Epi-convergence of stochastic optimization problems involving both random variables and measurability constraints approximations

J.-Ph. Chancelier¹

¹System and Optimization Working Group, CERMICS

CERMICS — ENPC 6 et 8 avenue Blaise Pascal Cité Descartes - Champs sur Marne 77455 Marne la Vallée Cedex 2

http://cermics.enpc.fr

Epi-convergence of stochastic optimization problems involving both random variables and measurability constraints approximations

Jean-Philippe Chancelier¹, SOWG²

- ¹ Cermics, École Nationale des Ponts et Chaussées, 6 et 8 avenue Blaise Pascal, 77455, Marne la Vallée, Cedex, France, e-mail: jpc@cermics.enpc.fr
- ² Systems Optimization Working Group (SOWG, École Nationale des Ponts et Chaussées) currently composed with Kengy Barty, Pierre Carpentier, Guy Cohen, Anes Dallagi, Michel de Lara, Babacar Seck, Cyrille Strugarek and the author of this article, who wishes to thank his colleagues for their work and support.

The date of receipt and acceptance will be inserted by the editor

Abstract We give sufficient conditions for the Mosco convergence of

$$F_n(\mathbf{u}) \stackrel{\text{def}}{=} \mathbb{E}\left[j(\boldsymbol{\xi}_n, \mathbf{u})\right] + \delta_{\{\mathcal{B}_n \text{-measurable}\}}(\mathbf{u})$$

for $\mathbf{u} \in L^p(\mathbb{R}^m)$ $(1 \le p < \infty)$ when \mathcal{B}_n Kudo converges and $\boldsymbol{\xi}_n$ converges to $\boldsymbol{\xi}$ in $L^1(\mathbb{R}^l)$ or in probability.

1 Introduction

In stochastic programming, the programmer takes its decisions subject to partial observations. A typical example is in multistage decision problems where the programmer is constrained to use *non-anticipative* strategies i.e admissible strategies are constrained to be adapted to the filtration generated by observations. Thus, just a partial knowledge of the probability space is available and this leads to constraints on admissible controls which are named measurability constraints. Investigating the dependence of optimal strategies and optimal values towards measurability constraints is of deep importance for both theory and applications. Moreover, if we are concerned with numerical solution of stochastic problems, approximation of random variables involved in the problem are also to be considered. We are thus led to examine optimal strategies and optimal values when both random variable approximation and measurability constraints approximation are considered. More precisely, let (T, \mathcal{F}, μ) be a probability space, we want to approximate an optimization problem

$$\min_{\mathbf{n}\in L^p(T,\mathcal{B},\mu;U)} \mathbb{E}\left[j(\boldsymbol{\xi},\mathbf{u})\right] \tag{1}$$

with measurability constraints (\mathcal{B} is a sub field of \mathcal{F}). Using approximations of both the random variable $\boldsymbol{\xi}$ and the σ -field \mathcal{B} we are lead to solve a sequence of approximated problems :

υ

$$\min_{\mathbf{u}\in L^p(T,\mathcal{B}_n,\mu;U)} \mathbb{E}\left[j(\boldsymbol{\xi}_n,\mathbf{u})\right]$$

Using the sequence of functions $F_n(\mathbf{u}) \stackrel{\text{def}}{=} \mathbb{E}[j(\boldsymbol{\xi}_n, \mathbf{u})] + \delta_{\{\mathcal{B}_n\text{-measurable}\}}(\mathbf{u})$ (defined more precisely latter on by (5)) we can rewrite the sequence as :

$$\min_{\mathbf{u}\in L^p(U)}F_n(\mathbf{u})$$

We give, in this paper, sufficient conditions under which the sequence F_n Mosco converges to F_∞ on $L^p(U)$. Mosco convergence will give us epiconvergence of the sequence of optimization problems in both the strong and weak topologies. As it is well known [2], epi-convergence of a sequence of functions f_n mixed with compactness assumptions will permit to approximate the infimal value and the set of minimizers of the epi-limit of the sequence f_n [2, Theorem 1.10, Theorem 2.11]. Related results can be found in [1] and [25] where the dependence of the optimal strategies and optimal values according to σ -field variation are examined. An added constraint on the control u given by $\mathbb{E}[u] \in C$, where C is a convex set, is also considered in both papers. In the context of stochastic optimization where the measurability constraint is not present, approximation of optimization problem through random variables approximation have been widely studied [7,3,19,14,24]. Note also that mixed random variable approximations and measurability constraint approximations are considered in [12] and [16] with specific assumptions which link the random variable approximations to the measurability constraint approximations.

The paper is organized as follows, In section 2 and 3 we give the notations and preliminaries on epi and Mosco convergence. In section 4 we give a set of propositions and theorems on normal integrands, conditional expectations, integral functionals and convergence of σ -fields. Section 4 ends with Proposition 8 which relates the Mosco convergences of F_n to respectively the strong epigraphical upper limit of the integral function of the sequence of conditional expectation of $j(\boldsymbol{\xi}_n, u)$ with respect to \mathcal{B}_n and to the weak epigraphical lower limit of the integral function of the sequence inequality for the strong epigraphical upper limit to the same inequality for the conditional expectation of integrands. This inequality is derived from [6] where slice-convergence of integral functional are related to slice-convergence of integrands. Subsection 4.2 is devoted to inequality for the weak epigraphical lower limit. Two results are given, in the first one (Theorem 3), more stringent condition on the $\boldsymbol{\xi}_n$ convergence are required but the convergence can be extended to more general functional spaces. The second result (Proposition 10) is based on a theorem of Ioffe [10]. Previous propositions are gathered in the last subsection to obtain a Mosco convergence theorem.

2 Definitions and notations

In the sequel, except explicitly stated, $U \stackrel{\text{def}}{=} \mathbb{R}^m$ and we will follow the notations of [6]. We denote by $\mathcal{P}(U)$ the space of all subsets of U and note $\mathbb{B}(U)$ for the Borel σ -field of U.

Let T be an abstract space. A multifunction or set-valued mapping F denoted by $F: T \rightrightarrows U$ is a map from T to $\mathcal{P}(U)$. The domain of F is the subset of T defined by dom $(F) \stackrel{\text{def}}{=} \{t \in T/F(t) \neq \emptyset\}.$

A selection of a multifunction F is a function $s: T \to U$ such that for all $t \in \text{dom}(F)$, $s(t) \in F(t)$ and we denote by Sel(F) the set of all the selections of F.

Suppose now that (T, \mathcal{F}) is a measurable space. A multifunction is said to be *measurable* if for every open set O in U we have $F^-O \in \mathcal{F}$, where $F^-O \stackrel{\text{def}}{=} \{t \in T/F(t) \cap O \neq \emptyset\}$. A *measurable selection* of F is a selection of F that is $(\mathcal{F}, \mathbb{B}(U))$ -measurable

Every map $f: T \times U \mapsto \overline{\mathbb{R}}$ such that f(t, u) is measurable with respect to t for every $u \in U$ is called an *integrand* [21, p. 661].

The *epigraphical multifunction* associated to an integrand f is

$$S_f: t \mapsto \operatorname{epi} f(t, .) = \{(u, \alpha) \in U \times \mathbb{R} \mid f(t, u) \le \alpha\}$$

and its domain mapping D_f is

$$D_f: t \mapsto \operatorname{dom} f(t, .) = \{ u \in U \,|\, f(t, u) < +\infty \} \,.$$

If the epigraphical multifunction S_f is measurable and closed-valued then f is said to be a normal integrand (we will also use \mathcal{F} -normal integrand when the σ -field used for measurability is to be specified) [21, Definition 14.27, p. 661]. Moreover if for almost every $t \in T$ the function f(t, .) satisfies a property P, then the integrand f is said to satisfy P. For example, f is a convex integrand if it is an integrand and f(t, .) is convex for every $t \in T$.

Let (T, \mathcal{F}, μ) be a complete probability space, that is \mathcal{F} is a complete σ -field. We write $L^0(U)$ for the space of $(\mathcal{F}, \mathbb{B}(U))$ -measurable functions on T with values in U. We write $L^p(U)$ ($1 \leq p \leq +\infty$) for the space of functions f such that $t \mapsto ||f(t)||_U$ belongs to $L^p(T, \mathcal{F}, \mu; \mathbb{R})$. The selections of F which are in $L^p(U)$ are denoted by $S^p(F) \stackrel{\text{def}}{=} L^p(U) \cap \text{Sel}(F)$. We will use bold faced fonts for elements of L^p spaces (*i.e* $\mathbf{u} \in L^p(U)$) and standard fonts for their values or for elements of U (*i.e* $u(t) \in U, u \in U$). The *integral function* associated to a normal integrand is the functional defined on $L^0(U)$ by:

$$I(f)(\mathbf{u}) \stackrel{\text{def}}{=} \int f(t, u(t))\mu(dt)$$
(2)

With the following traditional convention [21, Proposition 14.56] : if $\max(f(t, u(t)), 0) \in L^1(\mathbb{R})$ then $I(f)(\mathbf{u})$ is defined by (2) else $I(f)(\mathbf{u}) = +\infty$.

3 Epiconvergence

Let (Y, ρ) be an abstract topological space and $f_n : Y \to \overline{\mathbb{R}}$ a sequence of functions. We denote by V(u) the family of neighborhoods of $u \in Y$ relatively to topology ρ . The following functions are said to be, respectively, the ρ -epigraphical (or ρ -epi) lower limit and the ρ -epigraphical (or ρ -epi) upper limit of the sequence (f_n) :

$$(\rho - \operatorname{li}_e f_n)(u) \stackrel{\text{def}}{=} \sup_{\mathcal{V} \in V(u)} \liminf_{n} \inf_{\overline{u} \in \mathcal{V}} f_n(\overline{u})$$
$$(\rho - \operatorname{ls}_e f_n)(u) \stackrel{\text{def}}{=} \sup_{\mathcal{V} \in V(u)} \limsup_{n} \inf_{\overline{u} \in \mathcal{V}} f_n(\overline{u})$$

When these two functions are equal, the common value is called the ρ -epigraphical (or ρ -epi) limit of (f_n) , it is denoted by ρ -lim_e f_n , and the sequence (f_n) is said to be epi-convergent. Moreover, let us notice [2, Theorem 1.13] that if Y is a metric space for a topology denoted here by ρ , we have

$$(\rho-\mathrm{li}_e f_n)(u) = \min\left\{\liminf_n f_n(u_n) \,|\, u_n \text{ such that } u = \rho-\mathrm{lim}\,u_n\right\}.$$
 (3)

$$(\rho\text{-ls}_e f_n)(u) = \min\left\{\limsup_n f_n(u_n) \,|\, u_n \text{ such that } u = \rho\text{-lim}\,u_n\right\}.$$
 (4)

Let σ be another topology on Y. The sequence (f_n) is said to be Mosco convergent to f with respect to ρ and σ , and we write $f = M(\rho, \sigma)$ -lim f_n if $f = \rho$ -lim $f_n = \sigma$ -lim f_n .

If Y is a metric space with two topologies $\rho \leq \sigma$, using (3) and (4) we have [2, Proposition 1.14] that $f = M(\rho, \sigma)$ -lim f_n (with $\rho \leq \sigma$) if and only if conditions below are satisfied:

- (i) for every $u \in Y$, there exists a sequence (u_n) in Y with $u = \sigma$ -lim u_n such that $\limsup f_n(u_n) \leq f(u)$
- (ii) for every $u \in Y$, for any sequence (u_n) such that $u = \rho$ -lim u_n , then $\liminf f_n(u_n) \ge f(u)$.

(i) is equivalent to σ -ls_e $f_n(u) \leq f(u)$ and (ii) is equivalent to $f(u) \leq \rho$ -ls_e $f_n(u)$. We will use this criteria when σ and ρ are respectively the strong (s-) and weak (w-) topologies on L^p spaces.

4 Propositions

Let $(f_n) : T \times U \mapsto \overline{\mathbb{R}}$ a sequence of integrands and (\mathcal{B}_n) a sequence of sub σ -fields of \mathcal{F} defined for $n \in \overline{\mathbb{N}}$, we define a sequence of functionals $(F_n(\mathbf{u})) : L^0(U) \mapsto \overline{\mathbb{R}}$ for $n \in \overline{\mathbb{N}}$ as follows :

$$F_n(\mathbf{u}) \stackrel{\text{\tiny def}}{=} \begin{cases} \int f_n(t, u(t)) \mu(dt) & \text{if } \mathbf{u} \text{ is } \mathcal{B}_n \text{-measurable} \\ +\infty \text{ otherwise} \end{cases}$$
(5)

For a sequence (\mathcal{B}_n) of sub σ -fields of \mathcal{F} , when the f_n are normal integrands, we consider the sequence of integrands $g_n : T \times U \mapsto \overline{\mathbb{R}}$ for $n \in \overline{\mathbb{N}}$ defined as follows :

$$g_n(t,u) \stackrel{\text{\tiny def}}{=} \mathbb{E}\left[f_n(t,u) \,|\, \mathcal{B}_n\right]. \tag{6}$$

The next theorem shows that the (g_n) sequence is well defined when the f_n are \mathcal{F} -quasi integrable and that g_n is a \mathcal{B}_n -normal integrand.

Definition 1 An integrand f is said to be \mathcal{F} -quasi integrable if there exists a sequence $\alpha_m \in L^1(T; \mathbb{R})$ such that

$$\forall m \in \mathbb{N}, \forall t \in T, \inf_{u \in mB} f(t, u) \ge \alpha_m(t)$$

where B is the unit ball of U.

Theorem 1 (L. Thibault). Let \mathcal{B} be a sub σ -field of \mathcal{F} and $f: T \times U \to \overline{\mathbb{R}}$ be a \mathcal{F} -normal integrand which is \mathcal{F} -quasi integrable Then, there exists a $\mathcal{B} \times \mathbb{B}(U)$ -measurable mapping $g_{\mathcal{B}}: T \times U \to \overline{\mathbb{R}}$ such that

(i) for all $B \in \mathcal{B}$ and all $\mathbf{u} \in L^0(U)$ and \mathcal{B} -measurable

$$\int_B f(t, u(t))\mu(dt) = \int_B g_{\mathcal{B}}(t, u(t))\mu(dt).$$

- (ii) $g_{\mathcal{B}}$ is unique up to indistinguishability and is a \mathcal{B} -normal integrand. For all $(t, u) \in T \times mB$, $g_{\mathcal{B}}(t, u) \geq \mathbb{E} [\alpha_m | \mathcal{B}]$.
- (iii) Moreover, if f(t,.) is convex for almost all t then g(t,.) is also convex convex for almost all t.

Proof: The proof of (i) and (ii) can be found in [23, Proposition 12] and the proof of (iii) in [23, Proposition 15]. For (i) and (ii), a weaker assumption is given in [5] and a survey can be found in [22].

Definition 2 Denote by U^* the dual topological space of U, and by $\langle \cdot, \cdot \rangle$ the pairing product. Let $f: T \times U \mapsto \overline{\mathbb{R}}$ be a \mathcal{F} -normal integrand. The conjugate of f, is the mapping $f^*: T \times U^* \mapsto \overline{\mathbb{R}}$ defined by

$$\forall (t, u^{\star}) \in T \times U^{\star}, \ f^{\star}(t, u^{\star}) = \sup_{u \in U} (\langle u, u^{\star} \rangle - f(t, u)).$$
(7)

Proposition 1 Let $f : T \times U \mapsto \overline{\mathbb{R}}$ be a \mathcal{F} -normal integrand. Then, its conjugate $f^* : T \times U^* \mapsto \overline{\mathbb{R}}$ is a \mathcal{F} -normal integrand.

Proof: Proposition 1 can be found in [21, Theorem 14.50] when $U \stackrel{\text{def}}{=} \mathbb{R}^m$ and when U is a separable Banach space [20,4] or [8,9] (whithout completeness hypothesis on μ).

Proposition 2 (L. Thibault) Suppose that there exists $\mathbf{u}^* \in L^0(U^*)$ such that $\|\mathbf{u}\|$ and $f(t, \mathbf{u}(t))$ are \mathcal{F} -integrable then f is \mathcal{F} -quasi integrable

Proof:[23, Proposition 4]

Proposition 3 Let \mathcal{B} a sub σ -field of \mathcal{F} and g(t, u) a convex normal integrand which is \mathcal{B} measurable. For $\mathbf{u} \in L^p(U)$ $(p \ge 1)$ we have

$$I(g)(\mathbf{u}) \ge I(g)(\mathbb{E}\left[\mathbf{u} \mid \mathcal{B}\right])$$

Proof: we have $\mathbf{u} \in L^1(U)$ and the result follows from [1, Proposition 3.2]

Proposition 4 Let f_n a \mathcal{F} -normal integrand fulfilling conditions of Theorem 1. We consider the integral function sequence $(I(f_n))$. For $\mathbf{u} \in L^0(U)$ if \mathbf{u} is \mathcal{B}_n -measurable then $I(f_n)(\mathbf{u}) = I(g_n)(\mathbf{u}) = F_n(\mathbf{u})$, where g_n and F_n are defined by (6) and (5).

Proof: If **u** is \mathcal{B}_n -measurable then $I(f_n)(\mathbf{u}) = F_n(\mathbf{u})$ by definition of F_n and using Theorem 1 we obtain $I(f_n)(\mathbf{u}) = I(g_n)(\mathbf{u})$.

Let $(\mathcal{B}_n)_n$ be a sequence of sub- σ -fields. H. Kudo introduced in [13] the notions of upper and lower limits of the sequence (\mathcal{B}_n) that we recall here :

Definition 3 A σ -field is called the upper limit (resp. lower limit) of $(\mathcal{B}_n)_n$, and denoted by \mathcal{B}^{\sharp} (resp. \mathcal{B}^{\flat}) if it is the minimal (resp. maximal) sub- σ -field among sub- σ -fields \mathcal{B} of \mathcal{F} such that for every $f \in L^{\infty}_{\mathbb{R}}(\mathcal{F})$

 $\limsup \|\mathbb{E}[f \mid \mathcal{B}_n]\|_1 \le \|\mathbb{E}[f \mid \mathcal{B}]\|_1$

(resp. $\liminf \|\mathbb{E}[f | \mathcal{B}_n]\|_1 \ge \|\mathbb{E}[f | \mathcal{B}]\|_1$)

When \mathcal{B}^{\sharp} is equal to \mathcal{B}^{\flat} , their common value is denoted by \mathcal{B}_{∞} and the sequence is said to Kudo converge $(\mathcal{B}_n \to \mathcal{B}_{\infty})$. When $1 \leq p < \infty$, $\mathcal{B}_n \to \mathcal{B}_{\infty}$ if and only if $\mathbb{E}[f | \mathcal{B}_n]$ converges to $\mathbb{E}[f | \mathcal{B}_{\infty}]$ for every $f \in L^p_{\mathbb{R}}(\mathcal{F})$, where L^p is endowed with the strong or weak topology [18, Theorem 2.2] (Note that the case $p = \infty$ cannot be treated with strong topology [17])

Proposition 5 Let $(f_n)_{n\in\overline{\mathbb{N}}}$ a sequence of convex normal integrands and suppose that the $(\mathcal{B}_k)_{k\in\mathbb{N}}$ sequence converges in the Kudo sense to \mathcal{B}_{∞} . For $\mathbf{u} \in L^p(U)$ we consider the sequence of integral functions $(I(f_n))$. If $I(f_{\infty})(\mathbf{u}) \leq \tau - li_e I(f_n)(\mathbf{u})$ then $F_{\infty}(\mathbf{u}) \leq \tau - li_e F_n(\mathbf{u})$ where τ can be the strong or weak topology in $L^p(U)$ with $1 \leq p < +\infty$.

Proof: Let $\mathbf{u}_n \ \tau$ -converging to \mathbf{u} and $\nu(n)$ a sub-sequence such that $\lim_{n\to\infty} F_{\nu(n)}(\mathbf{u}_{\nu(n)}) \leq \tau$ -li_e $F_n(\mathbf{u}) + \epsilon$. If there is a subsequence $\beta(n)$ of $\nu(n)$ for which $\mathbf{u}_{\beta(n)}$ is not $\mathcal{B}_{\beta(n)}$ -measurable the result is obvious. Thus we may assume that \mathbf{u}_n is \mathcal{B}_n -measurable and consequently $F_n(\mathbf{u}_n) = I(f_n)(\mathbf{u}_n)$ and $I(f_{\infty})(\mathbf{u}) \leq \tau$ -li_e $F_n(\mathbf{u}) + \epsilon$. Using [18, theorem 2.3] $\mathbb{E}[\mathbf{u}_n | \mathcal{B}_n] \tau$ -converges to $\mathbb{E}[\mathbf{u} | \mathcal{B}_{\infty}]$ the result being true for the strong or weak topology of L^p when $1 \leq p < +\infty$. Since \mathbf{u}_n is \mathcal{B}_n -measurable we have $\mathbf{u}_n = \mathbb{E}[\mathbf{u}_n | \mathcal{B}_n]$ and we obtain that \mathbf{u} is \mathcal{B}_{∞} -measurable and thus $I(f_{\infty})(\mathbf{u}) = F_{\infty}(\mathbf{u})$. \Box

Proposition 6 Let $(f_n)_{n\in\overline{\mathbb{N}}}$ a sequence of convex normal integrands and suppose that the $(\mathcal{B}_k)_{k\in\mathbb{N}}$ sequence converges in the Kudo sense to \mathcal{B}_{∞} . For $\mathbf{u} \in L^p(U)$ If $F_{\infty}(\mathbf{u}) \leq \tau$ -li_e $F_n(\mathbf{u})$ and \mathbf{u} is \mathcal{B}_{∞} -measurable then $I(g_{\infty})(\mathbf{u}) \leq \tau$ -li_e $I(g_n)(\mathbf{u})$ where τ can be the strong or weak topology in $L^p(U)$ with $1 \leq p < +\infty$.

Proof: Let $\mathbf{u}_n \tau$ -converging to \mathbf{u} and $\mathbf{w}_n \stackrel{\text{def}}{=} \mathbb{E}[\mathbf{u}_n | \mathcal{B}_n]$. \mathbf{w}_n converges to $\mathbb{E}[\mathbf{u} | \mathcal{B}_\infty]$ [18, theorem 2.3] and we thus have :

$$F_{\infty}(\mathbf{u}) \leq \liminf_{n \mapsto \infty} F_n(\mathbf{w}_n)$$

Moreover, \mathbf{w}_n is \mathcal{B}_n -measurable and by Proposition 4 $F_n(\mathbf{w}_n) = I(g_n)(\mathbf{w}_n)$. Since \mathbf{u} is assumed to be \mathcal{B}_∞ -measurable we also have $F_\infty(\mathbf{u}) = I(g_\infty)(\mathbf{u})$. Thus $I(g_\infty)(\mathbf{u}) \leq \liminf_{n \mapsto \infty} I(g_n)(\mathbf{w}_n)$. Using Proposition 3 we have

$$I(g_n)(\mathbf{w}_n) \le I(g_n)(\mathbf{u}_n)$$

which combined with the previous inequality end the proof.

Proposition 7 Let $(f_n)_{n\in\overline{\mathbb{N}}}$ a sequence of convex normal integrands and suppose that the $(\mathcal{B}_n)_{n\in\mathbb{N}}$ sequence converges in the Kudo sense to \mathcal{B}_{∞} . If τ -ls_e $I(g_n)(\mathbf{u}) \leq I(g_{\infty})(\mathbf{u})$ then τ -ls_e $F_n(\mathbf{u}) \leq F_{\infty}(\mathbf{u})$

Proof: The only case to consider is when \mathbf{u} is \mathcal{B}_{∞} -measurable. Let \mathbf{u}_n a $L^p(U)$ sequence such that $\limsup_{n \to \infty} I(g_n)(\mathbf{u}) \leq I(g_{\infty})(\mathbf{u})$. Since \mathbf{u} is \mathcal{B}_{∞} -measurable we have by Proposition 4 $I(g_{\infty})(\mathbf{u}) = I(f_{\infty})(\mathbf{u}) = F_{\infty}(\mathbf{u})$. Let $\mathbf{w}_n \stackrel{\text{def}}{=} \mathbb{E}[\mathbf{u}_n | \mathcal{B}_n]$, the sequence \mathbf{w}_n converges to $\mathbb{E}[\mathbf{u} | \mathcal{B}_{\infty}] = \mathbf{u}$ in $L^p(U)$. using again Proposition 4 and Proposition 3 we have $F_n(\mathbf{w}_n) = I(g_n)(\mathbf{w}_n) \leq I(g_n)(\mathbf{u}_n)$ and thus :

$$\limsup_{n \to \infty} F_n(\mathbf{w}_n) \le \limsup_{n \to \infty} I(g_n)(\mathbf{u}_n) \le I(g_\infty)(\mathbf{u}) = F_\infty(\mathbf{u})$$
(8)

Proposition 8 Let f_n a sequence of of convex normal integrands and assume the same hypothesis as in Propositions 5 and 7. If :

$$\mathbf{s} - ls_e I(g_n)(\mathbf{u}) \le I(g)(\mathbf{u}) \text{ and } I(f)(\mathbf{u}) \le \mathbf{w} - li_e I(f_n)(\mathbf{u})$$
(9)

then the sequence F_n Mosco converges to F in $L^p(U)$.

Proof: The proof easily follows from Proposition 5 and 7. \Box

4.1 Inequality for the s-epigraphical upper limit

We give results here to obtain $\mathbf{s}\text{-ls}_e I(g_n)(\mathbf{u}) \leq I(g)(\mathbf{u})$ where (g_n) is given by (6). The proof is based on [6, theorem 4.1] in which U can be supposed to be a separable Banach case, in the special case where $U = \mathbb{R}^m$ similar results were proved in [11].

Theorem 2 (Couvreux [6, theorem 4.1]). Let U be a separable Banach space, $\gamma_n : T \times U \to (-\infty, +\infty]$ $(n \ge 1)$ a sequence of normal proper integrands, $\gamma : T \times U \to (-\infty, +\infty]$ a proper integrand, p, q with $1 \le p < +\infty$, and $p^{-1} + q^{-1} = 1$, satisfying the following assumptions:

- (a) for almost every $t \in T$ and each $u \in dom \gamma(t, .), \gamma(t, u) \ge \mathbf{s} \cdot ls_e \gamma_n(t, u);$
- (b) there exists a sequence (\mathbf{u}_n) in $L^p(U)$ and functions k and k_0 in $L^p(\mathbb{R})$, such that for each n, $||u_n(t)|| \le k(t)$ and $\gamma_n(t, u_n(t)) \le k_0(t)$ a.s.;
- (c) for almost every $t \in T$, each $u \in U$ and each $n \geq 1$, $\gamma_n(t,u) \geq -h(t)||u|| h_0(t)$, where h and h_0 belong to $L^q(\mathbb{R})$ with h(t) > 0 a.s.

Then for every function \mathbf{u} in $L^p(U)$, $I(\gamma)(\mathbf{u}) \geq \mathbf{s} \cdot ls_e I(\gamma_n)(\mathbf{u})$.

Proposition 9 For a sequence $\boldsymbol{\xi}_n$ of random variable we define a sequence (f_n) by $f_n(t, u) \stackrel{\text{def}}{=} j(\boldsymbol{\xi}_n(t), u)$ and then the associated sequences g_n by (6). Suppose that

- (i) For almost every $t \in T$ and each $u \in domg(t, .), g(t, u) \ge \mathbf{s} \cdot ls_e g_n(t, u)$;
- (ii) $j : \mathbb{R}^l \times U \to \mathbb{R}$ is lsc as a function of (ξ, u) and proper.
- (iii) There exists a measurable function $h(\xi)$ such that $j(\xi, u) \ge h(\xi)$ and such that $h(\boldsymbol{\xi}_n) \in L^q(\mathbb{R})$ and $\mathbb{E}[h(\boldsymbol{\xi}_n) | \mathcal{B}_n] \ge \tilde{h}$ with $\tilde{h} \in L^q(\mathbb{R})$.
- (iv) for $n \in \overline{N}$ there exists $\mathbf{u}_n \in L^p(U)$ and functions \mathbf{k} and \mathbf{k}_0 in $L^p(\mathbb{R})$ such that $\|\mathbf{u}_n\|_U \leq \mathbf{k}$ and $j(\boldsymbol{\xi}_n, \mathbf{u}_n) \leq \mathbf{k}_0$.

(v) k_0 given by (iv) is such that $\mathbb{E}[\mathbf{k}_0 | \mathcal{B}_n] \leq \mathbf{k}_1$ with \mathbf{k}_1 in $L^p(\mathbb{R})$

Then for every function \mathbf{u} in $L^p(U)$, $I(g)(\mathbf{u}) \geq \mathbf{s}$ - $ls_e I(g_n)(\mathbf{u})$.

Proof: The aim of the proof is simply to apply the previous theorem to g_n . The first step is to check that g_n gives a sequence of normal proper integrands. Using [21, Proposition 14.45] f_n is a sequence of normal integrand if j is a normal integrand and the $\boldsymbol{\xi}_n$ are measurable functions. j is a normal integrand since it is an autonomous integrand which is lsc as a function of (ξ, u) [21, Example 14.31]. The sequence f_n is also proper using (ii). Consequently, the sequence g_n is well defined, using (iii) and Theorem 1 each g_n is a normal integrand, which is also proper using again (ii). Now :

(a) : is just given by (i)

- (b) : is we easily deduced from (iv) and (v) with k_0 replaced k_1 of (v).
- (c) : We have using (*iii*), $f_n(t, u) \ge -||u||_U + h(\boldsymbol{\xi}_n)$ and with $h_1 \stackrel{\text{def}}{=} 1$ and $h_{2,n} \stackrel{\text{def}}{=} \mathbb{E}[h(\boldsymbol{\xi}_n) | \mathcal{B}_n]$ we obtain

$$g_n(t,u) \ge -h_1(t) \|u\| + h_2(t) \ge -h_1(t) \|u\| + h(t)$$

where h_1 and h_2 belong to L^q with $h_1(t) > 0$ a.s.. This gives (c) for the g_n sequence.

Remark 1 Under the added following assumption

(*ii*) $j(\xi, .)$ is a convex function for all $\xi \in \mathbb{R}^l$,

we obtain that the f_n is a sequence of convex normal integrands and using (iii) and [23, Proposition 15] g_n is also convex.

Remark 2 The key condition to verify is (i) and we list here some conditions on the sequences $(\boldsymbol{\xi}_n)$ and on the sub σ -fields \mathcal{B}_n in order to fulfill it.

- If $\mathcal{B}_n \uparrow \mathcal{B}_\infty$ (*i.e* increasing sequence Kudo converging to \mathcal{B}_∞), $\boldsymbol{\xi}_n \stackrel{\text{def}}{=} \boldsymbol{\xi}$ for all $n \in \mathbb{N}$ and for each fixed $u \in \text{dom}g(t, .)$ we have $j(\boldsymbol{\xi}_n, u) \in L^1(\mathbb{R})$ then $g_n(t, u)$ is a Martingale for each fixed value of u which converges a.s $(\sup_n \mathbb{E}[|g_n(t, u)|] = \mathbb{E}[|f(t, u)|] < \infty)$ to g(t, u). Then we trivially get using $u_n \stackrel{\text{def}}{=} u$ that $g(t, u) \geq \mathbf{s}\text{-ls}_e g_n(t, u)$.
- If $\mathcal{B}_n \uparrow \mathcal{B}_\infty$, $f_n(t, u) \leq f(t, u)$ and for each fixed $u \in \text{dom}g(t, .)$ we have $j(\boldsymbol{\xi}, u) \in L^1(\mathbb{R})$ then $\mathbb{E}[f_n(t, u) | \mathcal{B}_n] \leq \mathbb{E}[f(t, u) | \mathcal{B}_n]$. The right hand side of previous equation converges a.s to $\mathbb{E}[f(t, u) | \mathcal{B}_\infty] = g(t, u)$, thus we have again $\limsup_{n \mapsto \infty} g_n(t, u) \leq g(t, u)$ a.s.
- If $\mathcal{B}_n \uparrow \mathcal{B}_\infty$ and for each fixed $u \ j(\boldsymbol{\xi}_n, u) \mapsto j(\boldsymbol{\xi}, u) \in L^1(\mathbb{R})$ then g_n is a \mathcal{B}_n adapted sequence which converges in L^1 as we have already seen. It is thus uniformly integrable. Suppose now that g_n is a Martingale in the limit, namely

$$\lim_{n \ge m \mapsto \infty} \mathbb{E}\left[g_n \,|\, \mathcal{B}_m\right] - g_m = 0 \text{ a.e}$$

then [15, Theorem 2] g_n converges almost everywhere to g and we obtain (i) $g(t, u) \ge \mathbf{s} \cdot \mathbf{ls}_e g_n(t, u)$;.

4.2 Inequality for the w-epigraphical lower limit

We give results here to obtain $I(f)(\mathbf{u}) \leq \mathbf{w}\text{-li}_e I(f_n)(\mathbf{u})$. This proof is based on the previous theorem but applied to the sequence (f_n^{\star}) . Following [2, p271], a sequence (f_n) is said to be *uniformly proper* if there exists a bounded sequence $x_n \in U$ such that $\sup_{n \in \mathbb{N}} f_n(x_n) < +\infty$. **Theorem 3** Let U be a separable reflexive Banach space with a separable dual U^* , $f_n : T \times U \to (-\infty, +\infty]$ $(n \ge 1)$ a sequence of normal proper integrands uniformly proper, $\gamma : T \times U \to (-\infty, +\infty]$ a proper integrand, p, q with $1 \le p < +\infty$, and $p^{-1} + q^{-1} = 1$, satisfying the following assumptions:

- (a') for almost every $t \in T$ and each $u \in dom f^{\star}(t, .), f_n(t, u) \leq \mathbf{w}$ -li_e $f_n(t, u);$
- (b') there exists a sequence (\mathbf{u}_n) in $L^p(U)$ and functions $k \in L^p$ and k_0 in $L^p(\mathbb{R})$, such that for each n, $\|\mathbf{u}_n(t)\| \leq k(t)$ and $f_n(t, \mathbf{u}_n(t)) \leq k_0(t)$ a.s.;
- (c') there exists a sequence (\mathbf{u}_n^{\star}) in $L^q(U^{\star})$ and functions h and h_0 in $L^q(\mathbb{R})$, such that for each n, $\|u_n^{\star}(t)\| \leq h(t)$ and $f_n^{\star}(t, u_n^{\star}(t)) \leq h_0(t)$ a.s.;

Then for every function \mathbf{u} in $L^p(U)$, $I(f)(\mathbf{u}) \leq \mathbf{w} - li_e I(f_n)(\mathbf{u})$.

Proof: we first prove that we can apply Theorem 2 to the sequence f_n^* . First note the the sequence f_n^* is a well defined sequence of proper normal convex integrands [20, Proposition 2]. If U is a reflexive Banach space, (f_n) a sequence of normal proper convex integrands which is also uniformly proper then we have [2, theorem 3.7, p 271]

$$(\operatorname{seq} \mathbf{w} - \operatorname{li}_e f_n)^* = \mathbf{s} - \operatorname{ls}_e f_n^* \tag{10}$$

Thus we obtain condition (a) of Theorem 2 form (a'). It is then easy to see that condition (b) (resp. (c)) of Theorem 2 is implied by condition (c') (resp. (b')). We then apply Theorem 2 on the conjugate sequence f_n^* to obtain \mathbf{s} -ls_e $I(f_n^*) \leq I(f^*)$ which combined with properties of conjugacy for integral functionals $I(f_n^*) = I(f_n)^*$ (This is true on $L \stackrel{\text{def}}{=} L^p(U)$ for $1 \leq p \leq \infty$ if L is decomposable and $I(f)(\mathbf{u}) < +\infty$ for at least one $\mathbf{u} \in L$ [20, Theorem 2]) gives \mathbf{s} -ls_e $I(f_n)^* \leq I(f)^*$. In order to conclude we have to use again [2, theorem 3.7,p 271] but on the $I(f_n)$ sequence which is uniformly proper by (b') $(k \in L^1)$ and on the reflexive Banach space $L^p(U)$.

Remark 3 Note that assumption (a') is obtained if we suppose that $\boldsymbol{\xi}_n$ converges a.s to $\boldsymbol{\xi}$ and that $f(\boldsymbol{\xi}, u)$ is jointly l.s.c.

4.3 Inequality for the w-epigraphical lower limit

We give here an other result to obtain $I(f)(\mathbf{u}) \leq \mathbf{w}\text{-li}_e I(f_n)(\mathbf{u})$. In the previous sub-section, almost sure convergence of $\boldsymbol{\xi}_n$ was proved to be sufficient to obtain the Inequality for the \mathbf{w} -epigraphical lower limit, we will here show that convergence in probability is sufficient to obtain the same inequality. Note however that the previous applies when U is a separable reflexive Banach space with a separable dual U^* and here $X = \mathbb{R}^l$. The proof is based on [10, Theorem 1]. We assume here that μ is a finite positive nonatomic complete measure. The Ioffe theorem [10, Theorem 1] applies to extendedreal-valued functions $\gamma(t, \xi, u)$ on $\Omega \times \mathbb{R}^l \times U$ such that $\gamma(t, \xi(t), u(t))$ is measurable for any measurable $\boldsymbol{\xi} \in L^0(\mathbb{R}^l)$ and $\mathbf{u} \in L^0(U)$. **Theorem 4** (Ioffe). Here μ is assumed to be also nonatomic. Let L and M satisfy (H_1) and (H_2) . Assume that $f(t, \xi, u)$ is $\mathcal{F} \times \mathbb{B}(\mathbb{R}^l \times U)$ -measurable, lsc in (ξ, u) and convex in u. In order that I(f) be lower semicontinuous on $L \times M$ and everywhere on $L \times M$ more than $-\infty$, it is necessary and (if I(f) is finite for at least one point in $L \times M$) sufficient that f satisfy the lower compactness property.

We do not here recall the technical hypotheses (H_1) and (H_2) but citing [10] we recall that L^p spaces $(1 \le p \le \infty)$ with norm or weak topologies satisfy these hypotheses. It is also important to note that (H_2) states that the topology in L is not weaker than the topology of convergence in measure.

 γ is assumed to satisfy the *lower compactness property* on $L \times M$ if any sequence $\gamma^{-}(t, \boldsymbol{\xi}_{k}(t), u_{k}(t))$ is weakly precompact in L^{1} whenever the $\boldsymbol{\xi}_{k}$ converges in L, the \mathbf{u}_{k} converges in M and $I(\gamma)(\boldsymbol{\xi}_{k}, \mathbf{u}_{k}) \leq a < \infty$ for all k (here $f^{-} = \min(f, 0)$).

A first result is that the function f satisfy the *lower compactness property* if it is a positive function [10, Theorem 3]

Proposition 10 For a sequence $\boldsymbol{\xi}_n$ of random variables, we define define a sequence (f_n) by $f_n(t, u) \stackrel{\text{def}}{=} j(\boldsymbol{\xi}_n(t), u)$. Suppose that Suppose that for $1 \leq p < \infty$

(i) $j : \mathbb{R}^l \times U \to \mathbb{R}$ is lsc as a function of (ξ, u) and proper.

(ii) $j(\xi, .)$ is a convex function for all $\xi \in \mathbb{R}^l$.

(iii) $j(\xi, u)$ is positive.

(iv) There exists $\mathbf{u} \in L^p(U)$ such that $I(f)(\mathbf{u})$ is finite.

(vi) the sequence $\boldsymbol{\xi}_n$ of measurable functions converges to $\boldsymbol{\xi}$ in measure.

Then for every function \mathbf{u} in $L^p(U)$, $I(f)(\mathbf{u}) \leq \mathbf{w} \cdot li_e I(f_n)(\mathbf{u})$

Proof: As in the proof of previous Proposition the sequence (f_n) defined by $f_n(t, u) = j(\xi_n(t), u)$ is a sequence of normal integrand and thus has the requested measurability assumption. We are using [10, Theorem 1] in a context of an homogeneous integrand and the result is then just a rephrase of [10, Theorem 1].

Remark 4 It is possible to lower assumption (iii) but it will restrict the values of p for which the conclusion of the Proposition is valid. For example using in the homogeneous case [10, Theorem 5 (Olech)]

 $(iii') \ j(\xi, u) \ge -c(\|\xi\| + \|u\|) + b \text{ with } (c, b) \in \mathbb{R}^2$

and assuming that ξ_n converges in $L^1(\mathbb{R}^l)$, we obtain

$$I(f)(\mathbf{u}) \leq \mathbf{w} - \mathrm{li}_e I(f_n)(\mathbf{u})$$

for **u** in $L^1(U)$ only.

Remark 5 Note that we want to use Proposition 10 in conjunction with Proposition 9 and we can notice that conditions which are requested for the existence of conditional expectations also gives lower compactness property. For example, from the condition (c') of Theorem 3, we easily obtain from the inequality

$$f_n(t,u) \ge \langle u_n^\star, u \rangle - f^\star(t, u_n^\star(t)) \tag{11}$$

that each f_n is \mathcal{F} -quasi integrable (This is exactly [23, Proposition 4]). And as we have seen in Theorem 1 the existence of conditional expectation can be proved assuming quasi integrability. But equation 11 is also used to obtain (s, q)-lsc properties (lsc properties in (ξ, u)) for all s (lsc properties in (ξ, u)) of I(f) [10, Berkovitz theorem, p 524].

4.4 Mosco convergence

Gathering Propositions 8, 9 and 10 we can conclude by the following main theorem :

Theorem 5 Under the hypotheses of Propositions 8 and 9 (or Propositions 8 and 9) the F_n sequence given by (5) Mosco converges to F in $L^p(U)$.

We illustrate the previous theorems on a small example [12, Example IV.4]. Our aim here is to illustrate the fact that σ -fields discretisation and random variable discretisation can be done independently. Here $T \stackrel{\text{def}}{=} (-1, 1)$ with its Borel σ -field and μ is the uniform law on T. We consider $b: T \mapsto T$ defined by $b(t) \stackrel{\text{def}}{=} |t|$ and the measurability constraint is $\mathcal{B} \stackrel{\text{def}}{=} \sigma(b)$. Thus \mathcal{B} functions are even functions. It is easy to verify [12, Lemme IV.32] that for a given H random variable :

$$\mathbb{E}\left[H \mid \mathcal{B}\right] = \frac{H + \bar{H}}{2} \quad \text{with} \quad \bar{H}(t) \stackrel{\text{\tiny def}}{=} H(-t) \tag{12}$$

We now consider a sequence of piece-wize constant functions (Q_n) :

$$Q_n: [0,1] \mapsto [0,1] \quad \text{such that} \quad Q_n(t) \stackrel{\text{def}}{=} \sum_{k=1}^n \frac{k}{n} \mathbb{I}_{I_{n,k}}(t) \quad \text{and} \quad I_{n,k} \stackrel{\text{def}}{=} (\frac{k-1}{n}, \frac{k}{n}]$$
(13)

and then $b_n(t) \stackrel{\text{def}}{=} Q_n(|t|)$ and the associated sequence of σ -fields $\mathcal{B}_n \stackrel{\text{def}}{=} \sigma(b_n)$. It is easy to check that b_n converges almost surely to b. Moreover a \mathcal{B}_n -measurable function is even and thus we have $\sigma(\mathcal{B}_n) \subset \sigma(\mathcal{B})$. Using [12, theorem III.30] we conclude that \mathcal{B}_n Kudo-converges to \mathcal{B} . Now given a sequence of bounded equi-continuous random variables $(\boldsymbol{\xi}_n)$ converging almost surely to $\boldsymbol{\xi}$ a uniformy distributed random variable, we consider the convex normal integrand $j(\xi, u) \stackrel{\text{def}}{=} (u - \xi)^2$ and problem (1). In order to get Mosco-convergence of the associated sequence $F_n(u)$, we have to check a

suficient condition to insure condition (i) of Proposition 9, i.e for any fixed $u \in \mathbb{R}$ we have amost surely

$$\lim_{n \to \infty} \mathbb{E}\left[(u - \boldsymbol{\xi}_n)^2 \,|\, \mathcal{B}_n \right] = \mathbb{E}\left[(u - \boldsymbol{\xi})^2 \,|\, \mathcal{B} \right].$$

We have :

$$\mathbb{E}\left[(u-\boldsymbol{\xi}_{n})^{2} \,|\, \mathcal{B}_{n}\right] = \sum_{k=1}^{n} \mathbb{I}_{I_{n,k}}(t) \int_{I_{n,k}} \frac{(u-\boldsymbol{\xi}_{n}(x))^{2} + (u-\boldsymbol{\xi}_{n}(-x))^{2}}{2n} dx.$$

And almost sure convergence to $\frac{(u-\boldsymbol{\xi}_n(t))^2+(u-\boldsymbol{\xi}_n(-t))^2}{2}$ which is $\mathbb{E}\left[(u-\boldsymbol{\xi})^2 \mid \boldsymbol{\mathcal{B}}\right]$ easily follows from equi-continuity and almost sure convergence of the sequence $(\boldsymbol{\xi}_n)$.

References

- 1. Zvi Arstein. Sensitivity to σ -fields of information in stochastic allocation. Stochastics and Stochastics reports, pages 41–63, 1991.
- 2. H. Attouch. Variational convergence for functions and operators. Pitman, 1984.
- J.R. Birge and R.J-B. Wets. Designing approximation schemes for stochastic optimization problems in particular for stochastic programs with recourse. *Math. Prog. Stud.*, (27):54–102, 1986.
- C. Castaing and M. Valadier. Convex analysis and measurable multifunctions, volume 580. Springer, Berlin, Germany, 1977.
- C. Choirat, C. Hess, and R. Seri. A functional version of the birkhoff ergodic theorem for a normal integrand : a variational approach. *The annals of Probability*, 31(1):63–92, 2003.
- J. Couvreux. From the mosco and slice convergences for convex normal integrands to that of integral functionals on L^p spaces. Siam journal on Control and Optimisation, 39(1):179–195, 2000.
- J. Dupačová and R.J.-B. Wets. Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *Ann. Stat.*, 16(4):1517–1549, 1988.
- 8. C. Hess. Conditions d'optimalité pour des fonctionnelles intégrales convexes sur les espaces $\mathcal{L}^{p}(E)$. CEREMADE, Université Paris-Dauphine, Paris, France, 1982.
- C. Hess. On the measurability of the conjugate and the subdifferential of a normal integrand. CEREMADE, Université Paris-Dauphine, Paris, France, 1994.
- A. D. Ioffe. On lower semicontinuity of integral functionnals. Siam journal on Control and Optimisation, 15(4):521–538, 1977.
- 11. J.L. Joly and F. De Thelin. Convergence of convex integrals in \mathcal{L}^p spaces. Journal of Mathematical Analysis and Applications, 54:230–244, 1976.
- K.Barty. Contributions à la discrétisation des contraintes de mesurabilité pour les problèmes d'optimisation stochastique. Thèse de doctorat http://cermics.enpc.fr/theses/, École Nationale des Ponts et Chaussées, 2004.

- 13. H. Kudo. A note on the strong convergence of σ -algebras. Ann. Probability, 2(1):76–83, 1974.
- R. Lucchetti and R.J-B. Wets. Convergence of minima of integral functionals, with applications to optimal control and stochastic optimization. *Statist. decisions*, 1(11):69–84, 1993.
- A.G. Mucci. Limits for martingale-like sequences. Pacific Journal Of Mathematics, 48(1):197–202, 1973.
- T. Pennanen. Epi-convergent discretizations of multistage stochastic programs. Mathematics of Operations Research, 30:245–256, 2005.
- 17. L. Piccinini. Convergence of nonmonotone sub- σ -fields and convergence of associated subspaces $L^{p}(\mathcal{B}_{n})$ ($p \in [1, +\infty]$). Preprint du Departement de Mathematiques de Montpellier, Journal of Mathematical Analysis and Applications, 225(1):73–90, September 1998.
- L. Piccinini. A new version of the multivalued Fatou lemma. Journal of Applied Analysis, 4(2):231–244, 1998.
- S.M. Robinson and R.J-B. Wets. Stability in two-stage stochastic programming. SIAM J. Control Optim., 6(25):1409–1416, 1987.
- R. T. Rockafellar. Convex integral functionals and duality. In Proceedings of the University of Wisconsin, Madison, Wisconsin, pages 215–236, 1971.
- R. T. Rockafellar and R. J-B. Wets. Variational Analysis, volume 317 of A series of Comprehensive Studies in Mathematics. Springer Verlag, 1998.
- C. Strugarek. Interchange of minimization and integration with measurability constraints. *Preprint* CERMICS http://cermics.enpc.fr/reports, 321, 2006.
- L. Thibault. Espérances conditionnelles d'intégrandes semi-continus. Annales de l'I.H.P, section B, 17(4):337–350, 1981.
- 24. T.Pennanen and M.Koivu. Epi-convergent discretization of stochastic programs via integration quadratures. *Numerische mathematik*, 2003.
- Timothy Van Zandt. Information, measurability, and continuous behavior. Journal of Mathematical Economics, pages 293–309, 2002.