to \mathcal{U}_t^{ad} , and thus the convergence of their Cartesian products.

4. Examples The continuity of J and Θ as operators required in Theorem 2 is an abstract assumption. This section presents conditions for some classical constraint and objective functions to be representable by continuous operators. Before presenting those results we show a technical lemma that allows us to prove convergence for the topology of convergence in probability by considering sequences of random variables converging almost surely.

4.1. A technical Lemma

LEMMA 2. Let $\Theta : E \rightarrow F$, where $(E, \tau_{\mathbb{P}})$ is a space of random variables endowed with the topology of convergence in probability, and (F, τ) is a topological space. Assume that Θ is such that if $(\mathbf{U}_n)_{n \in \mathbb{N}}$ converges almost surely toward \mathbf{U} , then $\Theta(\mathbf{U}_n) \rightarrow_{\tau} \Theta(\mathbf{U})$. Then Θ is a continuous operator from $(E, \tau_{\mathbb{P}})$ into (F, τ) .

Proof We recall first a well known property (see for example [6, Th 2.3.3]). Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in a topological space. If from any sub-sequence $(x_{n_k})_{k \in \mathbb{N}}$ we can extract a sub-sub-sequence $(x_{\sigma(n_k)})_{k \in \mathbb{N}}$ converging to x^* , then $(x_n)_{n \in \mathbb{N}}$ converges to x^* . Indeed suppose that $(x_n)_{n \in \mathbb{N}}$ does not converges toward x^* . Then there exist an open set \mathcal{O} containing x^* and a sub-sequence $(x_{n_k})_{k \in \mathbb{N}}$ such that for all $k \in \mathbb{N}$, $x_{n_k} \notin \mathcal{O}$, and no sub-sub-sequence can converges to x^* , hence a contradiction.

Let $(\mathbf{U}_n)_{n \in \mathbb{N}}$ be a sequence converging in probability to \mathbf{U} . We consider the sequence $(\Theta(\mathbf{U}_n))_{n \in \mathbb{N}}$ in F . We choose a sub-sequence $(\Theta(\mathbf{U}_{n_k}))_{k \in \mathbb{N}}$. By assumption $(\mathbf{U}_n)_{n \in \mathbb{N}}$ converges in probability toward \mathbf{U} , thus we have $\mathbf{U}_{n_k} \rightarrow_{\mathbb{P}} \mathbf{U}$. Consequently there exist a sub-sub-sequence $\mathbf{U}_{\sigma(n_k)}$ converging almost surely to \mathbf{U} , and consequently $\Theta(\mathbf{U}_{\sigma(n_k)}) \rightarrow \Theta(\mathbf{U})$. Therefore Θ is sequentially continuous, and as the topology of convergence in probability is metrizable, Θ is continuous.

REMARK 5. This Lemma does not imply the equivalence between convergence almost sure and convergence in probability as you cannot endow \mathcal{U} with the “topology of almost sure convergence” as almost sure convergence is not generally induced by a topology.

However note that $(\mathbf{U}_n)_{n \in \mathbb{N}}$ converges in probability toward \mathbf{U} iff from any sub-sequence of $(\mathbf{U}_n)_{n \in \mathbb{N}}$ we can extract a further sub-sequence converging almost surely to \mathbf{U} (see [6, Th 2.3.2]).

4.2. Objective function Let \mathcal{U} be a space of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, with value in a Banach \mathbb{U} .

The most classical objective function is given as $J(\mathbf{U}) := \mathbb{E}(j(\mathbf{U}))$, where $j : \mathbb{U} \rightarrow \mathbb{R}$ is a measurable, bounded cost function. This objective function expresses a risk-neutral attitude; indeed a random cost with high variance or a deterministic cost with the same expectation are considered equivalent. Recently in order to capture risk-averse attitudes, coherent risk measures (as defined in [1]), or more generally convex risk measures (as defined in [7]), have been prominent in the literature.

Following [21], we call *convex risk measure* an operator $\rho : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ verifying

- Convexity: for all $\lambda \in [0, 1]$ and all $X, Y \in \mathcal{X}$, we have

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y);$$

- Monotonicity: for all $X, Y \in \mathcal{X}$ such that $X \leq Y$ we have $\rho(X) \leq \rho(Y)$;
 - Translation equivariance: for all constant $c \in \mathbb{R}$ and all $X \in \mathcal{X}$, we have $\rho(X + c) = \rho(X) + c$,
- where \mathcal{X} is a linear space of measurable functions. We focus on the case where $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$.

PROPOSITION 6. Let \mathcal{U} be a set of random variables endowed with the topology of convergence in probability, and $J(\mathbf{U}) := \rho(j(\mathbf{U}))$, where $j : \mathbb{U} \rightarrow \mathbb{R}$ is continuous and bounded, and ρ a convex risk measure. Then, $J : \mathcal{U} \rightarrow \mathbb{R}$ is continuous.

Proof Note that as j is bounded, $j(\mathbf{U}) \in \mathcal{X}$ for any $u \in \mathcal{U}$. Then we know that ([21]) there is a convex set of probabilities \mathcal{P} such that

$$\rho(\mathbf{X}) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\mathbb{Q}}(\mathbf{X}) - g(\mathbb{Q}),$$

where g is convex and weak*-lowersemicontinuous on the space of finite signed measures on (Ω, \mathcal{F}) . Moreover any probability in \mathcal{P} is absolutely continuous w.r.t \mathbb{P} .

Consider a sequence $(\mathbf{U}_n)_{n \in \mathbb{N}}$ of elements of \mathcal{U} converging in probability toward $\mathbf{U} \in \mathcal{U}$. Note that as j is bounded, we have $\rho(j(\mathbf{U})) < \infty$ by monotonicity of ρ . By definition of ρ , for all $\varepsilon > 0$ there is a probability $\mathbb{P}_\varepsilon \in \mathcal{P}$ such that

$$\mathbb{E}_{\mathbb{P}_\varepsilon}(j(\mathbf{U})) - g(\mathbb{P}_\varepsilon) \geq \rho(j(\mathbf{U})) - \varepsilon.$$

As \mathbb{P}_ε is absolutely continuous w.r.t \mathbb{P} , the convergence in probability under \mathbb{P} of $(\mathbf{U}_n)_{n \in \mathbb{N}}$ implies the convergence of probability under \mathbb{P}_ε and in turn the convergence in law under \mathbb{P}_ε . By definition of convergence in law we have that

$$\lim_n \mathbb{E}_{\mathbb{P}_\varepsilon}(j(\mathbf{U}_n)) - g(\mathbb{P}_\varepsilon) = \mathbb{E}_{\mathbb{P}_\varepsilon}(j(\mathbf{U})) - g(\mathbb{P}_\varepsilon).$$

Let η be a positive real, and set $\varepsilon = \eta/2$, and $N \in \mathbb{N}$ such that for all $n \geq N$,

$$|\mathbb{E}_{\mathbb{P}_\varepsilon}(j(\mathbf{U}_n)) - \mathbb{E}_{\mathbb{P}_\varepsilon}(j(\mathbf{U}))| \leq \frac{\eta}{2}. \tag{13}$$

Then, recalling that

$$\rho(j(\mathbf{U})) \geq \mathbb{E}_{\mathbb{P}_{\frac{\eta}{2}}}(j(\mathbf{U})) - g(\mathbb{P}_{\frac{\eta}{2}}) \geq \rho(j(\mathbf{U})) - \frac{\eta}{2}, \tag{14}$$

we have that for all $n \geq N$,

$$\begin{aligned} \rho(j(\mathbf{U}_n)) &= \sup_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\mathbb{Q}}(j(\mathbf{U}_n)) - g(\mathbb{Q}) \\ &\geq \mathbb{E}_{\mathbb{P}_{\frac{\eta}{2}}}(j(\mathbf{U}_n)) - g(\mathbb{P}_{\frac{\eta}{2}}) \\ &\geq \mathbb{E}_{\mathbb{P}_{\frac{\eta}{2}}}(j(\mathbf{U})) - g(\mathbb{P}_{\frac{\eta}{2}}) && \text{by (13),} \\ &\geq \rho(j(\mathbf{U})) - \eta && \text{by (14),} \end{aligned}$$

and thus

$$\rho(j(\mathbf{U})) + \frac{\eta}{2} \geq \rho(j(\mathbf{U}_n)) \geq \rho(j(\mathbf{U})) - \eta.$$

Thus $\lim_n \rho(j(\mathbf{U}_n)) = \rho(j(\mathbf{U}))$. Hence the continuity of J .

The assumptions of this Proposition can be relaxed in different ways.

In a first place, if the convex risk measure ρ is simply the expectation then we can simply endow \mathcal{U} with the topology of convergence in law. In this case the continuity assumption on j can also be relaxed. Indeed if $(\mathbf{U}_n)_{n \in \mathbb{N}}$ converges in law toward \mathbf{U} , and if the set K of points where j is continuous is such that $\mathbb{P}(\mathbf{U} \in K) = 1$, then $\mathbb{E}(j(\mathbf{U}_n))$ converges toward $\mathbb{E}(j(\mathbf{U}))$.

Otherwise assume that \mathcal{U} is a set of random variables endowed with the topology of convergence in probability, and that j continuous. Moreover if we can ensure that $j(\mathbf{U})$ is dominated by some integrable (for all probability of \mathcal{P}) random variable, then $J : \mathcal{U} \rightarrow \mathbb{R}$ is continuous. Indeed we consider a sequence $(\mathbf{U}_n)_{n \in \mathbb{N}}$ almost surely converging to \mathbf{U} . We modify the proof of Proposition 6 by using a dominated convergence theorem to show that $\lim_n \mathbb{E}_{\mathbb{P}_\varepsilon}(j(\mathbf{U}_n)) = \mathbb{E}_{\mathbb{P}_\varepsilon}(j(\mathbf{U}))$. Lemma 2 concludes the proof.

4.3. Constraint operator We present some usual constraints and how they can be represented by an operator Θ that is continuous and take values into \mathcal{V} .

4.3.1. Almost sure constraint From Lemma 2, we obtain a first important example of continuous constraints.

PROPOSITION 7. Suppose that \mathcal{U} is the set of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, with value in \mathbb{U} , endowed with the topology of convergence in probability. Assume that $\theta : \mathbb{U} \rightarrow \mathbb{V}$ is continuous and bounded. Then the operator $\Theta(\mathbf{U})(\omega) := \theta(\mathbf{U}(\omega))$ maps \mathcal{U} into \mathcal{V} and is continuous.

Proof The function θ being continuous, is also Borel measurable. Thus for all $\mathbf{U} \in \mathcal{U}$, for all Borel set $V \subset \mathbb{V}$, we have

$$(\Theta(\mathbf{U}))^{-1}(V) = \{\omega \in \Omega \mid \mathbf{U}(\omega) \in \theta^{-1}(V)\} \in \mathcal{B},$$

thus $\Theta(\mathbf{U})$ is \mathcal{F} -measurable. Boundedness of θ insure the existence of moment of all order of $\Theta(\mathbf{U})$. Thus Θ is well defined.

Suppose that $(\mathbf{U}_n)_{n \in \mathbb{N}}$ converges to \mathbf{U} almost surely. Then by boundedness of θ , we have that $\left(\|\theta(\mathbf{U}_n) - \theta(\mathbf{U})\|_{\mathbb{V}}^p \right)_{n \in \mathbb{N}}$ is bounded, and thus by dominated convergence theorem we have that

$$\lim_{n \rightarrow \infty} \theta(\mathbf{U}_n) = \theta(\mathbf{U}) \quad \text{in } L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{V}),$$

which is exactly

$$\lim_{n \rightarrow \infty} \Theta(\mathbf{U}_n) = \Theta(\mathbf{U}).$$

Consequently by Lemma 2 we have the continuity of Θ .

We note that boundedness of θ is only necessary in order to use the dominated convergence theorem. Thus an alternative set of assumptions is given in the following proposition.

PROPOSITION 8. Let \mathcal{B} be a sub-field of \mathcal{F} . If $\mathcal{U} = L^{p'}(\Omega, \mathcal{B}, \mathbb{P})$, with the topology of convergence in probability, and if θ is γ -Hölder, with $\gamma \leq p'/p$ then $\Theta(\mathbf{U})(\omega) := \theta(\mathbf{U}(\omega))$ is well defined and continuous as an operator mapping \mathcal{U} into \mathcal{V} .

Proof By definition a function θ mapping \mathbb{U} into \mathbb{V} is γ -Hölder if there exist a constant $C > 0$ such that for all u, u' in \mathbb{U} we have

$$\|\theta(u) - \theta(u')\|_{\mathbb{V}} \leq C \|u - u'\|_{\mathbb{U}}^\gamma,$$

in particular the 1-Hölder continuity is the Lipschitz continuity.

Following the previous proof we just have to check that the sequence $\left(\|\theta(\mathbf{U}_n) - \theta(\mathbf{U})\|_{\mathbb{V}}^p\right)_{n \in \mathbb{N}}$ is dominated by some integrable variable. The Hölder assumption implies

$$\left\|\theta(\mathbf{U}_{n_k}) - \theta(\mathbf{U})\right\|_{\mathbb{V}}^p \leq C^p \left\|\mathbf{U}_{n_k} - \mathbf{U}\right\|_{\mathbb{U}}^{p\gamma}.$$

And as $p\gamma \leq p'$, and \mathbf{U}_n and \mathbf{U} are elements of $L^{p'}(\Omega, \mathcal{F}, \mathbb{P})$, $\left\|\mathbf{U}_{n_k} - \mathbf{U}\right\|_{\mathbb{U}}^{p\gamma}$ is integrable.

4.3.2. Measurability constraint When considering a dynamic stochastic optimization problem, measurability constraints are used to represent the nonanticipativity constraints. They can be expressed by stating that a random variable and its conditional expectation are equal.

PROPOSITION 9. We set $\mathcal{U} = L^{p'}(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{U})$, with $p' \geq p$. Assume that

- either \mathcal{U} is equipped with the strong topology, and \mathcal{V} is equipped with the strong or weak topology,
- or \mathcal{U} and \mathcal{V} are equipped with the weak topology.

If \mathcal{B} is a sub-field of \mathcal{F} , then $\Theta(\mathbf{U}) := \mathbb{E}(\mathbf{U} \mid \mathcal{B}) - \mathbf{U}$, is well defined and continuous.

Proof In a first place note that as $p' \geq p$, and $\mathcal{F}' \subset \mathcal{F}$, $\mathcal{U} \subset \mathcal{V}$; and if $\mathbf{V} \in \mathcal{V}$ then $\mathbb{E}(\mathbf{V} \mid \mathcal{B}) \in \mathcal{V}$ as the conditional expectation is a contraction. Thus for all $\mathbf{U} \in \mathcal{U}$, we have $\Theta(\mathbf{U}) \in \mathcal{V}$.

Consider a sequence $(\mathbf{U}_n)_{n \in \mathbb{N}}$ of \mathcal{U} strongly converging in $L^{p'}$ toward $\mathbf{U} \in \mathcal{U}$. We have

$$\begin{aligned} \|\Theta(\mathbf{U}_n) - \Theta(\mathbf{U})\|_p &\leq \|\mathbf{U}_n - \mathbf{U}\|_p + \|\mathbb{E}(\mathbf{U}_n - \mathbf{U} \mid \mathcal{B})\|_p \\ &\leq 2\|\mathbf{U}_n - \mathbf{U}\|_p \\ &\leq 2\|\mathbf{U}_n - \mathbf{U}\|_{p'} \rightarrow 0. \end{aligned}$$

Thus the strong continuity of Θ is proven.

Now consider $(\mathbf{U}_n)_{n \in \mathbb{N}}$ converging weakly in $L^{p'}$ toward $\mathbf{U} \in \mathcal{U}$. We have, for all $\mathbf{Y} \in L^q$,

$$\begin{aligned} \mathbb{E}\left(\mathbb{E}(\mathbf{U}_n \mid \mathcal{B}) \cdot \mathbf{Y}\right) &= \mathbb{E}\left(\mathbf{U}_n \mathbb{E}(\mathbf{Y} \mid \mathcal{B})\right), \\ &\xrightarrow{n} \mathbb{E}\left(\mathbf{U} \mathbb{E}(\mathbf{Y} \mid \mathcal{B})\right), \\ &= \mathbb{E}\left(\mathbb{E}(\mathbf{U} \mid \mathcal{B}) \mathbf{Y}\right). \end{aligned}$$

Thus we have the weak convergence of the conditional expectation and therefore of Θ . Finally as the strong convergence imply the weak convergence we have the continuity from \mathcal{U} -strong into \mathcal{V} -weak.

Until now the topology of convergence in probability has been largely used. If we endow \mathcal{U} with the topology of convergence in probability in the previous proposition we will obtain continuity of Θ on a subset of \mathcal{U} . Indeed if a set of random variables \mathcal{U}^{ad} such that there exist a random variable in $L^{p'}(\Omega, \mathcal{F}, \mathbb{P})$ dominating every random variable in \mathcal{U}^{ad} , then a sequence converging almost surely will converge for the $L^{p'}$ norm and we can follow the previous proof to show the continuity of Θ on \mathcal{U}^{ad} .

4.3.3. Risk constraints Risk attitude can be expressed through the criterion or through constraints. We have seen that a risk measure can be chosen as objective function, we now show that conditional risk measure can be used as constraints.

Let ρ be a conditional risk mapping as defined in [20], and more precisely ρ maps \mathcal{U} into \mathcal{V} where $\mathcal{U} = L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{U})$ and $\mathcal{V} = L^p(\Omega, \mathcal{B}, \mathbb{P}; \mathbb{V})$, with $\mathcal{B} \subset \mathcal{F}$, and verifies the following properties

- Convexity: for all $\lambda \in \mathcal{U}$, $\lambda \in [0, 1]$ and all $X, Y \in \mathcal{V}$, we have

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda) \rho(Y);$$

- Monotonicity: for all $X, Y \in \mathcal{V}$ such that $X \leq Y$ we have $\rho(X) \leq \rho(Y)$;
- Translation equivariance: for all $c \in \mathcal{V}$ and all $X \in \mathcal{U}$, we have $\rho(X + c) = \rho(X) + c$.

PROPOSITION 10. Let \mathcal{U} be endowed with the topology of convergence in probability, and \mathcal{V} endowed with the strong topology. If ρ is a conditional risk mapping, θ is a continuous bounded cost function mapping \mathbb{U} into \mathbb{R} , and $a \in \mathcal{V}$, then $\Theta(\mathbf{U}) := \rho(\theta(\mathbf{U})) - a$ is continuous.

Proof Consider a sequence of random variables $(\mathbf{U}_n)_{n \in \mathbb{N}}$ converging in probability toward \mathbf{U}_∞ . Let $\pi : L^p(\Omega, \mathcal{B}, \mathbb{P}; \mathbb{U}) \rightarrow L^p(\Omega, \mathcal{B}, \mathbb{P}; \mathbb{U})$ be a selector of $\mathcal{V} = L^p(\Omega, \mathcal{B}, \mathbb{P}; \mathbb{U})$, i.e. for any $\mathbf{X} \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{U})$, $\pi(\mathbf{X}) \in \mathbf{X}$. For any $\omega \in \Omega$, any $\mathbf{X} \in L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{U})$ we define

$$\rho_\omega(\mathbf{U}) := \pi(\rho(\mathbf{U}))(\omega).$$

Note that for \mathbb{P} -almost all $\omega \in \Omega$, the function $\Theta_\omega(\mathbf{U}) := \rho_\omega(\theta(\mathbf{U}))$, satisfies the conditions of Proposition 6. Thus for \mathbb{P} -almost all $\omega \in \Omega$, $(\Theta_\omega(\mathbf{U}_n))_{n \in \mathbb{N}}$ converges toward $\Theta_\omega(\mathbf{U}_\infty)$. Thus we have shown that $(\Theta(\mathbf{U}_n))_{n \in \mathbb{N}}$ converges almost surely toward $\Theta(\mathbf{U}_\infty)$. By boundedness of θ and monotonicity of ρ we obtain the boundedness of $(\Theta(\mathbf{U}_n))_{n \in \mathbb{N}}$. Thus almost sure convergence and dominated convergence theorem insure that $(\Theta(\mathbf{U}_n))_{n \in \mathbb{N}}$ converges in L^p toward $\Theta(\mathbf{U}_\infty)$, hence the continuity of Θ .

Another widely used risk measure, even if it has some serious drawbacks, is the Value-at-Risk. If \mathbf{X} is a real random variable its value at risk of level α can be defined as $VaR_\alpha(\mathbf{X}) := \inf\{F_{\mathbf{X}}^{-1}(\alpha)\}$ where $F_{\mathbf{X}}(x) := \mathbb{P}(\mathbf{X} \leq x)$.

PROPOSITION 11. If $\theta : \mathbb{U} \rightarrow \mathbb{R}$ is continuous, and if \mathcal{U} is such that every $\mathbf{U} \in \mathcal{U}$ have a continuous distribution function, then $\Theta(\mathbf{U}) := VaR_\alpha(\theta(\mathbf{U}))$ is continuous if we have endowed \mathcal{U} with the topology of convergence in law, and a fortiori for the topology of convergence in probability.

Proof By definition of convergence in law, if $\mathbf{U}_n \rightarrow \mathbf{U}$ in law, then $(\theta(\mathbf{U}_n))_{n \in \mathbb{N}}$ converges in law toward $\theta(\mathbf{U})$ and we have, for all $x \in \mathbb{R}$, $F_{\theta(\mathbf{U}_n)}(x) \rightarrow F_{\theta(\mathbf{U})}(x)$. Thus $(\Theta(\mathbf{U}_n))_{n \in \mathbb{N}}$ converges almost surely toward $\Theta(\mathbf{U})$, and as $\Theta(\mathbf{U})$ is deterministic, Θ is continuous.

Note that in Proposition 11 the constraint function take deterministic values. Thus considering the conditional expectation of this constraint yields exactly the same constraint. However consider a constraint $\Theta_1 : \mathcal{U} \rightarrow \mathbb{R}$ of this form, and another constraint $\Theta_2 : \mathcal{U} \rightarrow \mathcal{V}$. Then if Θ_1 and Θ_2 are continuous, then so is the constraint $\Theta = (\Theta_1, \Theta_2) \rightarrow \mathbb{R} \times \mathcal{V}$. Thus we can apply Theorem 2 on the coupled constraint.

5. Dual Approximate Dynamic Programming In this section we say a few words about how the approximation of an almost sure constraint by a conditional expectation – as presented in section 3 – can be used.

5.1. Presentation of the problem We are interested in an electricity production problem with N power stations coupled by an equality constraint. At time step t , each power station i have an internal state \mathbf{X}_t^i , and is affected by a random exogenous noise \mathbf{W}_t^i . For each power station, and each time step t , we have a control $\mathbf{Q}_t^i \in \mathcal{Q}_{t,i}^{ad}$ that must be measurable with respect to \mathcal{F}_t where \mathcal{F}_t is the σ -algebra generated by all past noises: $\mathcal{F}_t = \sigma(\mathbf{W}_s^i)_{1 \leq i \leq n, 0 \leq s \leq t}$. Moreover there is a coupling constraint expressing that the total production must be equal to the demand. This constraint is represented as $\sum_{i=1}^N \theta_t^i(\mathbf{Q}_t^i) = 0$, where θ_t^i is a continuous bounded function from $\mathcal{Q}_{t,i}^{ad}$ into \mathbb{V} , for all $i \in \llbracket 1, n \rrbracket$. The cost to be minimized is a sum over time and power stations of all current local cost $L_t^i(\mathbf{X}_t^i, \mathbf{Q}_t^i, \mathbf{W}_t^i)$.

Finally the problem reads

$$\min_{\mathbf{X}, \mathbf{Q}} \mathbb{E} \left(\sum_{i=1}^N \sum_{t=0}^T L_t^i(\mathbf{X}_t^i, \mathbf{Q}_t^i, \mathbf{W}_t^i) \right) \quad (15a)$$

$$s.t. \quad \mathbf{X}_{t+1}^i = f_t^i(\mathbf{X}_t^i, \mathbf{Q}_t^i, \mathbf{W}_t^i) \quad \forall t, \quad \forall i, \quad (15b)$$

$$\mathbf{X}_0^i = x_0^i \quad \forall i, \quad (15c)$$

$$\mathbf{Q}_t^i \in \mathcal{Q}_{t,i}^{ad} \quad \forall t, \quad \forall i, \quad (15d)$$

$$\mathbf{Q}_t^i \preceq \mathcal{F}_t \quad \forall t, \quad \forall i, \quad (15e)$$

$$\sum_{i=1}^N \theta_t^i(\mathbf{Q}_t^i) = 0 \quad \forall t, \quad \forall i. \quad (15f)$$

For the sake of brevity, we denote by \mathcal{A} the set of random processes (\mathbf{X}, \mathbf{Q}) verifying constraints (15b), (15c) and (15d).

Let assume that all random variables are in L^2 spaces and dualize the coupling constraint (15f). We do not study here the relation between the primal and the following dual problem (see [15] and [18] for an alternative formulation involving duality between L^1 and its dual).

$$\max_{\lambda \in L^2} \min_{(\mathbf{X}, \mathbf{Q}) \in \mathcal{A}} \mathbb{E} \left(\sum_{i=1}^N \sum_{t=0}^T L_t^i(\mathbf{X}_t^i, \mathbf{Q}_t^i, \mathbf{W}_t^i) + \lambda_t \theta_t^i(\mathbf{Q}_t^i) \right) \quad (16)$$

$$s.t. \quad \mathbf{Q}_t^i \preceq \mathcal{F}_t \quad \forall t, \quad \forall i.$$

We solve this problem using a gradient-like algorithm on λ . Thus for a fixed $\lambda^{(k)}$ we have to solve N problems of smaller size than Problem (16).

$$\min_{(\mathbf{X}, \mathbf{U}) \in \mathcal{A}} \mathbb{E} \left(\sum_{t=0}^T L_t^i(\mathbf{X}_t^i, \mathbf{S}_t^i, \mathbf{W}_t^i) + \lambda_t^{(k)} \theta_t^i(\mathbf{Q}_t^i) \right) \quad (17)$$

$$s.t. \quad \mathbf{Q}_t^i \preceq \mathcal{F}_t \quad \forall t, \quad \forall i.$$

Note that the process $\lambda^{(k)}$ has no given dynamics but can be chosen to be adapted to the filtration $(\mathcal{F}_t)_{t=1, \dots, T}$. Consequently solving Problem (17) by Dynamic Programming is possible but numerically difficult as we need to keep all the past realizations of the noises in the state. In fact the so-called curse of dimensionality prevent us to solve numerically this problem.

Nevertheless it has recently been proposed in [3] to replace λ_t by $\mathbb{E}(\lambda_t \mid \mathbf{Y}_t)$, where \mathbf{Y}_t is a random variable measurable with respect to $(\mathbf{Y}_{t-1}, \mathbf{W}_t)$ instead of λ_t . This is similar to a decision rule approach for the dual as we are restraining the control to a certain class, the \mathbf{Y}_t -measurable λ in our case. Thus Problem (17) can be solved by Dynamic Programming with the augmented state $(\mathbf{X}_t^i, \mathbf{Y}_t)$. It has also been shown that, under some non-trivial conditions, using $\mathbb{E}(\lambda_t \mid \mathbf{Y}_t)$ instead of λ_t is equivalent to solving

$$\min_{(\mathbf{X}, \mathbf{Q}) \in \mathcal{A}} \mathbb{E} \left(\sum_{i=1}^N \sum_{t=0}^T L_t^i(\mathbf{X}_t^i, \mathbf{Q}_t^i, \mathbf{W}_t^i) \right) \quad (18a)$$

$$s.t. \quad \mathbf{Q}_t^i \preceq \mathcal{F}_t \quad \forall t, \quad \forall i, \quad (18b)$$

$$\mathbb{E} \left(\sum_{i=1}^N \theta_t^i(\mathbf{Q}_t^i) \mid \mathbf{Y}_t \right) = 0 \quad \forall t, \quad \forall i. \quad (18c)$$

Problem (18) is a relaxation of Problem (15) where the almost sure constraint (15f) is replaced by the constraint (18c). Now consider a sequence of information processes $(\mathbf{Y}^{(n)})_{n \in \mathbb{N}}$ each generating

a σ -algebra \mathcal{F}_n , and their associated relaxation (\mathcal{P}_n) (as specified in Problem 18) of Problem (15) (denoted (\mathcal{P})). Those problems correspond to Problems (11) and (12) with

$$J(\mathbf{U}) = \mathbb{E} \left(\sum_{i=1}^N \sum_{t=0}^T L_t^i(\mathbf{X}_t^i, \mathbf{Q}_t^i, \mathbf{W}_t^i) \right),$$

where $\mathbf{U} = (\mathbf{Q}^{(i)})_{i \in \llbracket 1, N \rrbracket}$ and \mathbf{X}_t^i follow the dynamic equation (15b). We also have

$$\Theta_t(\mathbf{U}_t) = \sum_{i=1}^N \theta_t^i(\mathbf{Q}_t^i)$$

and $C_t = \{0\}$.

Assume that for all $t \in \llbracket 1, T \rrbracket$, and all $i \in \llbracket 1, N \rrbracket$ the cost functions L_t^i and constraint function θ_t^i are continuous, and that $\mathcal{Q}_{t,i}^{ad}$ is a compact subset of an euclidian space. Moreover we assume that the noise variables \mathbf{W}_t^i are essentially bounded. Finally we endow the space of control processes with the topology of convergence in probability. Then by induction we have that the state processes and the control processes are essentially bounded, thus so is the cost $L_t^i(\mathbf{X}_t^i, \mathbf{U}_t^i, \mathbf{W}_t^i)$. Thus the cost function can be effectively replaced by bounded functions. Consequently Proposition 6 insures that J is continuous if \mathcal{U} is equipped with the topology of convergence in probability. Similarly Proposition 7 insures that Θ is continuous.

Thus Theorem 4 implies that our sequence of approximated problems (\mathcal{P}_n) converges toward the initial problem (\mathcal{P}) . More precisely assume that $(\mathbf{U}_n)_{n \in \mathbb{N}}$ is a sequence of ε_n -optimal solution of \mathcal{P}_n , i.e. \mathbf{U}_n verifying constraint (18c) and $J(\mathbf{U}_n) < \inf_{\mathcal{U} \in \mathcal{U}_n^{ad}} J + \varepsilon_n$, with $(\varepsilon_n)_{n \in \mathbb{N}}$ a sequence of positive real number converging to 0. Then we can extract a subsequence $(\mathbf{U}_{n_k})_{k \in \mathbb{N}}$ converging almost surely to an optimal solution of (\mathcal{P}) , and the limit of the approximated value of (\mathcal{P}_n) converges to the value of (\mathcal{P}) .

Conclusion In this paper we have used the standard theory of epi-convergence on a stochastic optimization framework combined with results on the Kudo convergence of σ -algebras. The abstract continuity conditions have been shown on some interesting frameworks under more classical conditions. Consequently most objective functions and constraints of a stochastic optimization problem, including tools such as convex conditional risk measures, are continuous as operators, and thus lead to a converging relaxation. Finally this relaxation naturally appears in a spatial decomposition-coordination approach in multistage stochastic optimization problem known as DADP.

Acknowledgments. The author thanks P.Carpentier, J.P.Chancelier and M.De Lara for their useful help and comments on this paper.

References

- [1] Artzner, P., F. Delbaen, J.M. Eber, D. Heath. 1999. Coherent measures of risk. *Mathematical Finance* **9** 203–228.
- [2] Attouch, H. 1984. *Variational convergence for functions and operators*, vol. 1. Pitman Advanced Pub. Program.
- [3] Barty, K., P. Carpentier, P. Girardeau. 2010. Decomposition of large-scale stochastic optimal control problems. *RAIRO. Recherche opérationnelle* **44**(3) 167–183.
- [4] Cotter, Kevin D. 1986. Similarity of information and behavior with a pointwise convergence topology. *Journal of mathematical economics* **15**(1) 25–38.
- [5] Dupacová, J., R.J.B. Wets. 1988. Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *The annals of statistics* **16**(4).

- [6] Durrett, R., R. Durrett. 2010. *Probability: theory and examples*. Cambridge Univ Pr.
- [7] Fllmer, H., A. Schied. 2002. Convex measures of risk and trading constraints. *Finance and Stochastics* **6** 429–447.
- [8] King, A.J., R.J.B. Wets. 1991. Epi-consistency of convex stochastic programs. *Stochastics and Stochastic Reports* **34**(1-2).
- [9] Kudo, H. 1974. A note on the strong convergence of σ -algebras. *The Annals of Probability* **2**(1) 76–83.
- [10] Kuhn, Daniel, Wolfram Wiesemann, Angelos Georghiou. 2011. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming* **130**(1).
- [11] Pennanen, Teemu. 2005. Epi-convergent discretizations of multistage stochastic programs. *Mathematics of Operations Research* **30**(1) 245–256.
- [12] Pennanen, Teemu. 2009. Epi-convergent discretizations of multistage stochastic programs via integration quadratures. *Mathematical Programming* **116**(1-2) 461–479.
- [13] Pennanen, Teemu, Matti Koivu. 2005. Epi-convergent discretizations of stochastic programs via integration quadratures. *Numerische mathematik* **100**(1) 141–163.
- [14] Piccinini, L. 1998. A new version of the multivalued fatou lemma. *Journal of Applied Analysis* **4**(2).
- [15] Rockafellar, R. T., R. J.B. Wets. 1978. The optimal recourse problem in discrete time: l^1 -multipliers for inequality constraints. *SIAM Journal on Control and Optimization* **16**(1).
- [16] Rockafellar, R Tyrrell, Roger J-B Wets. 1998. *Variational Analysis*, vol. 317. Springer Verlag.
- [17] Rockafellar, R.T., S. Uryasev. 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* **26**(7) 1443–1471.
- [18] Rockafellar, R.T., R.J.B. Wets. 1977. Measures as lagrange multipliers in multistage stochastic programming. *Journal of mathematical analysis and applications* **60**(2).
- [19] Rudin, W. 1991. *Functional analysis. International series in pure and applied mathematics*. McGraw-Hill, New York.
- [20] Ruszczyński, A., A. Shapiro. 2006. Conditional risk mappings. *Mathematics of Operations Research* 544–561.
- [21] Ruszczyński, A., A. Shapiro. 2006. Optimization of convex risk functions. *Mathematics of Operations Research* 433–452.
- [22] Shapiro, Alexander, Darinka Dentcheva, Andrzej Ruszczyński. 2009. *Lectures on stochastic programming: modeling and theory*, vol. 9. Society for Industrial and Applied Mathematics.
- [23] Shapiro, Alexander, Arkadi Nemirovski. 2005. On complexity of stochastic programming problems. *Continuous optimization*. Springer.
- [24] Zervos, M. 1999. On the epiconvergence of stochastic optimization problems. *Mathematics of Operations Research* **24**(2).