KYBERNETIKA — VOLUME *35* (1999), NUMBER 3, PAGES 367 – 381

TWO DIMENSIONAL PROBABILITIES WITH A GIVEN CONDITIONAL STRUCTURE

Josef Štěpán[1] and Daniel Hlubinka

A properly measurable set $\mathcal{P} \subset \mathbb{X} \times M_1(\mathbb{Y})$ (where $\mathbb{X}, \mathbb{Y}$ are Polish spaces and $M_1(\mathbb{Y})$ is the space of Borel probability measures on $\mathbb{Y}$) is considered. Given a probability distribution $\lambda \in M_1(\mathbb{X})$ the paper treats the problem of the existence of $\mathbb{X} \times \mathbb{Y}$-valued random vector (ξ, η) for which $\mathcal{L}(\xi) = \lambda$ and $\mathcal{L}(\eta|\xi = x) \in \mathcal{P}_x$ λ-almost surely that possesses moreover some other properties such as "$\mathcal{L}(\xi, \eta)$ has the maximal possible support" or "$\mathcal{L}(\eta|\xi = x)$'s are extremal measures in $\mathcal{P}_x$'s". The paper continues the research started in [7].

1. INTRODUCTION

To clarify the purpose of the paper consider the following model for a transport that starts randomly at a locality $x \in \mathbb{X}$ and reaches a random locality $y \in \mathbb{Y}$: If (ξ, η) denotes the $(\mathbb{X} \times \mathbb{Y})$-valued random vector which value $(\xi(\omega), \eta(\omega)) = (x, y)$ designates the particular transport from x to y, we ask the probability distribution of the (ξ, η) to respect in the first place that

(i) the conditional distribution of terminals y given a departure point x should be subjected to a restriction $\mathcal{L}(\eta|\xi = x) \in \mathcal{P}_x$ almost surely, where $\mathcal{P}_x$ is a set of (admissible) probability distributions for the transport that originates at the x, while the departure distribution is given by a fixed probability distribution λ.

Moreover, we may venture to ask $\mathcal{L}(\xi, \eta)$ to follow some additional rules on the top of (i):

(ii) For each $x \in \mathbb{X}$ there is a prescribed terminal region $A_x \subset \mathbb{Y}$ and the transport should made as many localities $y \in A_x$ as possible accessible from the starting point x i.e., we ask for a transport (ξ, η) such that with the probability one the conditional distribution $\mathcal{L}(\eta|\xi = x)$ is supported by the set A_x and it possesses the maximal possible support.

(iii) If $F(x, \mu)$ is the payoff we receive for the transport that originates at an $x \in \mathbb{X}$ using a target probability distribution $\mu \in \mathcal{P}_x$ we ask for a transport (ξ, η)

[1] The paper was prepared with the support of Grant Agency of Charles University under contract 3051-10/716.

that provides the maximal payoff with the probability one, i.e. $\mathcal{L}(\eta|\xi = x) = \arg\max\{F(x,\mu), \mu \in \mathcal{P}_x\}$ almost surely.

(iv) If $\mathcal{P}_x$'s are convex sets of probability distributions we wish to design a simple (discrete) transport (ξ,η) such that $\mathcal{L}(\eta|\xi = x)$ is an extremal distribution in $\mathcal{P}_x$ almost surely, or, on the contrary,

(v) having a measure m on the target space $\mathbb{Y}$ we prefer an m-continuous solution (ξ,η), i.e. such that $\mathcal{L}(\eta|\xi = x)$ is a distribution absolutely continuous with respect to m almost surely.

If we interpret the $\mathcal{P}_x$'s in (i) as the sections of a Borel set $\mathcal{P}$ in $\mathbb{X} \times M_1(\mathbb{Y})$ we are able to prove (Theorem 1) the existence of a transport (ξ,η) that respects (i) whatever probability distribution λ supported by $\mathrm{pr}_{\mathbb{X}}(\mathcal{P})$ we may prescribe for the random variable ξ. If we interpret the A_x's in (ii) as the values of a multifunction $A : \mathbb{X} \to 2^{\mathbb{Y}}$ which graph is a Borel set in $\mathbb{X} \times \mathbb{Y}$, Theorems 2 and 3 propose sufficient conditions for the existence of a transport that respects both (i) and (ii). The Corollaries 2,3 and 4 deal with a possibility to construct a transport (ξ,η) that satisfies the rules (i,iii), (i,iv) and (i,v), respectively.

A typical example of a set $\mathcal{P}$ we have on mind is a set $\mathcal{P} \subset \mathbb{X}$ each of which sections $\mathcal{P}_x$'s is defined as a moment problem. The Corollary 1 treats the situation.

The techniques used in our proofs depend heavily on the results coming from the theory of the analytic sets, on its cross-section theorems in the first place. We refer to [3] for the elements of the theory. The paper introduces also a concept of an universally measurable (closed valued) multifunction to generalize that of a lower semicontinuous multifunction (see [1]). A characterization of the universal measurability, given by our Lemma 1 may be of some interest by itself.

Generally, the paper is a contribution to the research on a possibility to construct a probability distribution with given moments, marginals and a conditional structure, see [2] for the latest developments. Actually, the paper continues and in a way completes the research started in [7]. Most importantly, the present paper clarifies the problem met in [7] when trying to construct the transports with the properties (i) and (ii) and introduces further nontrivial examples of the $\mathcal{P}$-sets the theory may be applied to (Corollaries 2 and 4).

2. DEFINITIONS AND RESULTS

Fix first metric spaces $\mathbb{X}$ and $\mathbb{Y}$ and denote by $\mathcal{F}(\mathbb{X}), \mathcal{G}(\mathbb{X}), \mathcal{B}(\mathbb{X}), \mathcal{A}(\mathbb{X})$, and $\mathcal{U}(\mathbb{X})$ all closed, open, Borel, analytic, and universally measurable sets in $\mathbb{X}$. Recall that a set $A \subset \mathbb{X}$ is analytic if there exists a Polish space $\mathbb{Z}$ and continuous map $\phi : \mathbb{Z} \to \mathbb{X}$ such that $A = \phi(\mathbb{Z})$, that

$$\mathcal{B}(\mathbb{X}) \subset \mathcal{A}(\mathbb{X}) \subset \mathcal{U}(\mathbb{X}) \quad \text{and}$$
$$\mathcal{B}(\mathbb{X} \times \mathbb{Y}) = \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}) \subset \mathcal{U}(\mathbb{X}) \otimes \mathcal{U}(\mathbb{Y}) \subset \mathcal{U}(\mathbb{X} \times \mathbb{Y})$$

and also recall that

$$\mathcal{U}(\mathbb{X}) = \{U \subset \mathbb{X} : \ \forall \mu \in M_1(\mathbb{X}) \ \exists B_1 \subset U \subset B_2, \ B_i \in \mathcal{B}(\mathbb{X}), \mu(B_2 \setminus B_1) = 0\},$$

where we have denoted the space of all Borel probability measures on $\mathbb{X}$ by $M_1(\mathbb{X})$. Let us agree that having a $\mu \in M_1(\mathbb{X})$, we denote by μ also its uniquely determined extension from $\mathcal{B}(\mathbb{X})$ to $\mathcal{U}(\mathbb{X})$. Moreover, using the notation λ^* for outer measures, we denote

$$M_1^*(B) = \{\lambda \in M_1(\mathbb{X}) : \lambda^*(B) = 1\} \text{ for a } B \subset \mathbb{X}.$$

Whenever speaking about a topology on $M_1(\mathbb{X})$ we mean its standard *weak topology* that makes the space metric and Polish if the space $\mathbb{X}$ has the property.

Agree that any map $A : \mathbb{X} \to 2^{\mathbb{Y}}$ will be referred to as a *multifunction from* $\mathbb{X}$ *to* $\mathbb{Y}$, we shall write $A : \mathbb{X} \rightrightarrows \mathbb{Y}$ in this case and denote

$$\mathrm{Graph}(A) := \{(x, y) \in \mathbb{X} \times \mathbb{Y} : y \in A_x\},$$

where $A_x \subset \mathbb{Y}$ is the value of A at a point $x \in \mathbb{X}$.

Define $A : \mathbb{X} \rightrightarrows \mathbb{Y}$ to be *U-measurable* and *strongly U-measurable* if

$$\begin{aligned} &\{x \in \mathbb{X} : A_x \cap G \neq \emptyset\} \in \mathcal{U}(\mathbb{X}), \ \forall G \in \mathcal{G}(\mathbb{Y}) \text{ and} \\ &\{x \in \mathbb{X} : A_x \cap B \neq \emptyset\} \in \mathcal{U}(\mathbb{X}), \ \forall B \in \mathcal{B}(\mathbb{Y}), \text{ respectively.} \end{aligned}$$

Observe that if we fix $V \in \mathcal{G}(\mathbb{Y})$ and $Z \subset \mathbb{X}$, $Z \notin \mathcal{U}(\mathbb{X})$, put $A_x = V$ for $x \notin Z$, $A_x = \overline{V}$ for $x \in Z$, we have exhibited an example of a multifunction $A = (A_x,\ x \in \mathbb{X})$ that is U-measurable but not strongly U-measurable.

A multifunction $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ will be called a *closed valued multifunction (CVM)* if $F_x \in \mathcal{F}(\mathbb{Y})$ for all $x \in \mathbb{X}$ and a *lower semicontinuous multifunction* if it is closed valued and $\{x \in \mathbb{X} : F_x \cap G \neq \emptyset\} \in \mathcal{G}(\mathbb{X})$ for all $G \in \mathcal{G}(\mathbb{Y})$. We refer to Lemma 1 for a necessary and sufficient condition for a CVM F to be (strongly) U-measurable, and observe that a multifunction $A : \mathbb{X} \rightrightarrows \mathbb{Y}$ is U-measurable iff the CVM $A_C := \{\overline{A_x}, x \in \mathbb{X}\}$ has the property. Thus

$$\mathrm{Graph}(A) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y}) \Rightarrow \mathrm{Graph}(A_C) \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}) \tag{1}$$

according to Lemma 1 (iv) and (i). Especially, we observe that

$$\mathrm{Graph}(A) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y}), A_x \in \mathcal{F}(\mathbb{Y}) \text{ for } x \in \mathbb{X} \Rightarrow \mathrm{Graph}(A) \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}) \tag{2}$$

Putting $S_\mu = \mathrm{supp}(\mu)$ for $\mu \in M_1(\mathbb{Y})$ where $\mathbb{Y}$ is a separable metric space we get an example of CVM $S = \big(S_\mu,\ \mu \in M_1(\mathbb{Y})\big)$ from $M_1(\mathbb{Y})$ to $\mathbb{Y}$ that is obviously lower semicontinuous. Recall that for a finite Borel measure μ on $\mathbb{Y}$ we define

$$\begin{aligned} \mathrm{supp}(\mu) &:= \bigcap \{F, F \in \mathcal{F}(\mathbb{Y}), \mu(F) = \mu(\mathbb{Y})\} \\ &= \{y \in \mathbb{Y} : \mu(G) > 0,\ \forall G \in \mathcal{G}(\mathbb{Y}), y \in G\}. \end{aligned}$$

For the rest of the paper we shall assume the fixed spaces $\mathbb{X}$ *and* $\mathbb{Y}$ *to be Polish.*

Our results concern subsets $\mathcal{P}$ in $\mathbb{X} \times M_1(\mathbb{Y})$ such that

$$\mathcal{P} \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big) \cup \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$$

mostly. To such a set we may attach naturally a set Output $\mathcal{P} \subset \mathbb{X} \times \mathbb{Y}$ defined by[2]

$$\begin{aligned} \text{Output}\,\mathcal{P} &:= \{(x,y) \in \mathbb{X} \times \mathbb{Y} : \exists\, \mu \in \mathcal{P}_x,\ y \in \text{supp}(\mu)\},\ \text{i.e.} \\ (\text{Output}\,\mathcal{P})_x &= \bigcup\{\text{supp}(\mu), \mu \in \mathcal{P}_x\},\ x \in \mathbb{X}. \end{aligned}$$

See Lemma 2 for a result that claims a topological stability of the $\mathcal{P} \to \text{Output}\,\mathcal{P}$ operation.

To illustrate this, consider a multifunction $A : \mathbb{X} \rightrightarrows \mathbb{Y}$ with $A_x \in \mathcal{U}(\mathbb{Y})$ and put $\mathcal{P}_A := \{(x,\mu) \in \mathbb{X} \times M_1(\mathbb{Y}) : \mu(A_x) = 1\}$. It is easy to verify that Output $\mathcal{P}_A = A_C$. Hence Lemma 4 (ii), (iii) together with Lemma 2 (ii), (iii) state that

$$\begin{aligned} \text{Graph}(A) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y}) &\Rightarrow \mathcal{P}_A \in \mathcal{A}(\mathbb{X} \times M_1(\mathbb{Y})) \\ &\Rightarrow \text{Output}\,\mathcal{P}_A \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}) \\ \text{Graph}(A) \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}) &\Rightarrow \mathcal{P}_A \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(M_1(\mathbb{Y})) \\ &\Rightarrow \text{Output}\,\mathcal{P}_A \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}). \end{aligned} \tag{3}$$

Frequently we need $\mathcal{P} \subset \mathbb{X} \times M_1(\mathbb{Y})$ such that $\big((\text{Output}\,\mathcal{P})_x, x \in \mathbb{X}\big)$ is a closed valued multifunction $\mathbb{X} \rightrightarrows \mathbb{Y}$. We can achieve that assuming a weak form of convexity for all the sections $\mathcal{P}_x$'s (see [7] and our Lemma 3). We shall say that a $\mathcal{P} \subset \mathbb{X} \times M_1(\mathbb{Y})$ satisfies *CS-condition* if

$$\forall\ \big(x \in \mathbb{X}, (\mu_n, n \in \mathbb{N}) \subset \mathcal{P}_x\big)\ \exists\ \left(\alpha_n > 0, \sum_1^\infty \alpha_n = 1 : \sum_1^\infty \alpha_n \mu_n \in \mathcal{P}_x\right).$$

A typical example of a $\mathcal{P} \subset \mathbb{X} \times M_1(\mathbb{Y})$ our results may be applied to is a set $\mathcal{P}$ each of its sections is defined by a moment problem:

$$\mathcal{P}_x := \left\{\mu \in M_1(\mathbb{Y}) : \int_{\mathbb{Y}} f_i(x,y)\mu(\mathrm{d}y) = c_i(x), i \in I\right\},\ x \in \mathbb{X}, \tag{4}$$

where $I \neq \emptyset$ is an index set and for $i \in I$
$f_i : \mathbb{X} \times \mathbb{Y} \to [0,+\infty],\ c_i : \mathbb{X} \to [0,+\infty]$ are Borel measurable functions. (5)

Remark that if I is at most countable set then such a $\mathcal{P}$ belongs to $\mathcal{B}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$ by Lemma 4(i). If f_i's are bounded continuous, c_i's continuous then regardless the cardinality of the set I, $\mathcal{P} \in \mathcal{F}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$. Either situation provides a $\mathcal{P}$ for which the CS-condition holds.

Recall that a map $H : \mathbb{X} \to \mathbb{Y}$ is called *universally measurable* if it is a map that is measurable with respect to the σ-algebras $\mathcal{U}(\mathbb{X})$ and $\mathcal{U}(\mathbb{Y})$ which is as to say that it is measurable w.r.t. the σ-algebras $\mathcal{U}(\mathbb{X})$ and $\mathcal{B}(\mathbb{Y})$ according to Lemma 8.4.6. in [3]. A universally measurable map $x \to \mathsf{P}^x$ from $\mathbb{X}$ into $M_1(\mathbb{Y})$ will be called here a *universally measurable Markov kernel* (UMK). Note that $x \to \mathsf{P}^x$ is a UMK if

[2] We denote by A_x the section of $A \subset \mathbb{X} \times \mathbb{Y}$ at a point $x \in \mathbb{X}$

and only if $x \to \mathsf{P}^x(B)$ is a universally measurable ($\overset{\text{u.m.}}{\to}$) function for all $B \in \mathcal{B}(\mathbb{Y})$. Indeed since

$$x \overset{\text{u.m.}}{\to} \mathsf{P}^x \Rightarrow x \overset{\text{u.m.}}{\to} \mathsf{P}^x(B), \forall\, B \in \mathcal{B}(\mathbb{Y}) \Rightarrow x \overset{\text{u.m.}}{\to} \mathsf{P}^x(f), \forall\, f \in \mathrm{C_b}(\mathbb{Y}) \Rightarrow x \overset{\text{u.m.}}{\to} \mathsf{P}^x,$$

where the first implication follows by the well known fact that $\mu \to \mu(B)$ are for all $B \in \mathcal{B}(\mathbb{Y})$ $\big(\mathcal{B}(M_1), \mathcal{B}\big)$ measurable, the second implication can be verified by approximating $f \in \mathrm{C_b}$ by Borel step functions and the third follows by separability of $M_1(\mathbb{Y})$ that implies $\mathcal{B}\big(M_1(\mathbb{Y})\big) = \sigma\{\mu : |\mu(f) - \mu_0(f)| < \varepsilon; \varepsilon > 0, \mu_0 \in M_1(\mathbb{Y}), f \in \mathrm{C_b}\}$. Hence, for a $\lambda \in M_1(\mathbb{X})$ and a UMK $x \to \mathsf{P}^x$ we define correctly a probability measure $\mathsf{P}^\lambda \in M_1(\mathbb{X} \times \mathbb{Y})$ by

$$\mathsf{P}^\lambda(A \times B) = \int_A \mathsf{P}^x(B)\, \lambda(\mathrm{d}x) \text{ where } A \times B \in \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}).$$

Remark 1. Let $f : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ be a universally measurable function. Then the sections $f(x, \cdot), x \in \mathbb{X}$ and $x \to \int_{\mathbb{Y}} f(x,y) \mathsf{P}^x(\mathrm{d}y)$ are universally measurable functions in the sense $\mathbb{Y} \to [0, \infty]$ and $\mathbb{X} \to [0, \infty]$, respectively. Moreover, if $\lambda \in M_1(\mathbb{X})$ then

$$\int_{\mathbb{X}\times\mathbb{Y}} f \,\mathrm{d}\mathsf{P}^\lambda = \int_{\mathbb{X}} \int_{\mathbb{Y}} f(x,y) \mathsf{P}^x(\mathrm{d}y)\, \lambda(\mathrm{d}x) \tag{6}$$

especially, $\mathsf{P}^\lambda(U) = \int_{\mathbb{X}} \mathsf{P}^x(U_x)\, \lambda(\mathrm{d}x),\ U \in \mathcal{U}(\mathbb{X} \times \mathbb{Y})$ defines the extension of P^λ from $\mathcal{B}(\mathbb{X} \times \mathbb{Y})$ to $\mathcal{U}(\mathbb{X} \times \mathbb{Y})$.

The universal measurability of the sections $f(x, \cdot)$ is an obvious statement. To verify the rest assume first that f *is Borel measurable*. Then the map $H_f : x \to \int_{\mathbb{Y}} f(x,y) \mathsf{P}^x(\mathrm{d}y)$ is received by substituting $x \to (x, \mathsf{P}^x)$ from $\mathbb{X}$ into $\mathbb{X} \times M_1(\mathbb{Y})$ to $(x, \mu) \to \int_{\mathbb{Y}} f(x,y)\, \mu(\mathrm{d}y)$ from $\mathbb{X} \times M_1(\mathbb{Y})$ into $[0, \infty]$. The former of the maps is easily seen to be measurable w.r.t. the σ-algebras $\mathcal{U}(\mathbb{X})$ and $\mathcal{B}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$ because $x \to \mathsf{P}^x$ is a UMK, while the latter one is a Borel measurable map by Lemma 4 (i) in Section 3. Hence the map H_f is universally measurable which implies, putting $f = I_C$ that $\mathsf{P}^\lambda(C) = \int_{\mathbb{X}} \mathsf{P}^x(C_x)\, \lambda(\mathrm{d}x)$ for $C \in \mathcal{B}(\mathbb{X} \times \mathbb{Y})$. A standard procedure extends the latter definition of P^λ to the equality (6). For a general f and $\lambda \in M_1(\mathbb{X})$ there are Borel measurable functions $f_1 \le f \le f_2$ such that $f_1 = f_2$ $[\lambda]$-almost surely. Then $H_{f_1} \le H_f \le H_{f_2}$ on $\mathbb{X}$, $H_{f_1} = H_{f_2}$ $[\lambda]$-almost surely according to (6) applied to f_1 and f_2. Hence, the H_f is universally measurable and

$$\int_{\mathbb{X}\times\mathbb{Y}} f \mathrm{d}\mathsf{P}^\lambda = \int_{\mathbb{X}\times\mathbb{Y}} f_1 \mathrm{d}\mathsf{P}^\lambda = \int_{\mathbb{X}} \int_{\mathbb{Y}} f_1(x,y) \mathsf{P}^x(\mathrm{d}y)\, \lambda(\mathrm{d}x) = \int_{\mathbb{X}} \int_{\mathbb{Y}} f(x,y) \mathsf{P}^x(\mathrm{d}y)\, \lambda(\mathrm{d}x)$$

according to the first part of our argument.

Let us agree that whenever we shall speak about an $(\mathbb{X} \times \mathbb{Y})$-*valued vector* (ξ, η) we mean a map defined on a probability space $(\Omega, \mathcal{E}, \mathsf{P})$ that is measurable with respect to the σ-algebras $\mathcal{E}$ and $\mathcal{U}(\mathbb{X} \times \mathbb{Y})$. This definition makes the random variables ξ and η to be measurable w.r.t. the σ-algebras $\mathcal{U}(\mathbb{X})$ and $\mathcal{U}(\mathbb{Y})$, respectively and it presents no loss of generality (see Lemma 8.4.6. in [3], again). Recall that if we have an $(\mathbb{X}\times\mathbb{Y})$-valued random vector (ξ,η), then a UMK $x\to \mathsf{P}^x$ from $\mathbb{X}$ into $M_1(\mathbb{Y})$ is called a *regular conditional distribution* of η given the values of ξ if

$$\mathsf{P}[\xi\in A,\eta\in B]=\int_A \mathsf{P}^x(B)\lambda(\mathrm{d}x),\ A\in\mathcal{B}(\mathbb{X}),B\in\mathcal{B}(\mathbb{Y}),\ \text{where } \lambda=\mathcal{L}(\xi). \tag{7}$$

It is a well known fact that a regular conditional distribution of η given the values of ξ exists and it is determined uniquely almost surely w.r.t. $\mathcal{L}(\xi)$ provided that $\mathbb{X}$ and $\mathbb{Y}$ are Polish spaces (see [8], p.126). We shall denote as usual $\mathsf{P}^x=\mathcal{L}(\eta|\xi=x)$ for any regular conditional distribution $x\to\mathsf{P}^x$ of η given the values of ξ.

Obviously we may paraphrase Remark 1 as

Remark 2. If (ξ,η) is an $(\mathbb{X}\times\mathbb{Y})$-valued random vector such that

$$\mathcal{L}(\xi)=\lambda \text{ and } \mathcal{L}(\eta|\xi=x)=\mathsf{P}^x\ \lambda\text{-almost surely} \tag{8}$$

holds for a $\lambda\in M_1(\mathbb{X})$ and a UMK $x\to\mathsf{P}^x$ then

$$\mathcal{L}(\xi,\eta)=\mathsf{P}^\lambda \text{ and } \mathsf{E}[f(\xi,\eta)|\xi=x]=\int_{\mathbb{Y}} f(x,y)\mathsf{P}^x(\mathrm{d}y)\ \lambda\text{-almost surely}$$

holds for any universally measurable function $f\in L_1(\mathsf{P}^\lambda)$.

A reverse statement to Remark 2 is provided by

Remark 3. Given a UMK $x\to\mathsf{P}^x$ and a $\lambda\in M_1(\mathbb{X})$ there is an $(\mathbb{X}\times\mathbb{Y})$-valued random vector (ξ,η) such that (8) holds.

To construct a vector (ξ,η) possessing the properties (8) put $(\Omega,\mathcal{F},\mathsf{P}):=\big(\mathbb{X}\times\mathbb{Y},\mathcal{U}(\mathbb{X}\times\mathbb{Y}),\mathsf{P}^\lambda\big)$ and $\xi:=\mathrm{pr}_{\mathbb{X}},\ \eta:=\mathrm{pr}_{\mathbb{Y}}$, where $\mathrm{pr}_{\mathbb{X}}:\mathbb{X}\times\mathbb{Y}\to\mathbb{X}$ denotes the canonical projection of $\mathbb{X}\times\mathbb{Y}$ onto $\mathbb{X}$.

More generally, given a $\mathcal{P}\subset\mathbb{X}\times M_1(\mathbb{Y})$ and $\lambda\in M_1(\mathbb{X})$ our results concern mainly the existence of an $(\mathbb{X}\times\mathbb{Y})$-valued random vector (ξ,η) such that

$$\mathcal{L}(\xi)=\lambda \text{ and } \mathcal{L}(\eta|\xi=x)\in\mathcal{P}_x \text{ almost surely w.r.t. } \lambda. \tag{9}$$

A random vector (ξ,η) with properties (9) shall be called a $(\mathcal{P},\lambda)$-*vector*. Observe that the random vector (ξ,η) the existence of which is stated by Remark 3 is in fact $(\mathcal{P},\lambda)$-vector with $\mathcal{P}=\mathrm{Graph}(x\to\mathsf{P}^x)$. A simple argument verifies that $\mathcal{P}\in\mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big)$ in this case as a consequence of the universal measurability of $x\to\mathsf{P}^x$.

Remark 4. If $A:\mathbb{X}\rightrightarrows\mathbb{Y}$ is a multifunction with $\mathrm{Graph}(A)\in\mathcal{U}(\mathbb{X}\times\mathbb{Y})$ and $\lambda\in M_1(\mathbb{X})$ then

(i) (ξ,η) is a $(\mathcal{P}_A,\lambda)$-vector.

(ii) $\mathsf{P}[(\xi,\eta)\in\mathrm{Graph}(A)|\xi=x]=1$ λ-almost surely.

(iii) $\mathsf{P}[(\xi,\eta)\in \mathrm{Graph}(A)]=1$

are equivalent statements because $\mathsf{P}[(\xi,\eta)\in \mathrm{Graph}(A)|\xi=x]=\mathsf{P}^x(A_x)$ according to Remark 2.

Finally, we shall say that a $(\mathcal{P},\lambda)$-vector is *maximally supported* if

$$\mathrm{supp}\big(\mathcal{L}(\eta|\xi=x)\big)\supset \mathrm{supp}\big(\mathcal{L}(\eta'|\xi'=x)\big)\lambda\text{-a.s. for any } (\mathcal{P},\lambda)\text{-vector } (\xi',\eta').$$

Note that if a $(\mathcal{P},\lambda)$-vector is maximally supported then according to Lemma 5 in Section 3 $\mathrm{supp}\big(\mathcal{L}(\xi,\eta)\big)\supset \mathrm{supp}\big(\mathcal{L}(\xi',\eta')\big)$ for any $(\mathcal{P},\lambda)$-vector (ξ',η') and that the implication can not be reversed according the counterexample that follows the proof of the lemma.

Our main results are

Theorem 1. Consider $\mathcal{Q}\subset \mathbb{X}\times M_1(\mathbb{Y})$, a multifunction $A:\mathbb{X}\rightrightarrows\mathbb{Y}$ and $\lambda\in M_1^*\big(D(\mathcal{Q},A)\big)$, where $D(\mathcal{Q},A):=\{x\in\mathbb{X}:\exists\,\mu\in\mathcal{Q}_x,\mu^*(A_x)=1\}$. Then either

$$\mathcal{Q}\in\mathcal{A}\big(\mathbb{X}\times M_1(\mathbb{Y})\big),\ \mathrm{Graph}(A)\in\mathcal{A}(\mathbb{X}\times\mathbb{Y})$$

or

$$\mathcal{Q}\in\mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big),\ \mathrm{Graph}(A)\in\mathcal{U}(\mathbb{X})\otimes\mathcal{B}(\mathbb{Y})$$

implies that there is a $(\mathcal{Q}\cap\mathcal{P}_A,\lambda)$-vector (ξ,η).

Observe that according to Remark 4 the theorem states exactly that there is a $(\mathcal{Q},\lambda)$-vector (ξ,η) such that $\mathsf{P}[(\xi,\eta)\in\mathrm{Graph}(A)]=1$.

Theorem 2. Assume that $\mathcal{P}\in\mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big)$ satisfies the CS-condition and is such that $\mathrm{Output}\,\mathcal{P}\in\mathcal{U}(\mathbb{X})\otimes\mathcal{B}(\mathbb{Y})$. Then for each $\lambda\in M_1^*(\mathrm{pr}_{\mathbb{X}}\mathcal{P})$ there exists a $(\mathcal{P},\lambda)$-vector (ξ,η) such that

$$\mathrm{supp}\big(\mathcal{L}(\eta|\xi=x)\big)=(\mathrm{Output}\,\mathcal{P})_x\ \lambda\text{-almost surely.}\tag{10}$$

Remark that a $(\mathcal{P},\lambda)$-vector (ξ,η) that possesses the property (10) is maximally supported. We do not know whether the implications $\mathcal{P}\in\mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big)\Rightarrow \mathrm{Output}\,\mathcal{P}\in\mathcal{U}(\mathbb{X})\otimes\mathcal{B}(\mathbb{Y})$ is true or not. Observe (3) for the positive answer for a very simple choice of $\mathcal{P}$.

Theorem 3. Assume that $\mathcal{R}\subset\mathbb{X}\times M_1(\mathbb{Y})$ and a multifunction $A:\mathbb{X}\rightrightarrows\mathbb{Y}$ are such that

$$\begin{aligned}&\mathrm{Graph}(A)\in\mathcal{A}(\mathbb{X}\times\mathbb{Y})\cap\mathcal{U}(\mathbb{X})\otimes\mathcal{B}(\mathbb{Y}),\\ &\mathcal{R}\in\mathcal{A}\big(\mathbb{X}\times M_1(\mathbb{Y})\big)\cap\mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big)\text{ and satisfies the CS-condition.}\end{aligned}\tag{11}$$

Then for each $\lambda\in M_1^*\big(D(\mathcal{R},A)\big):=M_1^*\{x\in\mathbb{X}:\exists\,\mu\in\mathcal{R}_x,\mu(A_x)=1\}$ there exists a maximally supported $(\mathcal{R}\cap\mathcal{P}_A,\lambda)$-vector (ξ,η).

Observe that Theorem 3 may be applied to $\mathcal{R}$ and A such that both $\mathcal{R}$ and $\mathrm{Graph}(A)$ are simply Borel sets and that, in this situation, provides a generalization to the second part of Theorem 1 in [7].

3. PROOFS

Lemma 1. Let $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a CVM, and $A : \mathbb{X} \rightrightarrows \mathbb{Y}$ a multifunction. Then

(i) F U-measurable

(ii) $\mathrm{Graph}(F) \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$

(iii) F strongly U-measurable,

are equivalent statements.
Moreover

(iv) $\mathrm{Graph}(A) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y}) \cup \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}) \Rightarrow A$ is strongly U-measurable.

(v) F lower semicontinuous $\Rightarrow \mathrm{Graph}(F) \in \mathcal{B}(\mathbb{X} \times \mathbb{Y})$.

Proof. It is sufficient to verify (i)⇒(ii), (iv), (v).
(i)⇒(ii): To verify this we simply write

$$\mathbb{X}\times\mathbb{Y}\setminus\mathrm{Graph}(F) = \{(x,y) : y \notin F_x\} = \bigcup_{G\in\mathcal{V}} \{x : F_x \cap G = \emptyset\} \times G \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}) \quad (12)$$

where $\mathcal{V}$ is a countable topological base in $\mathbb{Y}$.

(iv): Let $B \in \mathcal{B}(\mathbb{Y})$. Then $\{x : A_x \cap B \neq \emptyset\} = \mathrm{pr}_{\mathbb{X}}[\mathrm{Graph}(A) \cap (\mathbb{X} \times B)] \in \mathcal{U}(\mathbb{X})$ by 8.4.4. and 8.4.6. in [3] because $\mathrm{Graph}(A) \cap (\mathbb{X} \times B) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y}) \cup \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$

(v): It follows by (12) because $\{x : F_x \cap G = \emptyset\} = \mathbb{X} \setminus \{x : F_x \cap G \neq \emptyset\} \in \mathcal{F}(\mathbb{X})$ for $G \in \mathcal{G}(\mathbb{Y})$ as F is lower semicontinuous. □

Lemma 2. (see also Lemma in [7] for the implication (i) below)

(i) $\mathcal{P} \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big) \Rightarrow \mathrm{Output}\,\mathcal{P} \in \mathcal{A}(\mathbb{X} \times \mathbb{Y})$

(ii) $\mathcal{P} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big) \Rightarrow \mathrm{Output}\,\mathcal{P} \in \mathcal{U}(\mathbb{X} \times \mathbb{Y})$

(iii) $\mathcal{P} \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big), (\mathrm{Output}\,\mathcal{P})_x \in \mathcal{F}(\mathbb{Y})$ for all $x \in \mathbb{X} \Rightarrow \mathrm{Output}\,\mathcal{P} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$.

Proof. (iii) follows by (iv) and by [(iii)⇒(ii)] in Lemma 1 as $x \to (\mathrm{Output}\,\mathcal{P})_x$ represents a closed valued multifunction $\mathbb{X} \rightrightarrows \mathbb{Y}$.
We shall prove (i) and (ii): Put $D := \{(x, y, \mu) \in \mathbb{X} \times \mathbb{Y} \times M_1(\mathbb{Y}) : (x, \mu) \in \mathcal{P}, y \in \mathrm{supp}(\mu)\}$, observe that $\mathrm{Output}\,\mathcal{P} = \mathrm{pr}_{\mathbb{X}\times\mathbb{Y}}(D)$, and $D = (\mathcal{P} \times \mathbb{Y}) \cap (\mathbb{X} \times \mathrm{Graph}(S))$, where $S : M_1(\mathbb{Y}) \rightrightarrows \mathbb{Y}$ is the closed valued correspondence defined by $S_\mu = \mathrm{supp}(\mu)$. Because S is easily seen to be lower semicontinuous it follows by (v) in Lemma 1 that

$$\mathcal{P} \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big) \Rightarrow D \in \mathcal{A}\big(\mathbb{X} \times \mathbb{Y} \times M_1(\mathbb{Y})\big) \Rightarrow \mathrm{pr}_{\mathbb{X}\times\mathbb{Y}}(D) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y})$$

and

$$\mathcal{P} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big) \Rightarrow D \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(\mathbb{Y} \times M_1(\mathbb{Y})\big) \Rightarrow \mathrm{pr}_{\mathbb{X}\times\mathbb{Y}}(D) \in \mathcal{U}(\mathbb{X} \times \mathbb{Y})$$

(again by 8.4.4. and 8.4.6. in [3]). □

Lemma 3. Let $\mathcal{P} \subset \mathbb{X} \times M_1(\mathbb{Y})$ satisfies the CS-condition. Then

$$\forall\, x \in \mathrm{pr}_{\mathbb{X}}\mathcal{P}\ \exists\, \mu_x \in \mathcal{P}_x \text{ such that } \mathrm{supp}(\mu_x) = (\mathrm{Output}\,\mathcal{P})_x$$

and therefore $x \to (\mathrm{Output}\,\mathcal{P})_x$ is a closed valued multifunction $\mathbb{X} \rightrightarrows \mathbb{Y}$.

To verify the statement it is sufficient to read carefully the first part of the proof of Theorem 2 in [7]. We shall do it for the sake of completeness of our presentation.

P r o o f. Let $x \in \mathrm{pr}_{\mathbb{X}}\mathcal{P}$ and $\{\mu_1, \mu_2, \ldots\}$ a dense set in $\mathcal{P}_x$. By the CS-condition we have $\mu_x = \sum_1^\infty \alpha_n \mu_n \in \mathcal{P}_x$ for some $\alpha_n > 0$, $\sum_1^\infty \alpha_n = 1$. Obviously $\mathrm{supp}(\mu_x) \subset (\mathrm{Output}\,\mathcal{P})_x$, to verify the reverse inclusion choose $y \in (\mathrm{Output}\,\mathcal{P})_x$ and $\mathsf{V}_y \in \mathcal{G}(\mathbb{Y})$ its arbitrary neighbourhood. There is a $\nu \in \mathcal{P}_x$ such that $y \in \mathrm{supp}(\nu)$. If $\mu_{n_k} \to \nu$ weakly then for an arbitrary open neighbourhood V_y of y $\limsup \mu_{n_k}(\mathsf{V}_y) \geq \limsup \nu(\mathsf{V}_y) > 0$. Thus, $\mu_{n_k}(\mathsf{V}_y) > 0$ for a $k \in \mathbb{N}$, hence $\mu_x(\mathsf{V}_y) \geq \sum \alpha_{n_k}\mu_{n_k}(\mathsf{V}_y) > 0$. It follows that $y \in \mathrm{supp}(\mu_x)$. □

Lemma 4. Let $f : \mathbb{X} \times \mathbb{Y} \to [0, \infty]$ be a $\big(\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y}), \mathcal{B}(\mathbb{R}^+)\big)$ measurable function and $A : \mathbb{X} \rightrightarrows \mathbb{Y}$ a multifunction. Then

(i) $(x, \mu) \to \int_{\mathbb{Y}} f(x, y)\mu(\mathrm{d}y)$ is a $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$-measurable map from $\mathbb{X} \times M_1(\mathbb{Y})$ into $[0, \infty]$. Moreover, the Borel measurability of f implies that the map is Borel measurable.

(ii) If $\mathrm{Graph}(A) \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$ then $\mathcal{P}_A \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$.

(iii) If $\mathrm{Graph}(A) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y})$ then $\mathcal{P}_A \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$.

(iv) If $\mathrm{Graph}(A) \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$ then[3]

$$\mathcal{P}_{A,S} := \{(x, \mu) \in \mathbb{X} \times M_1(\mathbb{Y}) : \mu(A_x) = 1, \mathrm{supp}(\mu|A_x) = A_x\}$$

is a set in $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$.

Observe that $A_x \in \mathcal{B}(\mathbb{Y})$ and $A_x \in \mathcal{U}(\mathbb{Y})$ if $\mathrm{Graph}(A) \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$ and $\mathrm{Graph}(A) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y})$, respectively. Hence the sets $\mathcal{P}_A$, $\mathcal{P}_{A,S}$ are defined correctly. Observe also that we miss an analogue of (iv) when $\mathrm{Graph}(A) \in \mathcal{A}(\mathbb{X} \times \mathbb{Y})$.

P r o o f. (i) Assume first that $f = I_{U \times B}$ where $U \in \mathcal{U}(\mathbb{X})$, $B \in \mathcal{B}(\mathbb{Y})$. Then $\int_{\mathbb{Y}} f(x, y)\, \mu(\mathrm{d}y) = \mu(B) I_U(x)$ for $x \in \mathbb{X}$ and (i) follows easily observing that $\mu \to \mu(B)$ is a Borel measurable map $M_1(\mathbb{Y}) \to \mathbb{R}$. Theorem I.2.20 in [5] now extends the validity of (i) to f's that are bounded and $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$-measurable, which in fact verifies (i) generally. The "moreover part" of (i) may be proved in a similar way.

(ii) is an immediate consequence of (i) putting $f(x, y) = I_{A_x}(y)$.

(iii) Because $\mathrm{Graph}(A)$ is universally measurable in $\mathbb{X} \times \mathbb{Y}$ it follows that

$$\mu(A_x) = (\varepsilon_x \otimes \mu)\big(\mathrm{Graph}(A)\big) \text{ for } x \in \mathbb{X},$$

[3] As usual if $\mu \in M_1(\mathbb{Y})$ and $A \in \mathcal{U}(\mathbb{Y})$, $(\mu|A)$ denotes the restriction of μ to the Borel σ-algebra $\mathcal{B}(A)$, hence $\mathrm{supp}(\mu|A) \in \mathcal{F}(A)$ is the set defined equivalently by $\mathrm{supp}(\mu|A) = \{y \in A : \mu(G \cap A) > 0\ \forall\, G \in \mathcal{G}(\mathbb{Y}),\ y \in G\}$.

where ε_x denotes the probability measure that degenerates at x, hence

$$\mathcal{P}_A = \{(x,\mu) : (\varepsilon_x \otimes \mu)\big(\mathrm{Graph}(A)\big) = 1\}.$$

Thus, $\mathcal{P}_A$ is seen to be inverse image of $M_1^*\big(\mathrm{Graph}(A)\big)$ with respect to the continuous map $(x,\mu) \to (\varepsilon_x \otimes \mu)$ that maps $\mathbb{X} \times M_1(\mathbb{Y})$ into $M_1(\mathbb{X} \times \mathbb{Y})$. Because $M_1^*\big(\mathrm{Graph}(A)\big)$ is an analytic set in $M_1(\mathbb{X} \times \mathbb{Y})$ by Theorem 7, p. 385 in [6][4], (iii) follows directly by 8.2.6. in [3].
(iv) According to (iii) we have to prove that $\mathcal{P}_S := \{(x,\mu) : \mathrm{supp}(\mu|A_x) = A_x\}$ is a set in $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$. To see that we write $\mathcal{P}_S$ as the intersection of the sets

$$\Big[\Big(\{x : G \cap A_x \neq \emptyset\} \times M_1(\mathbb{Y}) \cap \{(x,\mu) : \mu(G \cap A_x) > 0\}\Big) \cup \Big(\{x : G \cap A_x = \emptyset\} \times M_1(\mathbb{Y})\Big)\Big]$$

where the G's are running through a countable topological base in $\mathbb{Y}$. To verify the above equality observe that

$$\mathrm{supp}(\mu|A_x) = A_x \ \text{ iff } \ [G \cap A_x \neq \emptyset, G \in \mathcal{V} \Rightarrow \mu(G \cap A_x) > 0], \ x \in \mathbb{X}.$$

To complete the proof apply (i) to see that

$$\{(x,\mu) : \mu(G \cap A_x) > 0\} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$$

and (iv) in Lemma 1 to see that $\{x : G \cap A_x \neq \emptyset\}$ and $\{x : G \cap A_x = \emptyset\}$ are sets in $\mathcal{U}(\mathbb{X})$. □

Lemma 5. Let (ξ,η) be a maximally supported $(\mathcal{P},\lambda)$-vector for a $\mathcal{P} \subset \mathbb{X} \times M_1(\mathbb{Y})$ and $\lambda \in M_1(\mathbb{X})$. Then

$$\mathrm{supp}\big(\mathcal{L}(\xi,\eta)\big) \supset \mathrm{supp}\big(\mathcal{L}(\xi',\eta')\big) \ \text{ for any } (\mathcal{P},\lambda)\text{-vector } (\xi',\eta').$$

Proof. Denote $\mathsf{P}^x = \mathcal{L}(\eta|\xi = x)$ and $\mathsf{Q}^x = \mathcal{L}(\eta'|\xi' = x)$. It follows by Remark 1 in Section 2 that $\int_{\mathbb{X}} \mathsf{P}^x\big[(\mathrm{supp}\mathsf{P}^\lambda)_x\big]\lambda(\mathrm{d}x) = \mathsf{P}^\lambda[\mathrm{supp}\mathsf{P}^\lambda] = 1$. Hence the sections $(\mathrm{supp}\mathsf{P}^\lambda)_x \in \mathcal{F}(\mathbb{Y})$ are such that $\mathsf{P}^x\big[(\mathrm{supp}\mathsf{P}^\lambda)_x\big] = 1$ almost surely w.r.t. λ and therefore $(\mathrm{supp}\mathsf{P}^\lambda)_x \supset \mathrm{supp}(\mathsf{P}^x)$. Observe that the latter inclusion and Remark 1 imply that

$$\begin{aligned} \mathsf{Q}^\lambda(\mathrm{supp}\mathsf{P}^\lambda) &= \int_{\mathbb{X}} \mathsf{Q}^x\big[(\mathrm{supp}\mathsf{P}^\lambda)_x\big]\lambda(\mathrm{d}x) \geq \int_{\mathbb{X}} \mathsf{Q}^x\big[\mathrm{supp}\mathsf{P}^x\big]\lambda(\mathrm{d}x) \\ &\geq \int_{\mathbb{X}} \mathsf{Q}^x\big[\mathrm{supp}\mathsf{Q}^x\big]\lambda(\mathrm{d}x) = 1 \end{aligned}$$

because $\mathrm{supp}\mathsf{P}^x \supset \mathrm{supp}\mathsf{Q}^x$ a.s $[\lambda]$. Thus $\mathrm{supp}\mathsf{P}^\lambda \supset \mathrm{supp}\mathsf{Q}^\lambda$ which, according to Remark 2, concludes the proof. □

It might be of some interest to note that the reverse implication to that of presented by Lemma 5 is not true: put $\mathsf{Q}^x = \varepsilon_x$ for $x \in [0,1]$ and $\mathsf{P}^x = \varepsilon_x$ for $x \in [0,1)$,

[4] The theorem states exactly that $M_1\big(\mathrm{Graph}(A)\big) \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$, but $M_1^*\big(\mathrm{Graph}(A)\big)$ is easily seen to be the image of the former set w.r.t. the continuous map $\lambda \to 1_{\mathrm{Graph}(A)} \circ \lambda$ where $1_{\mathrm{Graph}(A)} : \mathrm{Graph}(A) \to \mathbb{X} \times \mathbb{Y}$ is the identity map. Hence $M_1^*\big(\mathrm{Graph}(A)\big) \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$.

$\mathsf{P}^1 = \varepsilon_0$ and $\lambda = \frac{1}{2}(m + \varepsilon_1)$ where m is Lebesgue measure on $[0,1]$. Obviously, we have

$$\mathrm{supp}(\mathsf{Q}^\lambda) = \mathrm{Diag}\big([0,1]^2\big), \mathrm{supp}(\mathsf{P}^\lambda) = \mathrm{Diag}\big([0,1]^2\big) \cup \{(1,0)\}$$

hence

$$\mathrm{supp}(\mathsf{Q}^\lambda) \subset \mathrm{supp}(\mathsf{P}^\lambda),\ \mathrm{supp}(\mathsf{P}^1) = 0 \ \text{ and } \ \mathrm{supp}(Q^1) = \{1\}.$$

Putting $\mathcal{P} = \mathrm{Graph}(x \to \mathsf{P}^x) \cup \mathrm{Graph}(x \to \mathsf{Q}^x)$, $\mathcal{L}(\eta|\xi = x) = \mathsf{P}^x$, $\mathcal{L}(\eta'|\xi' = x) = \mathsf{Q}^x$, $\mathcal{L}(\xi) = \mathcal{L}(\xi') = \lambda$ we observe that the (ξ,η) is a $(\mathcal{P},\lambda)$-vector which distribution has the maximal support but it is not maximally supported.

We are prepared to complete our proofs.

Proof of Theorem 1. Put $\mathcal{P} := \mathcal{Q} \cap \mathcal{P}_A$. It follows by Lemma 4 (iii) and (ii) that either $\mathcal{P} \in \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$ or $\mathcal{P} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$ which in both cases implies that $D(\mathcal{Q}, A) = \mathrm{pr}_{\mathbb{X}}\mathcal{P} \in \mathcal{U}(\mathbb{X})$ (8.4.1., 8.2.6. and 8.4.4. in [3]). The cross section theorem (either 8.5.3.(b) or 8.5.4.(b) in [3]) verifies that there is a map $x \to \mathsf{P}^x$ from $D(\mathcal{Q}, A)$ into $M_1(\mathbb{Y})$ which is measurable w.r.t. the σ-algebras $\mathcal{U}(\mathbb{X}) \cap D(\mathcal{Q}, A)$ and $\mathcal{B}\big(M_1(\mathbb{Y})\big)$ such that $\mathsf{P}^x \in \mathcal{P}_x$ holds on $D(\mathcal{Q}, A)$, i.e. λ-almost surely. The map $x \to \mathsf{P}^x$ can be obviously extended (e.g. by any constant) to an universally measurable Markov kernel $x \to \mathsf{P}^x$ from $\mathbb{X}$ into $M_1(\mathbb{Y})$ and according to Remark (3) in Section 2 there exists a $(\mathbb{X} \times \mathbb{Y})$-valued vector (ξ, η) such that (8) holds. This of course means that the (ξ,η) is an $(\mathcal{Q} \cap \mathcal{P}_A, \lambda)$-vector. □

Proof of Theorem 2. Put $\mathcal{Q} := \mathcal{P} \cap \mathcal{P}_S$, where $\mathcal{P}_S := \{(x,\mu) \in \mathbb{X} \times M_1(\mathbb{Y}) : \mathrm{supp}(\mu) = (\mathrm{Output}\,\mathcal{P})_x\}$. Because $(\mathrm{Output}\,\mathcal{P})_x \in \mathcal{F}(\mathbb{Y})$ for each $x \in \mathbb{X}$ according to Lemma 3, we may apply Lemma 4 (iv) with $A = \{(\mathrm{Output}\,\mathcal{P})_x, x \in \mathbb{X}\}$ to verify that $\mathcal{P}_S \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$. Hence $\mathcal{Q}$ belongs to the σ-algebra also and Theorem 1, applied to the $\mathcal{Q}$ and to the CVM A with $\mathrm{Graph}(A) = \mathbb{X} \times \mathbb{Y}$, implies that there is a $(\mathcal{Q},\lambda)$-vector (ξ,η) because $\mathrm{pr}_{\mathbb{X}}\mathcal{Q} = \mathrm{pr}_{\mathbb{X}}\mathcal{P}$ according to Lemma 3 again. Hence, the (ξ,η) is a $(\mathcal{P},\lambda)$-vector such that (10) holds. □

Proof of Theorem 3. We plan to apply Theorem 2 to $\mathcal{P} = \mathcal{R} \cap \mathcal{P}_A$, where $\mathcal{P}_A$ and hence also $\mathcal{P}$ belong to $\mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big) \cap \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$ according to Lemma 4 (ii) and (iii). It is obvious that $\mathcal{P}$ satisfies the CS-condition and therefore $\mathrm{Output}\,\mathcal{P}$ is in $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$ according to Lemma 3 and Lemma 2 (iii). Because $D(\mathcal{R}, A) = \mathrm{pr}_{\mathbb{X}}\mathcal{P}$, it follows by Theorem 2 that there is a $(\mathcal{P},\lambda)$-vector (ξ,η) such that (10) holds. It follows directly from the definition of the set $\mathrm{Output}\,\mathcal{P}$ that the (ξ,η) is a maximally supported $(\mathcal{R} \cap \mathcal{P}_A, \lambda)$-vector. □

4. COROLLARIES

Using Theorem 1 and 3 we are able to generalize Corollary 1 in [7], namely to remove the requirement on the local compactness of the space $\mathbb{Y}$.

Corollary 1. Assume that $f_i(x,y), c_i(x)$ satisfy (5) for $i \in I$, I being an at most countable set. Consider $A \in \mathcal{A}(\mathbb{X}\times\mathbb{Y}) \cup \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$ and put

$$D(f,c,A) := \left\{ x \in \mathbb{X} : \exists\, \mu \in M_1(\mathbb{Y}), \mu(A_x) = 1, \int_{\mathbb{Y}} f_i(x,y)\, \mu(\mathrm{d}y) = c_i(x), i \in I \right\}.$$

Then to each $\lambda \in M_1^*\big(D(f,c,A)\big)$ such that $c_i \in L_1(\lambda)$ for $i \in I$ there exists an $(\mathbb{X}\times\mathbb{Y})$-valued random vector (ξ,η) for which

$$\mathcal{L}(\xi) = \lambda, \mathsf{P}[(\xi,\eta) \in A] = 1, \mathsf{E}[f_i(\xi,\eta)] < \infty, \mathsf{E}[f_i(\xi,\eta)|\xi] = c_i(\xi), i \in I \tag{13}$$

holds.

If moreover $A \in \mathcal{A}(\mathbb{X}\times\mathbb{Y}) \cap \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{Y})$ then a random vector (ξ,η) with the properties (13) may be chosen such that $\mathrm{supp}\big(\mathcal{L}(\xi,\eta)\big) \supset \mathrm{supp}\big(\mathcal{L}(\xi',\eta')\big)$ for any other random vector (ξ',η') that satisfies (13).

P r o o f . Put $\mathcal{Q} = \big\{(x,\mu) \in \mathbb{X}\times M_1(\mathbb{Y}) : \int_{\mathbb{Y}} f_i(x,y)\,\mu(\mathrm{d}y) = c_i(x), i \in I\big\}$ and consider the multifunction $B : \mathbb{X} \rightrightarrows \mathbb{Y}$ with $\mathrm{Graph}(B) = A$. Then, using the notation introduced in Theorem 1, we have $D(f,c,A) = D(\mathcal{Q},B)$ and $\mathcal{Q} \in \mathcal{B}\big(\mathbb{X}\times M_1(\mathbb{Y})\big)$ according to Lemma 4 (i). Observe also, that for a random vector (ξ,η), the properties (13) state equivalently that the (ξ,η) is a $(\mathcal{Q}\cap\mathcal{P}_B,\lambda)$-vector. The equivalence is an easy consequence of Remark 2 and 4 in Section 2 using the integrability of c_i's with respect to λ. Because the set $\mathcal{Q}$ satisfies obviously the CS-condition, Theorem 1 and Theorem 3 verify the statements of our Corollary. □

Remark that for a finite index set I

$$D(f,c,A) = \{x \in \mathbb{X} : (x) \in \mathrm{co}(\mathbf{f}(x,A_x))\},\ \mathbf{c} = (c_i, i \in I), \mathbf{f} = (f_i, i \in I),$$

where co denotes the convex hull (see [4], for example).

The theory we have presented is designed mostly with the purpose to prove the existence of a $(\mathcal{P},\lambda)$-vector with the maximal support of its probability distribution. The rest of our corollaries suggests some other possible applications.

Corollary 2. Consider a set $\mathcal{P} \in \mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big)$ and an upper bounded function $F : \mathbb{X}\times M_1(\mathbb{Y}) \to \mathbb{R}$ that is $\mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big)$-measurable. Denote

$$S_F(x) := \sup\{F(x,\mu), \mu \in \mathcal{P}_x\} \text{ for } x \in \mathbb{X} \text{ (i.e. } S_F(x) = -\infty \text{ for } x \notin \mathrm{pr}_{\mathbb{X}}(\mathcal{P}))$$
$$D(\mathcal{P},F) := \{x \in \mathbb{X} : S_F(x) = F(x,\mu) \text{ for some } \mu \in \mathcal{P}_x\}.$$

Consider moreover a measure $\lambda \in M_1^*\big(D(\mathcal{P},F)\big)$. Then there exists a $(\mathcal{P},\lambda)$-vector (ξ,η) such that

$$F\big(x,\mathcal{L}(\eta|\xi=x)\big) = S_F(x) \text{ holds } \lambda\text{-almost surely.} \tag{14}$$

P r o o f . Obviously, the random vector (ξ,η) which existence is stated is equivalently defined as a $(\mathcal{Q},\lambda)$-vector, where

$$\mathcal{Q} := \mathcal{P} \cap \mathcal{S}_F, \text{ where } \mathcal{S}_F := \{(x,\mu) : F(x,\mu) = S_F(x)\}.$$

Because $\mathrm{pr}_{\mathbb{X}}\mathcal{Q} = D(\mathcal{P}, F)$, we could use Theorem 1 (with $A : \mathbb{X} \rightrightarrows \mathbb{Y}$, such that $\mathrm{Graph}(A) = \mathbb{X} \times \mathbb{Y}$) to prove the existence of a $(\mathcal{Q}, \lambda)$-vector (ξ, η) if $\mathcal{S}_F$ would be a set in $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$. To verify this, it is sufficient to show that the function $S_F : \mathbb{X} \to [-\infty, +\infty)$ is universally measurable: Fix $a \in \mathbb{R}$ and observe that

$$\{x : S_F(x) > a\} = \{x : \exists\, \mu \in \mathcal{P}_x, F(x,\mu) > a\} = \mathrm{pr}_{\mathbb{X}}(\mathcal{P} \cap [F > a]),$$

where $[F > a] = \{(x,\mu) : F(x,\mu) > a\}$. Thus $\{x : S_F(x) > a\}$ is the projection of a set in $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$ and therefore a universally measurable set in $\mathbb{X}$ according to 8.4.4. in [3]. □

An obvious choice for the function $F(x,\mu)$ is given by

$$F(x,\mu) := \int_{\mathbb{Y}} f(x,y)\, \mu(\mathrm{d}y),\ x \in \mathbb{X},\ \ \mu \in M_1(\mathbb{Y}),$$

where $f : \mathbb{X}\times\mathbb{Y} \to \mathbb{R}$ is an upper bounded $\mathcal{U}(\mathbb{X})\otimes\mathcal{B}(\mathbb{Y})$-measurable function. A more sophisticated choice of the F allows to enrich the result given by Theorem 3 in [7]: For a $\mathcal{P} \subset \mathbb{X} \times M_1(\mathbb{Y})$ such that all its sections $\mathcal{P}_x$ are convex sets we denote $\mathcal{P}^e := \{(x,\mu) \in \mathcal{P} : \mu \in \mathrm{ex}\mathcal{P}_{\mathrm{x}}\}$ where $\mathrm{ex}\mathcal{P}_x$ denotes as usual the set of all extremal measures in $\mathcal{P}_x$ (might be an empty set). Theorem 4 in [7] states the existence of a $(\mathcal{P}^e, \lambda)$-vector (ξ, η) (i.e. $\mathcal{L}(\eta|\xi = x)$ is an extremal measure in $\mathcal{P}_x$ λ-almost surely), provided that the $\mathcal{P}$ is a closed set in $\mathbb{X} \times M_1(\mathbb{Y})$ and $\lambda \in M_1^*(\mathrm{pr}_{\mathbb{X}}(\mathcal{P}))$.

Corollary 3. Let $\mathcal{P} \in \mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big)$ is a set such that $\mathcal{P}_x$ is a compact convex set in $M_1(\mathbb{Y})$ for all $x \in \mathbb{X}$ and λ a measure in $M_1^*(\mathrm{pr}_{\mathbb{X}}(\mathcal{P}))$. Then there exists a $(\mathcal{P}, \lambda)$-vector (ξ, η) such that $\mathcal{L}(\eta|\xi = x) \in \mathrm{ex}\mathcal{P}_x$ λ-almost surely.

Proof. It is a well known fact that there exists a bounded continuous strictly convex function $A : M_1(\mathbb{Y}) \to \mathbb{R}$. For its construction we may refer to [8] (p.40) or simply suggest to put $A(\mu) := \sum_{n=1}^{\infty} 2^{-n} \left(\int_{\mathbb{Y}} f_n \mathrm{d}\mu\right)^2, \mu \in M_1(\mathbb{Y})$, where $0 \le f_n \le 1$ are continuous functions defined on $\mathbb{Y}$ such that $\int_{\mathbb{Y}} f_n d\mu = \int_{\mathbb{Y}} f_n \mathrm{d}\nu,\ n \in \mathbb{N}$ implies that $\mu = \nu$ for $\mu, \nu \in M_1(\mathbb{Y})$. Applying Corollary 2 to the continuous bounded function

$$F : \mathbb{X} \times M_1(\mathbb{Y}) \to \mathbb{R} \text{ defined by } F(x,\mu) = A(\mu) \text{ for } (x,\mu) \in \mathbb{X} \times M_1(\mathbb{Y}),$$

observing that $D(\mathcal{P}, F) = \mathrm{pr}_{\mathbb{X}}(\mathcal{P})$ in this case ($F(x,\cdot)$'s are continuous on compacts $\mathcal{P}_x$'s) we prove the existence of a $(\mathcal{P}, \lambda)$-vector (ξ, η) that possesses the property (14). It means that $A\big(\mathcal{L}(\eta|\xi = x)\big) = \max\{A(\mu) : \mu \in \mathcal{P}_x\}$ λ-almost surely, hence $\mathcal{L}(\eta|\xi = x) \in \mathrm{ex}\mathcal{P}_x$ λ-almost surely because A is a strictly convex function. □

Observe that Corollary 3 may be applied to a set $\mathcal{P}$ defined by

$$\mathcal{P} = \left\{(x,\mu) \in \mathbb{X} \times M_1(\mathbb{Y}) : \int_{\mathbb{Y}} f_i(x,y)\mu(\mathrm{d}y) = c_i(x), i \in \mathbb{N}\right\},$$

where $\mathbb{Y}$ is a *compact* metric space and $f_i : \mathbb{X}\times\mathbb{Y} \to [0,\infty)$, $c_i : \mathbb{X} \to [0,\infty]$ are Borel measurable such that $f_i(x,\cdot)$ is a bounded continuous for each $x \in \mathbb{X}$.

We shall close our presentation by a simple observation on the existence of $(\mathcal{P},\lambda)$-vectors (ξ,η) with the $\mathcal{L}(\eta|\xi = x)$'s that are absolutely continuous with respect to a σ-finite Borel measure on the space $\mathbb{Y}$.

Corollary 4. Let $\mathcal{P}$ is a set in $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$ and m a σ-finite Borel measure on $\mathbb{Y}$. Denote

$$D(\mathcal{P},m) := \{x \in \mathbb{X} : \exists\, \mu \in \mathcal{P}_x, \mu \ll m\}$$

and consider $\lambda \in M_1^*\big(D(\mathcal{P},m)\big)$. Then there exists a $(\mathcal{P},\lambda)$-vector (ξ,η) such that

$$\mathcal{L}(\eta|\xi = x) \ll m \ [\lambda] \text{ a.s. or equivalently } \mathcal{L}(\xi,\eta) \ll \lambda \otimes m. \tag{15}$$

If $\mathcal{P} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big) \cap \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$ satisfies moreover the CS-condition then there is a $(\mathcal{P},\lambda)$-vector such that (15) holds and such that

$$\operatorname{supp}\big(\mathcal{L}(\xi,\eta)\big) \supset \operatorname{supp}\big(\mathcal{L}(\xi',\eta')\big)\ \forall\ (\mathcal{P},\lambda)\text{-vector } (\xi',\eta') \text{ with the property (15).}$$

Proof. We shall use Theorem 1 and Theorem 3 with $\mathcal{Q} = \mathcal{P} \cap \mathcal{A}_m$ and $\mathcal{R} = \mathcal{P}\cap\mathcal{A}_m$, respectively and also with $A : \mathbb{X} \rightrightarrows \mathbb{Y}$ such that $\operatorname{Graph}(A) = \mathbb{X}\times\mathbb{Y}$, denoting $\mathcal{A}_m := \{(x,\mu) \in \mathbb{X} \times M_1(\mathbb{Y}) : \mu \ll m\}$. Observe that $D(\mathcal{P} \cap \mathcal{A}_m, A) = D(\mathcal{P},m) = \operatorname{pr}_{\mathbb{X}}(\mathcal{P} \cap \mathcal{A}_m)$ in this case. We state that $\mathcal{A}_m$ is a Borel set in $\mathbb{X} \times M_1(\mathbb{Y})$: Observe first that $Z = \big\{f \in L_1(m) : f \geq 0 \ \ m\text{-almost everywhere}, \int_{\mathbb{Y}} f \,\mathrm{d}m = 1\big\}$ is a closed, hence a Borel set in $L_1(m)$ that is a Polish space in its standard norm topology. Putting $H(f) = m_f$, where $f \in L_1(m)$ and $\mathrm{d}m_f = f\,\mathrm{d}m$, it follows easily that $H : Z \to M_1(\mathbb{Y})$ is a continuous injective map such that $\mathcal{A}_m = \mathbb{X} \times H(Z)$. Hence, $\mathcal{A}_m \in \mathcal{B}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$ according to 8.3.7. in [3].

Thus, $\mathcal{P} \cap \mathcal{A}_m$ is a set that satisfies the measurability requirement of Theorem 1 if $\mathcal{P} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$ and that of Theorem 3 if $\mathcal{P} \in \mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big) \cap \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$. Moreover, the set $\mathcal{P} \cap \mathcal{A}_m$ obviously satisfies the CS-condition if the set $\mathcal{P}$ does. Hence, for a $\mathcal{P}$ in $\mathcal{U}(\mathbb{X}) \otimes \mathcal{B}\big(M_1(\mathbb{Y})\big)$ there exists a $(\mathcal{P} \cap \mathcal{A}_m,\lambda)$-vector (ξ,η) according to Theorem 1 and for $\mathcal{P} \in \mathcal{U}(\mathbb{X})\otimes\mathcal{B}\big(M_1(\mathbb{Y})\big) \cap \mathcal{A}\big(\mathbb{X} \times M_1(\mathbb{Y})\big)$ there exists a maximally supported $(\mathcal{P} \cap \mathcal{A}_m,\lambda)$-vector (ξ,η) according to Theorem 3 which concludes the proof because

$$(\xi,\eta) \text{ is an } (\mathcal{A}_m,\lambda)\text{-vector iff } \mathcal{L}(\xi,\eta) \ll \lambda \otimes m$$

according to Remark 1 in Section 2. □

(Received August 5, 1997.)

REFERENCES

[1] J.-P. Aubin and H. Frankowska: Set Valued Analysis. Birkhäuser, Boston 1990.
[2] V. Beneš and J. Štěpán (eds.): Distributions with Given Marginals and Moment Problems. Kluwer, Dordrecht 1997.

[3] D. L. Cohn: Measure Theory. Birkhäuser, Boston 1980.
[4] J. H. B. Kempermann: The general moment problem, a geometric approach. Ann. Math. Statist. *39* (1968), 93–122.
[5] P. A. Meyer: Probability and Potentials. Blaisdell, Waltham 1966.
[6] L. Schwarz: Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures. Oxford University Press, Oxford 1973.
[7] J. Štěpán: How to construct two dimensional random vector with given marginal structure. In: Distributions with Given Marginals and Moment Problems (V. Beneš and J. Štěpán, eds.), Kluwer, Dordrecht 1997, pp. 161–171,
[8] G. Winkler: Choquet Order and Simplices. (Lectures Notes in Mathematics 1145.) Springer–Verlag, Berlin 1985.

Prof. RNDr. Josef Štěpán, DrSc. and Mgr. Daniel Hlubinka, Ph.D., Department of Probability and Statistics, Faculty of Mathematics and Physics – Charles University, Sokolovská 83, 186 00 Praha 8. Czech Republic.
e-mails: stepan@karlin.mff.cuni.cz, hlubinka@karlin.mff.cuni.cz