GLUING LEMMAS AND SKOROHOD REPRESENTATIONS

PATRIZIA BERTI, LUCA PRATELLI, AND PIETRO RIGO

Abstract. Let $(\mathcal{X}, \mathcal{E})$, $(\mathcal{Y}, \mathcal{F})$ and $(\mathcal{Z}, \mathcal{G})$ be measurable spaces. Suppose we are given two probability measures γ and τ, with γ defined on $(\mathcal{X} \times \mathcal{Y},\, \mathcal{E} \otimes \mathcal{F})$ and τ on $(\mathcal{X} \times \mathcal{Z},\, \mathcal{E} \otimes \mathcal{G})$. Conditions for the existence of random variables X, Y, Z, defined on the same probability space $(\Omega, \mathcal{A}, P)$ and satisfying

$$(X, Y) \sim \gamma \quad \text{and} \quad (X, Z) \sim \tau,$$

are given. The probability P may be finitely additive or σ-additive. As an application, a version of Skorohod representation theorem is proved. Such a version does not require separability of the limit probability law, and answers (in a finitely additive setting) a question raised in [2] and [4].

1. Introduction and motivations

This paper is split into two parts. The first focuses on gluing lemmas, while the second deals with Skorohod representation theorem. The second part is the natural continuation of some previous papers (see [1]-[4]) and the main reason for investigating gluing lemmas.

In the sequel, a gluing lemma is meant as follows. Let $(\mathcal{X}, \mathcal{E})$, $(\mathcal{Y}, \mathcal{F})$ and $(\mathcal{Z}, \mathcal{G})$ be measurable spaces. Suppose we are given two probability measures γ and τ, with γ defined on $(\mathcal{X} \times \mathcal{Y},\, \mathcal{E} \otimes \mathcal{F})$ and τ on $(\mathcal{X} \times \mathcal{Z},\, \mathcal{E} \otimes \mathcal{G})$. A gluing lemma gives conditions for the existence of three random variables X, Y, Z defined on the same probability space and satisfying

$$(X, Y) \sim \gamma \quad \text{and} \quad (X, Z) \sim \tau.$$

Without loss of generality, we shall assume that X, Y, Z are the coordinate projections

$$X(x, y, z) = x, \quad Y(x, y, z) = y, \quad Z(x, y, z) = z,$$

where $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Under this convention, the question reduces to whether there is a probability measure P on the product σ-field $\mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}$ such that

$$P\big[(X, Y) \in A\big] = \gamma(A) \quad \text{and} \quad P\big[(X, Z) \in B\big] = \tau(B) \tag{1}$$

whenever $A \in \mathcal{E} \otimes \mathcal{F}$ and $B \in \mathcal{E} \otimes \mathcal{G}$.

Gluing lemmas occur in various frameworks, mainly in connection with optimal transport, coupling and related topics; see e.g. [19]. Another application of gluing lemmas, as discussed below, concerns Skorohod representation theorem.

We also note that gluing lemmas, as defined in this paper, are connected to *transfer* results in the sense of [10, Theorem 6.10] and [16, pages 135-136 and 152-153]. Indeed, gluing and transfer lemmas are complementary, even if technically different, and results concerning one of the two fields might be useful in the other.

2010 *Mathematics Subject Classification.* 60B10, 60A05, 60A10.

Key words and phrases. Disintegration – Finitely additive probability – Gluing lemma – Separable probability measure – Skorohod representation theorem – Wasserstein distance.

The transfer idea has been around in some form for a long time, but it seems to have been first formalized in Thorisson's thesis; see [10, page 573] and [16, page 482].

An obvious necessary condition for (1) is

$$\gamma\big(A\times\mathcal{Y}\big)=\tau\big(A\times\mathcal{Z}\big)\quad\text{for all }A\in\mathcal{E}. \tag{2}$$

In this paper, it is shown that condition (2) is not enough for (1), even if $(\mathcal{X},\mathcal{E})=(\mathcal{Y},\mathcal{F})=(\mathcal{Z},\mathcal{G})$ with $\mathcal{X}$ separable metric and $\mathcal{E}$ the Borel σ-field. However, condition (2) suffices for (1) under some extra assumption. For instance, (2) implies (1) if one of γ and τ is disintegrable, or else if all but perhaps one of the marginals of γ and τ are perfect. See Example 1, Lemma 4 and Corollary 5.

In dealing gluing lemmas, one naturally comes across with *finitely additive probabilities*. We substantiate this claim with two results; see Lemma 2. Suppose the probability P involved in condition (1) is only requested to be finitely additive. Then, (1) admits a simple characterization. Indeed, (1) holds if and only if

$$\gamma_*\big(A\times\mathcal{Y}\big)\leq\tau^*\big(A\times\mathcal{Z}\big)\quad\text{for all }A\subset\mathcal{X}$$

where γ_* and τ^* are the inner and outer measures induced by γ and τ. Next, suppose $\mathcal{X}$ and $\mathcal{Z}$ are topological spaces (equipped with the Borel σ-fields $\mathcal{E}$ and $\mathcal{G}$). Then, (2) suffices for (1) provided B is restricted to be a continuity set for τ, in the sense that $B\in\mathcal{E}\otimes\mathcal{G}$ and $\tau^*(\partial B)=0$.

We next turn to Skorohod representation theorem (SRT). In addition to [1]-[4], related references are [9], [13], [15], [17].

Let S be a metric space, $\mathcal{B}$ the Borel σ-field on S, and $(\mu_n:n\geq 0)$ a sequence of probability measures on $\mathcal{B}$. Recall that a law μ on $\mathcal{B}$ is *separable* if $\mu(A)=1$ for some separable set $A\in\mathcal{B}$. According to SRT, if $\mu_n\to\mu_0$ weakly and μ_0 is separable, on some probability space there are S-valued random variables $(X_n:n\geq 0)$ such that $X_n\sim\mu_n$ for all $n\geq 0$ and $X_n\xrightarrow{a.s.}X_0$. See [7], [14] and [20]; see also [8, page 130] and [18, page 77] for historical notes.

In [2] and [4], the separability assumption on μ_0 is investigated. Suppose

$$d:S\times S\to[0,\infty)\ \text{ is measurable with respect to }\ \mathcal{B}\otimes\mathcal{B},$$

where d is the distance on S. Also, say that the sequence (μ_n) admits a *Skorohod representation* if

> On some probability space, there are S-valued random variables X_n such that $X_n\sim\mu_n$ for all $n\geq 0$ and $X_n\to X_0$ in probability.

If non separable laws on $\mathcal{B}$ actually exist, then:

- (i) It may be that $\mu_n\to\mu_0$ weakly but (μ_n) fails to admit a Skorohod representation. See [2, Example 4.1].
- (ii) It may be that (μ_n) has a Skorohod representation, but no S-valued random variables Y_n satisfy $Y_n\sim\mu_n$ for all $n\geq 0$ and $Y_n\xrightarrow{a.s.}Y_0$. See [2, Example 4.2].
- (iii) (μ_n) admits a Skorohod representation, for arbitrary μ_0, if the condition $\mu_n\to\mu_0$ weakly is suitably strengthened. See [2, Theorem 4.2] and [4, Theorems 1.1 and 1.2].

Hence, separability of μ_0 can not be dropped from SRT (by (i)) and almost sure convergence is too much (by (ii)). On the other hand, because of (iii), a possible *conjecture* is

$$(\mu_n) \text{ has a Skorohod representation} \quad \Leftrightarrow \quad \lim_n \rho(\mu_n, \mu_0) = 0$$

where ρ is some discrepancy measure between probability laws. If true for a reasonable ρ, such a conjecture would be a (nice) version of SRT not requesting separability of μ_0.

Two common choices of ρ are $\rho = L$ and $\rho = W$, where L is the bounded-Lipschitz-metric and W the Wasserstein distance. The definition of L is recalled in Subsection 3.1. As to W, if μ and ν are any probability measures on $\mathcal{B}$, then

$$W(\mu, \nu) = \inf_\gamma \int 1 \wedge d(x, y)\, \gamma(dx, dy)$$

where inf is over those probability measures γ on $\mathcal{B} \otimes \mathcal{B}$ with marginals μ and ν.

It is not hard to prove that $\lim_n W(\mu_n, \mu_0) = 0$ if (μ_n) has a Skorohod representation. Thus, $\rho = W$ is an admissible choice. Also, since $L \leq 2\,W$, if the conjecture works with $\rho = L$ then it works with $\rho = W$ as well. Accordingly, we let $\rho = W$.

Suppose $\lim_n W(\mu_n, \mu_0) = 0$. By definition, there is a sequence $(\gamma_n : n \geq 1)$ of probability measures on $\mathcal{B} \otimes \mathcal{B}$ such that each γ_n has marginals μ_0 and μ_n and

$$\lim_n \gamma_n\big\{(x, y) : d(x, y) > \epsilon\big\} = 0 \quad \text{for all } \epsilon > 0.$$

Thus, *one automatically obtains a Skorohod representation for* (μ_n) *if, on some probability space, there are* S*-valued random variables* $(X_n : n \geq 0)$ *such that*

$$(X_0, X_n) \sim \gamma_n \quad \text{for all } n \geq 1. \tag{3}$$

This is exactly the point where gluing lemmas come into play. Roughly speaking, they serve to paste in the γ_n in order to get condition (3). Unfortunately, Example 1 precludes to obtain (3) for an arbitrary sequence (γ_n) such that

$$\gamma_n(A \times S) = \mu_0(A) = \gamma_1(A \times S) \quad \text{for all } n \geq 1 \text{ and } A \in \mathcal{B}.$$

However, something can be said. Our main result is that $\lim_n W(\mu_n, \mu_0) = 0$ if and only if, on a finitely additive probability space $(\Omega, \mathcal{A}, P)$, there are S-valued random variable X_n such that

$$X_n \xrightarrow{P} X_0, \quad P(X_0 \in A) = \mu_0(A) \text{ for all } A \in \mathcal{B}, \quad \text{and}$$
$$P\big[(X_0, X_n) \in A\big] = \gamma_n(A) \quad \text{whenever } n \geq 1,\ A \in \mathcal{B} \otimes \mathcal{B} \text{ and } \gamma_n^*(\partial A) = 0.$$

Moreover, $P\big[(X_0, X_n) \in \cdot\big] = \gamma_n(\cdot)$ on all of $\mathcal{B} \otimes \mathcal{B}$ if μ_n is perfect.

To sum up, in a finitely additive setting, the above conjecture is true with $\rho = W$ provided $X_n \sim \mu_n$ is meant as $P(X_n \in A) = \mu_n(A)$ if $A \in \mathcal{B}$ and $\mu_n(\partial A) = 0$, or equivalently as

$$E_P\big\{f(X_n)\big\} = \int f\, d\mu_n \quad \text{for all bounded continuous } f : S \to \mathbb{R}.$$

We refer to Theorem 8 for details.

2. Gluing lemmas

In the sequel, the abbreviation "f.a.p." stands for *finitely additive probability.* A σ-additive f.a.p. is referred to as a *probability measure.*

Let $(\mathcal{X},\mathcal{E})$, $(\mathcal{Y},\mathcal{F})$, $(\mathcal{Z},\mathcal{G})$ be (arbitrary) measurable spaces, γ a f.a.p. on $\mathcal{E}\otimes\mathcal{F}$ and τ a f.a.p. on $\mathcal{E}\otimes\mathcal{G}$. Recall that, if Q is a f.a.p. on a field $\mathcal{U}$ on some set Ω, the outer and inner measures are defined by

$$Q^*(A)=\inf\{Q(B):A\subset B\in\mathcal{U}\}\quad\text{and}\quad Q_*(A)=1-Q^*(A^c)\quad\text{where } A\subset\Omega.$$

We begin with an example where condition (2) holds while condition (1) fails for any f.a.p. P, despite γ and τ are probability measures and $(\mathcal{X},\mathcal{E})=(\mathcal{Y},\mathcal{F})=(\mathcal{Z},\mathcal{G})$ with $\mathcal{X}$ separable metric and $\mathcal{E}$ the Borel σ-field.

Example 1. Let λ be the Lebesgue measure on the Borel σ-field on $[0,1]$. Take $I\subset[0,1]$ such that $\lambda^*(I)=1$ and $\lambda_*(I)=0$ and define $J=[0,1]\setminus I$ and

$$\mathcal{X}=\{(x,1):x\in[0,1]\}\cup\{(x,2):x\in I\}\cup\{(x,3):x\in J\}.$$

Then, $\mathcal{X}$ is a separable metric space under the distance

$$d[(x,r)\,,\,(y,k)]=1\ \text{ if } r\neq k\quad\text{and}\quad d[(x,r)\,,\,(y,k)]=|x-y|\ \text{ if } r=k.$$

Let $(\mathcal{X},\mathcal{E})=(\mathcal{Y},\mathcal{F})=(\mathcal{Z},\mathcal{G})$ where $\mathcal{E}$ is the Borel σ-field on $\mathcal{X}$. For each $A\in\mathcal{E}\otimes\mathcal{E}$, define also

$$\gamma(A)=\lambda^*\big\{x\in I:((x,1),(x,2))\in A\big\},$$
$$\tau(A)=\lambda^*\big\{x\in J:((x,1),(x,3))\in A\big\}.$$

Since $\lambda^*(I)=\lambda^*(J)=1$, both γ and τ are probability measures on $\mathcal{E}\otimes\mathcal{E}$. Let $B\in\mathcal{E}$ and $f(x)=(x,1)$ for $x\in[0,1]$. On noting that $f:[0,1]\to\mathcal{X}$ is Borel measurable and using $\lambda^*(I)=\lambda^*(J)=1$ again, one obtains

$$\gamma\big(B\times\mathcal{X}\big)=\lambda^*\big(I\cap\{f\in B\}\big)=\lambda(f\in B)=\lambda^*\big(J\cap\{f\in B\}\big)=\tau\big(B\times\mathcal{X}\big).$$

Hence, condition (2) holds. However, condition (1) fails for any f.a.p. P. Define in fact $h(x,r)=x$ for all $(x,r)\in\mathcal{X}$. If (1) holds for some f.a.p. P, then

$$P\big[h(X)=h(Y),\,Y\in\{(x,2):x\in I\}\big]$$
$$=\gamma\big\{((x,r)\,,\,(y,k)):x=y\in I,\,k=2\big\}=\lambda^*(I)=1.$$

Similarly, $P\big[h(X)=h(Z),\,Z\in\{(x,3):x\in J\}\big]=1$. Thus, one obtains the contradiction $P\big[h(X)\in I\cap J\big]=1$.

Because of Example 1, some condition for (1) is needed. Next lemma is actually fundamental for Theorem 8.

Lemma 2. *Let γ be a f.a.p. on $\mathcal{E}\otimes\mathcal{F}$ and τ a f.a.p. on $\mathcal{E}\otimes\mathcal{G}$. There is a f.a.p. P on $\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}$ satisfying condition* (1) *if and only if*

$$\gamma_*\big(A\times\mathcal{Y}\big)\leq\tau^*\big(A\times\mathcal{Z}\big)\quad\textit{for all } A\subset\mathcal{X}.$$

Moreover, if condition (2) *holds and $\mathcal{X}$ and $\mathcal{Z}$ are topological spaces (equipped with the Borel σ-fields $\mathcal{E}$ and $\mathcal{G}$) there is a f.a.p. P on $\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}$ such that*

$$P\big[(X,Y)\in A\big]=\gamma(A)\quad\textit{and}\quad P\big[(X,Z)\in B\big]=\tau(B)$$

whenever $A\in\mathcal{E}\otimes\mathcal{F}$, $B\in\mathcal{E}\otimes\mathcal{G}$ and $\tau^*(\partial B)=0$.

Proof. Suppose that (1) holds for some f.a.p. P. Let Q be a f.a.p. on the power set of $\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$ such that $Q=P$ on $\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}$. By definition of inner and outer measure and since Q extends P, it follows that

$$\gamma_*\big(A\times\mathcal{Y}\big)\le P_*(X\in A)\le Q(X\in A)\le P^*(X\in A)\le\tau^*\big(A\times\mathcal{Z}\big)$$

for all $A\subset\mathcal{X}$.

Conversely, suppose $\gamma_*\big(A\times\mathcal{Y}\big)\le\tau^*\big(A\times\mathcal{Z}\big)$ for all $A\subset\mathcal{X}$. We need the following result by Bhaskara Rao and Bhaskara Rao [5, Theorem 3.6.1].

(BR) For $j=1,2$, let $\mathcal{U}_j$ be a field on a set Ω and P_j a f.a.p. on $\mathcal{U}_j$. There is a f.a.p. P on the power set of Ω such that $P=P_1$ on $\mathcal{U}_1$ and $P=P_2$ on $\mathcal{U}_2$ if and only if $P_1(A)\le P_2(B)$ whenever $A\in\mathcal{U}_1$, $B\in\mathcal{U}_2$ and $A\subset B$.

Let $\Omega=\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$ and

$$\mathcal{U}_1=\big\{\{(X,Y)\in A\}:A\in\mathcal{E}\otimes\mathcal{F}\big\},\quad \mathcal{U}_2=\big\{\{(X,Z)\in B\}:B\in\mathcal{E}\otimes\mathcal{G}\big\},$$
$$P_1\big[(X,Y)\in A\big]=\gamma(A)\quad\text{and}\quad P_2\big[(X,Z)\in B\big]=\tau(B).$$

Fix $A\in\mathcal{E}\otimes\mathcal{F}$ and $B\in\mathcal{E}\otimes\mathcal{G}$ with $\{(X,Y)\in A\}\subset\{(X,Z)\in B\}$ and define

$$A_0=\{x\in\mathcal{X}:(x,y)\in A\text{ for some }y\in\mathcal{Y}\}$$

to be the projection of A onto $\mathcal{X}$. Since $A_0\times\mathcal{Z}\subset B$, then

$$P_1\big[(X,Y)\in A\big]=\gamma(A)\le\gamma_*\big(A_0\times\mathcal{Y}\big)\le\tau^*\big(A_0\times\mathcal{Z}\big)\le\tau(B)=P_2\big[(X,Z)\in B\big].$$

Therefore, in view of (BR), condition (1) holds for some f.a.p. P.

Finally, suppose condition (2) holds and $\mathcal{X}$ and $\mathcal{Z}$ are topological spaces equipped with the Borel σ-fields $\mathcal{E}$ and $\mathcal{G}$. Define the field

$$\mathcal{U}_2=\big\{\{(X,Z)\in B\}:B\in\mathcal{E}\otimes\mathcal{G},\ \tau^*(\partial B)=0\big\}$$

and take $\mathcal{U}_1$, P_1, P_2 as above. Fix $A\in\mathcal{E}\otimes\mathcal{F}$ and $B\in\mathcal{E}\otimes\mathcal{G}$ such that $\tau^*(\partial B)=0$ and $\{(X,Y)\in A\}\subset\{(X,Z)\in B\}$. Since $\overline{A_0}\in\mathcal{E}$ and $\tau^*(\partial B)=0$,

$$P_1\big[(X,Y)\in A\big]\le\gamma\big(\overline{A_0}\times\mathcal{Y}\big)=\tau\big(\overline{A_0}\times\mathcal{Z}\big)\le\tau^*(\overline{B})=\tau(B)=P_2\big[(X,Z)\in B\big].$$

An application of (BR) concludes the proof. □

Remark 3. Other statements, similar to Lemma 2, can be proved by the same argument. As an example, under condition (2), there is a f.a.p. Q on $\mathcal{E}\otimes\mathcal{F}\otimes\mathcal{G}$ such that $Q\big[X\in A,\,Y\in B\big]=\gamma(A\times B)$ and $Q\big[(X,Z)\in D\big]=\tau(D)$ whenever $A\in\mathcal{E}$, $B\in\mathcal{F}$ and $D\in\mathcal{E}\otimes\mathcal{G}$.

Usually, γ and τ are probability measures (and not merely f.a.p.'s) and a natural question is whether condition (1) holds under a (σ-additive) probability measure P. To address this issue, we first recall some definitions.

Let μ be a probability measure on $(\mathcal{X},\mathcal{E})$. Say that μ is *perfect* if, for each $\mathcal{E}$-measurable function $f:\mathcal{X}\to\mathbb{R}$, there is a Borel set $B\subset\mathbb{R}$ such that $B\subset f(\mathcal{X})$ and $\mu(f\in B)=1$. If $\mathcal{X}$ is separable metric and $\mathcal{E}$ the Borel σ-field, then μ is perfect if and only if it is tight. In particular, μ is perfect if $\mathcal{X}$ is a universally measurable subset of a Polish space (in particular, a Borel subset) and $\mathcal{E}$ the Borel σ-field.

Let γ be a probability measure on $\mathcal{E}\otimes\mathcal{F}$. Say that γ is *disintegrable* if γ admits a regular conditional distribution given the sub-σ-field $\{A\times\mathcal{Y}:A\in\mathcal{E}\}$. Equivalently, there is a collection $\{\alpha(x,\cdot):x\in\mathcal{X}\}$ such that:

– $\alpha(x,\cdot)$ is a probability measure on $\mathcal{F}$ for $x\in\mathcal{X}$;

– $x \mapsto \alpha(x, B)$ is $\mathcal{E}$-measurable for $B \in \mathcal{F}$;
– $\gamma(H) = \int \alpha(x, H_x)\, \mu(dx)$ for $H \in \mathcal{E} \otimes \mathcal{F}$, where μ is the marginal of γ on $\mathcal{E}$ and $H_x = \{y \in \mathcal{Y} : (x, y) \in H\}$.

The collection $\{\alpha(x, \cdot) : x \in \mathcal{X}\}$ is said to be a *disintegration* for γ.

A disintegration can fail to exist. However, for γ to admit a disintegration, it suffices that $\mathcal{F}$ is countably generated and the marginal of γ on $\mathcal{F}$ is perfect.

Some form of disintegrability yields condition (1) under a σ-additive P. This fact is essentially known and implicit in all existing gluing or transfer results; see e.g. [10, Section 6] and [16, pages 135-136 and 152-153]. For completeness (and since we do not know of any explicit statement) we also provide a proof.

Lemma 4. *Let γ be a probability measure on $\mathcal{E} \otimes \mathcal{F}$ and τ a probability measure on $\mathcal{E} \otimes \mathcal{G}$. If condition* (2) *holds and one of γ and τ is disintegrable, then condition* (1) *holds with P a probability measure on $\mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}$.*

Proof. Suppose γ disintegrable and take a disintegration $\{\alpha(x, \cdot) : x \in \mathcal{X}\}$ for γ. For all $H \in \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}$, define

$$P(H) = \int \alpha(x,\, H_{x,z})\, \tau(dx, dz) \quad \text{where } H_{x,z} = \{y \in \mathcal{Y} : (x, y, z) \in H\}.$$

Then, P is a probability measure on $\mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}$ and $P\big[(X, Z) \in B\big] = \tau(B)$ for all $B \in \mathcal{E} \otimes \mathcal{G}$. Because of (2), γ and τ have a common marginal on $\mathcal{E}$, say μ. Fix $A \in \mathcal{E} \otimes \mathcal{F}$ and take $H = \{(X, Y) \in A\}$. Since $H_{x,z} = \{y \in \mathcal{Y} : (x, y) \in A\} = A_x$ for all $(x, z) \in \mathcal{X} \times \mathcal{Z}$, it follows that

$$P\big[(X, Y) \in A\big] = \int \alpha(x,\, A_x)\, \mu(dx) = \gamma(A).$$

This concludes the proof if γ is disintegrable. If τ is disintegrable, it suffices to take a disintegration $\{\beta(x, \cdot) : x \in \mathcal{X}\}$ for τ and to let

$$P(H) = \int \beta(x,\, H_{x,y})\, \gamma(dx, dy) \quad \text{where } H_{x,y} = \{z \in \mathcal{Z} : (x, y, z) \in H\}.$$

□

A quick consequence of Remark 3 and Lemma 4 is the following.

Corollary 5. *Suppose condition* (2) *holds for the probability measures γ on $\mathcal{E} \otimes \mathcal{F}$ and τ on $\mathcal{E} \otimes \mathcal{G}$. Then, condition* (1) *holds with a σ-additive P if at least one of the following conditions is satisfied:*

(*j*) *$\mathcal{F}$ is countably generated and the marginal of γ on $\mathcal{F}$ is perfect;*
(*jj*) *$\mathcal{G}$ is countably generated and the marginal of τ on $\mathcal{G}$ is perfect;*
(*jjj*) *All but perhaps one of the marginals of γ and τ on $\mathcal{E}$, $\mathcal{F}$ and $\mathcal{G}$ are perfect.*

Proof. Under (j) or (jj), one of γ and τ is disintegrable, and the conclusion follows from Lemma 4. Suppose (jjj) holds, and let $\mathcal{R}$ be the field on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ generated by the rectangles $A \times B \times C$ with $A \in \mathcal{E}$, $B \in \mathcal{F}$ and $C \in \mathcal{G}$. By Remark 3, there is a f.a.p. P_0 on $\mathcal{R}$ such that $P_0\big[X \in A, Y \in B\big] = \gamma(A \times B)$ and $P_0\big[X \in A,\, Z \in C\big] = \tau(A \times C)$ whenever $A \in \mathcal{E}$, $B \in \mathcal{F}$ and $C \in \mathcal{G}$. (Just take P_0 to be the restriction of Q to $\mathcal{R}$, with Q as in Remark 3). The marginals of P_0 on $\mathcal{E}$, $\mathcal{F}$ and $\mathcal{G}$ are all σ-additive. Moreover, all but perhaps one of such marginals are perfect. Hence, P_0 is σ-additive by [12, Theorem 6]. Thus, it suffices to take P to be the σ-additive extension of P_0 to $\sigma(\mathcal{R}) = \mathcal{E} \otimes \mathcal{F} \otimes \mathcal{G}$. □

We close this section with a last result related to disintegrability and perfectness. Assume condition (2) and let μ denote the (common) marginal of γ and τ on $\mathcal{E}$. If

$$\gamma^*(A \times \mathcal{Y}) = \mu^*(A) \quad \text{for all } A \subset \mathcal{X}, \tag{4}$$

then

$$\gamma_*(A \times \mathcal{Y}) = 1 - \gamma^*(A^c \times \mathcal{Y}) = \mu_*(A) \le \tau_*(A \times \mathcal{Z}) \le \tau^*(A \times \mathcal{Z})$$

where the first inequality depends on the definition of inner measure while the second is trivial. By Lemma 2, condition (1) holds for some f.a.p. P. Similarly, condition (1) holds for some f.a.p. P whenever

$$\tau^*(A \times \mathcal{Z}) = \mu^*(A) \quad \text{for all } A \subset \mathcal{X}.$$

Thus, it may be useful to have conditions for (4).

Lemma 6. *Let γ be a probability measure on $\mathcal{E} \otimes \mathcal{F}$ with marginals μ and ν on $\mathcal{E}$ and $\mathcal{F}$, respectively. Condition (4) holds provided, for each $H \in \mathcal{E} \otimes \mathcal{F}$, there are sub-$\sigma$-fields $\mathcal{E}_0 \subset \mathcal{E}$ and $\mathcal{F}_0 \subset \mathcal{F}$ such that $H \in \mathcal{E}_0 \otimes \mathcal{F}_0$ and γ is disintegrable on $\mathcal{E}_0 \otimes \mathcal{F}_0$. In particular, condition (4) holds if ν is perfect.*

Proof. It suffices to prove that $\mu^*(A) \le \gamma^*(A \times \mathcal{Y})$ (the opposite inequality follows from the definition of outer measure). Fix $A \subset \mathcal{X}$ and take $H \in \mathcal{E} \otimes \mathcal{F}$ such that $H \supset A \times \mathcal{Y}$ and $\gamma(H) = \gamma^*(A \times \mathcal{Y})$. Take also $\mathcal{E}_0 \subset \mathcal{E}$ and $\mathcal{F}_0 \subset \mathcal{F}$ such that $H \in \mathcal{E}_0 \otimes \mathcal{F}_0$ and γ is disintegrable on $\mathcal{E}_0 \otimes \mathcal{F}_0$. Given a disintegration $\{\alpha(x, \cdot) : x \in \mathcal{X}\}$ for γ on $\mathcal{E}_0 \otimes \mathcal{F}_0$, define

$$B = \{x \in \mathcal{X} : \alpha(x, H_x) = 1\} \quad \text{where } H_x = \{y \in \mathcal{Y} : (x, y) \in H\}.$$

Since $H \supset A \times \mathcal{Y}$, then $\alpha(x, H_x) = \alpha(x, \mathcal{Y}) = 1$ for each $x \in A$. Hence, $A \subset B$. Further, $B \in \mathcal{E}_0 \subset \mathcal{E}$, so that

$$\gamma^*(A \times \mathcal{Y}) = \gamma(H) = \int \alpha(x, H_x)\, \mu(dx) \ge \mu(B) \ge \mu^*(A).$$

Finally, suppose ν is perfect. Then, it suffices to note that each $H \in \mathcal{E} \otimes \mathcal{F}$ actually belongs to $\mathcal{E} \otimes \mathcal{F}_0$, for some countably generated sub-σ-field $\mathcal{F}_0 \subset \mathcal{F}$, and γ is disintegrable on $\mathcal{E} \otimes \mathcal{F}_0$ for ν is perfect and $\mathcal{F}_0$ countably generated. □

If γ is disintegrable (on all of $\mathcal{E} \otimes \mathcal{F}$) one can take $\mathcal{E}_0 = \mathcal{E}$ and $\mathcal{F}_0 = \mathcal{F}$. Thus, since any product probability is disintegrable, Lemma 6 improves [8, Proposition 3.4.2] and [18, Lemma 1.2.5].

3. Skorohod representations

3.1. A Wasserstein-type "distance".

In this Section, (S, d) is a metric space, $\mathcal{B}$ the Borel σ-field on S and $\mathbb{P}$ the set of probability measures on $\mathcal{B}$. For each $n \ge 1$, $\mathcal{B}^n = \mathcal{B} \otimes \ldots \otimes \mathcal{B}$ denotes the product σ-field on $S^n = S \times \ldots \times S$. Similarly, $\mathcal{B}^\infty$ is the product σ-field on S^∞, where S^∞ is the set of sequences

$$\omega = (\omega_0, \omega_1, \ldots) \quad \text{with } \omega_n \in S \text{ for each } n \ge 0.$$

Also, for μ, $\nu \in \mathbb{P}$, we let $\mathcal{F}(\mu, \nu)$ be the collection of those probability measures γ on $\mathcal{B}^2$ such that

$$\gamma(A \times S) = \mu(A) \quad \text{and} \quad \gamma(S \times A) = \nu(A) \quad \text{for each } A \in \mathcal{B}.$$

If (S,d) is not separable, $\mathcal{B}^2$ may be strictly smaller than the Borel σ-field on S^2, and this could be a problem for defining a Wasserstein-type "distance". Accordingly, we assume

$$\sigma(d) \subset \mathcal{B}^2, \tag{5}$$

that is, the function $d : S^2 \to [0,\infty)$ is measurable with respect to $\mathcal{B}^2$.

Condition (5) is trivially true if (S,d) is separable, as well as in various non separable situations. For instance, (5) holds if d is the uniform distance on some space S of cadlag functions, or if d is the 0-1 distance and $\text{card}(S) = \text{card}(\mathbb{R})$. A necessary condition of (5) is that $\mathcal{B} \supset \mathcal{C}$ for some countably generated σ-field $\mathcal{C}$ including the singletons. Hence, (5) yields $\text{card}(S) \leq \text{card}(\mathbb{R})$.

In any case, under (5), we let

$$W(\mu,\nu) = \inf_{\gamma \in \mathcal{F}(\mu,\nu)} \int 1 \wedge d(x,y)\, \gamma(dx,dy) = \inf_{\gamma \in \mathcal{F}(\mu,\nu)} E_\gamma(1 \wedge d)$$

for all μ, $\nu \in \mathbb{P}$. We also introduce the bounded-Lipschitz-metric

$$L(\mu,\nu) = \sup_f \Big| \int f\, d\mu - \int f\, d\nu \Big|$$

where sup is over those functions $f : S \to [-1,1]$ such that $|f(x) - f(y)| \leq d(x,y)$ for all x, $y \in S$.

It is worth noting that $L \leq 2\,W$. Fix in fact $\gamma \in \mathcal{F}(\mu,\nu)$ and a function f such that $-1 \leq f(x) \leq 1$ and $|f(x) - f(y)| \leq d(x,y)$ for all x, $y \in S$. If X and Y are the coordinate projections on S^2, then

$$\Big| \int f\, d\mu - \int f\, d\nu \Big| = |E_\gamma\{f(X)\} - E_\gamma\{f(Y)\}| \leq E_\gamma |f(X) - f(Y)| \leq 2\, E_\gamma(1 \wedge d).$$

We do not know whether W is a distance on all of $\mathbb{P}$. However, W looks a reasonable discrepancy measure between elements of $\mathbb{P}$ and is a distance when restricted on the subset {separable laws on $\mathcal{B}$}.

Lemma 7. *For all μ, $\nu \in \mathbb{P}$,*

$$W(\mu,\nu) = W(\nu,\mu), \quad W(\mu,\nu) = 0 \iff \mu = \nu, \quad \textit{and}$$
$$W(\mu,\nu) \leq W(\mu,\lambda) + W(\lambda,\nu) \quad \textit{if } \lambda \in \mathbb{P} \textit{ is separable.}$$

Proof. It is straightforward that $W(\mu,\nu) = W(\nu,\mu)$ and $W(\mu,\nu) = 0$ if $\mu = \nu$. Since $L \leq 2\,W$ and L is a distance on $\mathbb{P}$, then $\mu = \nu$ whenever $W(\mu,\nu) = 0$. Let $\lambda \in \mathbb{P}$ be separable. Define

$$W_0(\mu,\nu) = \inf_{\gamma \in \mathcal{D}(\mu,\nu)} E_\gamma(1 \wedge d)$$

where $\mathcal{D}(\mu,\nu)$ is the set of *disintegrable* probability measures on $\mathcal{B}^2$ with marginals μ and ν. Since $\mathcal{D}(\mu,\nu) \subset \mathcal{F}(\mu,\nu)$, then $W_0(\mu,\nu) \geq W(\mu,\nu)$. Because of [2, Theorem 4.1], W_0 satisfies the triangle inequality and separability of λ implies $W_0(\mu,\lambda) = W(\mu,\lambda)$ and $W_0(\lambda,\nu) = W(\lambda,\nu)$. Thus,

$$W(\mu,\nu) \leq W_0(\mu,\nu) \leq W_0(\mu,\lambda) + W_0(\lambda,\nu) = W(\mu,\lambda) + W(\lambda,\nu).$$

$\square$

3.2. **A finitely additive Skorohod representation.** To state our main result, we let X_n denote the n-th coordinate projection on S^∞, namely

$$X_n(\omega) = \omega_n \quad \text{for all } n \geq 0 \text{ and } \omega = (\omega_0, \omega_1, \dots) \in S^\infty.$$

Theorem 8. *Suppose $\sigma(d) \subset \mathcal{B}^2$ and $(\mu_n : n \geq 0)$ is a sequence of probability measures on $\mathcal{B}$. Then,*

$$\lim_n W(\mu_n, \mu_0) = 0$$

if and only if there is a f.a.p. P on $\mathcal{B}^\infty$ such that

(a) $X_n \xrightarrow{P} X_0$;

(b) *there is a sequence $\gamma_n \in \mathcal{F}(\mu_0, \mu_n)$, $n \geq 1$, satisfying*

$$P\big[(X_0, X_n) \in A\big] = \gamma_n(A) \quad \textit{whenever } A \in \mathcal{B}^2 \textit{ and } \gamma_n^*(\partial A) = 0;$$

(c) *$P(X_0 \in A) = \mu_0(A)$ for all $A \in \mathcal{B}$.*

Moreover, for each $n \geq 1$, one also obtains

$$P\big[(X_0, X_n) \in \cdot\big] = \gamma_n(\cdot) \quad \textit{on all of } \mathcal{B}^2 \textit{ whenever } \mu_n \textit{ is perfect.}$$

Proof. We first recall a known fact. Let $f : S^2 \to \mathbb{R}$ be a bounded continuous function such that $\sigma(f) \subset \mathcal{B}^2$. Given a f.a.p. γ on $\mathcal{B}^2$, define the field $\mathcal{U} = \{A \in \mathcal{B}^2 : \gamma^*(\partial A) = 0\}$. Since $\partial\{f \leq t\} \subset \{f = t\} \in \mathcal{B}^2$ for all $t \in \mathbb{R}$, then $\{f \leq t\} \in \mathcal{U}$ except possibly for countably many values of t. Hence $f_k \to f$, uniformly, for some sequence (f_k) of $\mathcal{U}$-simple functions.

Letting $f = 1 \wedge d$ and

$$\mathcal{U}_n = \{A \in \mathcal{B}^2 : \gamma_n^*(\partial A) = 0\},$$

for each $n \geq 1$ there is a sequence $(f_{n,k} : k \geq 1)$ of $\mathcal{U}_n$-simple functions such that

$$f_{n,k} \longrightarrow 1 \wedge d \quad \text{uniformly as } k \to \infty.$$

Suppose now that conditions (a)-(b) hold for some $\gamma_n \in \mathcal{F}(\mu_0, \mu_n)$ and some f.a.p. P. For fixed $n \geq 1$ and $\epsilon \in (0,1)$, condition (b) yields

$$\begin{aligned} E_{\gamma_n}(1 \wedge d) &= \lim_k E_{\gamma_n}(f_{n,k}) = \lim_k E_P\big\{f_{n,k}(X_0, X_n)\big\} \\ &= E_P\big\{1 \wedge d(X_0, X_n)\big\} \leq \epsilon + P\big(d(X_0, X_n) > \epsilon\big). \end{aligned}$$

Since $\gamma_n \in \mathcal{F}(\mu_0, \mu_n)$, then $W(\mu_n, \mu_0) \leq E_{\gamma_n}(1 \wedge d)$. Hence, condition (a) implies

$$\limsup_n W(\mu_n, \mu_0) \leq \limsup_n E_{\gamma_n}(1 \wedge d) \leq \epsilon \quad \text{for all } \epsilon \in (0,1).$$

Conversely, suppose $W(\mu_n, \mu_0) \to 0$, and take a sequence $\gamma_n \in \mathcal{F}(\mu_0, \mu_n)$, $n \geq 1$, such that $E_{\gamma_n}(1 \wedge d) \to 0$. Define

$$\mathcal{V}_n = \mathcal{B}^2 \text{ if } \mu_n \text{ is perfect and } \mathcal{V}_n = \mathcal{U}_n \text{ otherwise.}$$

Apply Lemma 2 to $\gamma = \gamma_1$ and $\tau = \gamma_2$. If μ_2 is perfect, apply also Lemma 6 to $\gamma = \gamma_2$. It follows that there is a f.a.p. Q_2 on $\mathcal{B}^3$ such that

$$Q_2\big[(X_0, X_1) \in A\big] = \gamma_1(A) \text{ for } A \in \mathcal{B}^2 \quad \text{and} \quad Q_2\big[(X_0, X_2) \in B\big] = \gamma_2(B) \text{ for } B \in \mathcal{V}_2.$$

In particular,

$$Q_2(X_0 \in \cdot) = \mu_0(\cdot) \text{ on } \mathcal{B} \text{ and } Q_2\big[(X_0, X_j) \in \cdot\big] = \gamma_j(\cdot) \text{ on } \mathcal{V}_j \text{ for } j = 1, 2.$$

By induction, let Q_n be a f.a.p. on $\mathcal{B}^{n+1}$ satisfying

$$Q_n(X_0 \in \cdot) = \mu_0(\cdot) \text{ on } \mathcal{B} \text{ and } Q_n\big[(X_0, X_j) \in \cdot\big] = \gamma_j(\cdot) \text{ on } \mathcal{V}_j \text{ for } j = 1, \dots, n.$$

Define $(\mathcal{X},\mathcal{E})=(\mathcal{Z},\mathcal{G})=(S,\mathcal{B})$, $(\mathcal{Y},\mathcal{F})=(S^n,\mathcal{B}^n)$, and note that

$$Q_n\big(A\times S^n\big)=\mu_0(A)=\gamma_{n+1}(A\times S)\quad\text{for } A\in\mathcal{B}.$$

Apply Lemma 2 to $\gamma=Q_n$ and $\tau=\gamma_{n+1}$. If μ_{n+1} is perfect, apply also Lemma 6 to $\gamma=\gamma_{n+1}$. Then, there is a f.a.p. Q_{n+1} on $\mathcal{B}^{n+2}$ such that

$$Q_{n+1}\big(A\times S\big)=Q_n(A) \text{ for } A\in\mathcal{B}^{n+1} \text{ and } Q_{n+1}\big[(X_0,X_{n+1})\in\cdot\big]=\gamma_{n+1}(\cdot) \text{ on } \mathcal{V}_{n+1}.$$

Finally, for each $n\geq 2$, take a f.a.p. P_n on $\mathcal{B}^\infty$ such that

$$P_n\big[(X_0,\ldots,X_n)\in A\big]=Q_n(A)\quad\text{for } A\in\mathcal{B}^{n+1}.$$

Define also

$$P(A)=\int P_n(A)\,\pi(dn)\quad\text{for } A\in\mathcal{B}^\infty,$$

where π is a f.a.p. on the power set of $\{1,2,\ldots\}$ such that $\pi\{n\}=0$ for all n. Then, P is a f.a.p. on $\mathcal{B}^\infty$ and

$$P\big[(X_0,\ldots,X_j)\in A\big]=\int_{\{n:n\geq j\}}P_n\big[(X_0,\ldots,X_j)\in A\big]\,\pi(dn)=Q_j(A)$$

for all $j\geq 1$ and $A\in\mathcal{B}^{j+1}$. Hence, conditions (b)-(c) are satisfied and

$$P\big[(X_0,X_j)\in\cdot\big]=\gamma_j(\cdot)\ \text{ on all of } \mathcal{B}^2 \text{ whenever } \mu_j \text{ is perfect.}$$

As to (a), the remark at the beginning of this proof yields

$$E_P\big\{1\wedge d(X_0,X_n)\big\}=E_{\gamma_n}(1\wedge d)\longrightarrow 0.$$

$\square$

Motivations for Theorem 8 have been given in Section 1. Here, we make a last remark.

For $n\geq 1$, Theorem 8 implies $P(X_n\in A)=\mu_n(A)$ if $A\in\mathcal{B}$ and $\mu_n(\partial A)=0$, or equivalently $E_P\big\{f(X_n)\big\}=\int f\,d\mu_n$ for all bounded continuous $f:S\to\mathbb{R}$. Unless μ_n is perfect, however, one does not obtain $P(X_n\in\cdot)=\mu_n(\cdot)$ on all of $\mathcal{B}$. This is certainly a drawback. On the other hand, this is also a typical finitely additive situation. We mention [6] and [11] as remarkable examples.

References

[1] Berti P., Pratelli L., Rigo P. (2007) Skorohod representation on a given probability space, *Probab. Theory Rel. Fields*, 137, 277-288.

[2] Berti P., Pratelli L., Rigo P. (2010) Skorohod representation theorem via disintegrations, *Sankhya*, 72, 208-220.

[3] Berti P., Pratelli L., Rigo P. (2011) A Skorohod representation theorem for uniform distance, *Probab. Theory Rel. Fields*, 150, 321-335.

[4] Berti P., Pratelli L., Rigo P. (2013) A Skorohod representation theorem without separability, *Electr. Commun. Probab.*, 18, 1-12.

[5] Bhaskara Rao K.P.S., Bhaskara Rao M. (1983) *Theory of charges*, Academic Press.

[6] Dubins L.E. (1999) Paths of finitely additive Brownian motion need not be bizarre, *Semin. de Probab. XXXIII*, L.N.M., Springer, 1709, 395-396.

[7] Dudley R.M. (1968) Distances of probability measures and random variables, *Ann. Math. Statist.*, 39, 1563-1572.

[8] Dudley R.M. (1999) *Uniform central limit theorems*, Cambridge University Press.
[9] Jakubowski A. (1998) The almost sure Skorokhod representation for subsequences in nonmetric spaces, *Theory Probab. Appl.*, 42, 167-174.
[10] Kallenberg O. (2002) *Foundations of modern probability*, Second edition, Springer.
[11] Karandikar R.L. (1988) A general principle for limit theorems in finitely additive probability: the dependent case, *J. Multivariate Anal.*, 24, 189-206.
[12] Ramachandran D. (1996) The marginal problem in arbitrary product spaces, In: *Distributions with fixed marginals and related topics* (Ruschendorf, Schweizer and Taylor eds.) IMS Lect. Notes Monog. Series, Vol. 28, 260-272.
[13] Sethuraman J. (2002) Some extensions of the Skorohod representation theorem, *Sankhya*, 64, 884-893.
[14] Skorohod A.V. (1956) Limit theorems for stochastic processes, *Theory Probab. Appl.*, 1, 261-290.
[15] Thorisson H. (1995) Coupling methods in probability theory, *Scand. J. Statist.*, 22, 159-182.
[16] Thorisson H. (2000) *Coupling, stationarity, and regeneration*, Springer.
[17] Thorisson H. (2014) Convergence in density in finite time windows and the Skorohod representation, *Submitted*.
[18] van der Vaart A., Wellner J.A. (1996) *Weak convergence and empirical processes*, Springer.
[19] Villani C. (2009) *Optimal transport, old and new*, Springer.
[20] Wichura M.J. (1970) On the construction of almost uniformly convergent random variables with given weakly convergent image laws, *Ann. Math. Statist.*, 41, 284-291.

Patrizia Berti, Dipartimento di Matematica Pura ed Applicata "G. Vitali", Universita' di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy
E-mail address: `patrizia.berti@unimore.it`

Luca Pratelli, Accademia Navale, viale Italia 72, 57100 Livorno, Italy
E-mail address: `pratel@mail.dm.unipi.it`

Pietro Rigo (corresponding author), Dipartimento di Matematica "F. Casorati", Universita' di Pavia, via Ferrata 1, 27100 Pavia, Italy
E-mail address: `pietro.rigo@unipv.it`