Logical Methods in Computer Science
Volume 17, Issue 2, 2021, pp. 3:1–3:12
https://lmcs.episciences.org/

Submitted Nov. 01, 2019
Published Apr. 14, 2021

SEMIPULLBACKS OF LABELLED MARKOV PROCESSES

JAN PACHL [a] AND PEDRO SÁNCHEZ TERRAF [b]

[a] Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada.

[b] Universidad Nacional de Córdoba. Facultad de Matemática, Astronomía, Física y Computación. Centro de Investigación y Estudios de Matemática (CIEM-FaMAF), Conicet. Córdoba. Argentina.

Abstract. A *labelled Markov process (LMP)* consists of a measurable space S together with an indexed family of Markov kernels from S to itself. This structure has been used to model probabilistic computations in Computer Science, and one of the main problems in the area is to define and decide whether two LMP S and S' "behave the same". There are two natural categorical definitions of sameness of behavior: S and S' are *bisimilar* if there exist an LMP T and measure preserving maps forming a diagram of the shape $S \leftarrow T \rightarrow S'$; and they are *behaviorally equivalent* if there exist some U and maps forming a dual diagram $S \rightarrow U \leftarrow S'$.

These two notions differ for general measurable spaces but Doberkat (extending a result by Edalat) proved that they coincide for analytic Borel spaces, showing that from every diagram $S \rightarrow U \leftarrow S'$ one can obtain a bisimilarity diagram as above. Moreover, the resulting square of measure preserving maps is commutative (a *semipullback*).

In this paper, we extend the previous result to measurable spaces S isomorphic to a universally measurable subset of a Polish space with the trace of the Borel σ-algebra, using a version of Strassen's theorem on common extensions of finitely additive measures.

1. Introduction

Markov decision processes have been considered in the Computer Science literature as a model for probabilistic computation. In this context, a *labelled Markov process (LMP)* is a structure $\mathbf{S} = (S, \Sigma, \{\tau_a : a \in L\})$ where (S, Σ) is a measurable space and for $a \in L$, $\tau_a : S \times \Sigma \rightarrow [0, 1]$ is a *Markov kernel*, i.e., a function such that for each fixed $s \in S$, $\tau(s, \cdot)$ is a finite positive measure bounded above by 1, and for each fixed $Q \in \Sigma$, $\tau(\cdot, Q)$ is a Σ-$\mathfrak{B}([0, 1])$-measurable function. In one interpretation of this computational model, the system $\mathbf{S}$ stands at any particular time at a *current state* $s_0 \in S$, but this information is hidden from the hypothetical *users* of $\mathbf{S}$, whose only interaction with the system is through L. Intuitively, the user is presented with a black box with buttons labelled by L, and a button a is available to be pressed whenever $\tau_a(s_0, S) > 0$. A detailed discussion of LMP and many motivating examples are to be found in Desharnais' thesis [Des99].

Key words and phrases: semipullback, Radon measure, labelled Markov process, bisimulation, universally measurable set. *MSC 2010:* 28A35; 28A60, 68Q85. *ACM class:* F.4.1; F.1.2.

Supported by Secyt-UNC project 33620180100465CB.

LOGICAL METHODS IN COMPUTER SCIENCE

DOI:10.23638/LMCS-17(2:3)2021

Of primary importance is to be able to determine when two such systems $\mathbf{S}$ and $\mathbf{S}'$ behave the same way from the user viewpoint. That is, when a user doing repeated experiments with $\mathbf{S}$ and $\mathbf{S}'$ would conclude that they are indistinguishable. Actually, for such probabilistic systems there are at least two different ways to formalize a notion of behavior, and they are intimately related to measure-preserving maps.

Definition 1.1. Let $\mathbf{S} = (S, \Sigma, \{\tau_a : a \in L\})$ and $\mathbf{S}' = (S', \Sigma', \{\tau'_a : a \in L\})$ be LMP. A *zigzag morphism* $f : \mathbf{S} \to \mathbf{S}'$ is a surjective measurable map $f : (S, \Sigma) \to (S', \Sigma')$ such that for all $a \in L$ we have:

$$\forall s \in S\ \forall Q \in \Sigma' : \tau_a(s, f^{-1}(Q)) = \tau'_a(f(s), Q).$$

We say that $\mathbf{S}$ and $\mathbf{S}'$ are *bisimilar* if there exists an LMP $\mathbf{T}$ and zigzag morphisms forming a diagram of the shape $\mathbf{S} \leftarrow \mathbf{T} \to \mathbf{S}'$. This definition, in this categorical form, can be traced to Joyal *et al.* [JNW96] and it provides one of the possible formalizations of the concept of equality of behavior. The second one is given by the dual diagram: $\mathbf{S}$ and $\mathbf{S}'$ are *behaviorally equivalent* if there exists an LMP $\mathbf{U}$ and morphisms forming a diagram of the shape $\mathbf{S} \to \mathbf{U} \leftarrow \mathbf{S}'$. This notion, in turn, was introduced by Danos *et al.* [DDLP06], and it can be shown by functorial manipulations that behavioral equivalence (also known as "event bisimilarity", i.e. the greatest *event bisimulation*) is a transitive relation in the category of LMP. It can be proved that bisimilar LMP are behaviorally equivalent. Also, a neat logical characterization of this last relation is given in [DDLP06]. The originating papers of the concepts of LMP and bisimilarity with its logical characterization are Blute *et al.* [BDEP97] and Desharnais *et al.* [DEP98], and the presentation of the results of both papers were soon after streamlined in [DEP02]. An alternative general source on the topic is Doberkat [Dob09].

Some of the main problems in this area are to find conditions for the relation of bisimilarity to be transitive, and more strongly, for behavioral equivalence to entail bisimilarity. This is not true in the general case [ST11], but there are various important positive results which restrict or otherwise modify the category of processes and measurable spaces considered. The first one was obtained by Edalat [Eda99] for a category of LMP with a relaxed measurability condition on Markov kernels (these are only required to be *universally measurable*) over analytic spaces: In such category of *generalized* LMP, every *cospan* $\mathbf{S} \to \mathbf{U} \leftarrow \mathbf{S}'$ can be completed to a commutative square by finding an appropriate $\mathbf{T}$ and arrows to $\mathbf{S}$, $\mathbf{S}'$. This $\mathbf{T}$ is called the *semipullback* of the cospan. Later, Doberkat [Dob05] obtained the same result now properly for the category of LMP (with kernels as defined above) over analytic spaces. He specifically showed the existence of semipullbacks in the category of Markov kernels (that is, LMP with a singleton label set L) over analytic state spaces and Borel zigzag maps; from this, the result for general label sets follows.

In the present paper we will show that the existence of semipullbacks holds in the larger category of Markov kernels over universally measurable spaces. Our proof does not rely on the existence of disintegrations (regular conditional probabilities) as in [Eda99], but we use a result about common extensions of finitely additive measures (Lemma 3.3, a version of Strassen's theorem). In Section 2 we present a related category, that of *probability kernels*. The main technical result of this paper is to show that this category has semipullbacks. In Section 3 we gather some results on extensions of finitely additive measures. Section 4 presents the construction of the semipullback S_3 of a given cospan of probability kernels $S_1 \to S_0 \leftarrow S_2$; this is essentially built over the set-theoretic pullback of that diagram. The reduction of the problem of Markov kernels and general LMP to our result is done in

Section 5; in particular we show that LMP over coanalytic Borel spaces have semipullbacks. We conclude with some counterexamples in the last section.

2. Probability kernels

We find it technically convenient to describe the main construction in terms of probability kernels from a fixed measurable space. Let (X, Ξ) and (S, Σ) be two measurable spaces. As in [Kal02, Ch.1], a mapping $\mu\colon X \times \Sigma \to [0, \infty)$ is a *kernel from X to S* if $\mu(x, \cdot)$ is a measure on Σ for each $x \in X$ and $\mu(\cdot, Q)$ is a Ξ–$\mathfrak{B}([0, 1])$-measurable function on X for each $Q \in \Sigma$. From now on we write $\mu^x(Q)$ instead of $\mu(x, Q)$ for $x \in X$, $Q \in \Sigma$.

Recall that a *Radon measure* is a (non-negative) measure defined on the σ-algebra of Borel sets of a topological space such that it is inner regular with respect to compact sets; that is, $\mu(B) = \sup \mu(K)$ where the supremum is over compact subsets K of B, for every Borel set B. We say that μ is a *probability kernel* if $\mu^x(S) = 1$ for all $x \in X$, and a *subprobability kernel* if $\mu^x(S) \leq 1$ for all $x \in X$. We say that μ is a *Radon (sub)probability kernel* if moreover S is a topological space, Σ is its Borel σ-algebra and every μ^x is a Radon (sub)probability measure on Σ.

Thus a Markov kernel in the definition of LMP above is a subprobability kernel from S to itself.

When μ is a kernel from X to S, we write (S, Σ, μ) instead of μ when (X, Ξ) is understood and we wish to make S and Σ explicit.

For a fixed (X, Ξ), kernels from X form a category with surjective measure-preserving maps as morphisms:

Definition 2.1. Let (X, Ξ) be a fixed measurable space. For $j = 1, 2$ and $x \in X$ let (S_j, Σ_j, μ_j^x) be a measure space such that μ_j is a kernel from X to S_j. A mapping $h\colon S_1 \to S_2$ is a *kernel morphism from μ_1 to μ_2* if it is Σ_1–Σ_2 measurable, $h(S_1) = S_2$, and $\mu_1^x(h^{-1}(A)) = \mu_2^x(A)$ for all $x \in X$, $A \in \Sigma_2$. A morphism h from μ_1 to μ_2 is sometimes written $h\colon (S_1, \Sigma_1, \mu_1) \to (S_2, \Sigma_2, \mu_2)$, or simply $h\colon S_1 \to S_2$ when Σ_j and μ_j are understood.

We find that notation is somewhat simpler when we work with kernel morphisms rather than zigzag morphisms. Once we prove the existence of a semipullback for kernel morphisms, the existence for zigzag morphisms will easily follow (Section 5).

In the present paper we prove:

Theorem 2.2. *Let (X, Ξ) be a fixed measurable space. Consider the category in which each object is a Radon subprobability kernel from X to a separable metric space, and morphisms are kernel morphisms. Every cospan $(S_1, \Sigma_1, \mu_1) \to (S_0, \Sigma_0, \mu_0) \leftarrow (S_2, \Sigma_2, \mu_2)$ has a semipullback (S_3, Σ_3, μ_3) such that S_3 is the set pullback of $S_1 \to S_0 \leftarrow S_2$. Moreover S_3 is a measurable subset of $S_1 \times S_2$.*

We will first prove the theorem for probability kernels. For this we fix, up to Section 4, three Radon probability kernels (S_j, Σ_j, μ_j) (for $j = 0, 1, 2$) as in the statement of Theorem 2.2, and for $j = 1, 2$, kernel morphisms $h_j\colon S_j \to S_0$. Our goal is to construct a semipullback of h_1, h_2, i.e. (S_3, Σ_3, μ_3) and for $j = 1, 2$, morphisms $k_j\colon S_3 \to S_j$ such that $h_1 \circ k_1 = h_2 \circ k_2$ (see Figure 1).

We first proceed to construct the pullback of the mappings h_1 and h_2 in the category of measurable spaces, whose upper vertex will be the underlying space of the semipullback.

$$\begin{tikzcd}
 & (S_3,\Sigma_3,\mu_3) \arrow[dl, "k_1"'] \arrow[dr, "k_2"] & \\
(S_1,\Sigma_1,\mu_1) \arrow[dr, "h_1"'] & & (S_2,\Sigma_2,\mu_2) \arrow[dl, "h_2"] \\
 & (S_0,\Sigma_0,\mu_0) &
\end{tikzcd}$$

Figure 1: A semipullback.

Let $\pi_j : S_1 \times S_2 \to S_j$, $j = 1, 2$, be the natural projections. Denote by $\Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2$ the smallest σ-algebra on $S_1 \times S_2$ for which π_j are Σ_j-measurable.

Define

$$\begin{aligned} S_3 &:= \{(x_1, x_2) \in S_1 \times S_2 : h_1(x_1) = h_2(x_2)\} \\ \Sigma_3 &:= \{A \cap S_3 : A \in \Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2\} \\ k_j &:= \text{ restriction of } \pi_j \text{ to } S_3, \text{ for } j = 1, 2. \end{aligned} \tag{2.1}$$

Then $k_j(S_3) = S_j$ for $j = 1, 2$ (because $h_j(S_j) = S_0$).

All that remains now is to construct the probability kernel μ_3; this will be done in several steps. We define a countable algebra $\mathfrak{A} \subseteq \Sigma_3$ that generates Σ_3 as a σ-algebra, and finitely additive measures ν_3^x on $\mathfrak{A}$. In defining ν_3^x we use a constructive variant of the Hahn-Banach theorem, to ensure that ν_3^x is a measurable function of x. Then we prove that ν_3^x is countably additive, so that it extends to a countably additive measure μ_3^x on Σ_3.

3. PRELIMINARIES

In this section we establish notation and several results that will be needed in the main construction.

If S is a set and V is a set of real-valued functions on S, write $V^+ := \{f \in V : f \geq 0\}$. Here $\geq$ is the pointwise partial order.

When $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are algebras of subsets of S_1 and S_2, we denote by $\mathfrak{B}_1 \otimes \mathfrak{B}_2$ the algebra of subsets of $S_1 \times S_2$ consisting of finite unions of sets of the form $B_1 \times B_2$, $B_j \in \mathfrak{B}_j$ for $j = 1, 2$. Recall that we denote by $\mathfrak{B}_1 \mathbin{\widehat{\otimes}} \mathfrak{B}_2$ the σ-algebra generated by $\mathfrak{B}_1 \otimes \mathfrak{B}_2$.

If $\mathfrak{A}$ is an algebra of sets, denote by $\mathsf{L}(\mathfrak{A})$ the space of simple $\mathfrak{A}$-measurable functions; that is, functions of the form $\sum_{i \in F} r_i \chi_{A_i}$ where F is a finite set, $A_i \in \mathfrak{A}$ and $r_i \in \mathbb{R}$ for $i \in F$. If in addition ν is a finitely additive measure on $\mathfrak{A}$, denote by $\overline{\nu}$ its integral defined on $\mathsf{L}(\mathfrak{A})$:

$$\overline{\nu}(f) := \int f \mathrm{d}\nu \quad \text{for } f \in \mathsf{L}(\mathfrak{A}).$$

In other words, $\overline{\nu}(f) = \sum_{i \in F} r_i \nu(A_i)$ when $f = \sum_{i \in F} r_i \chi_{A_i}$. This is well-defined— see e.g. [DS57, Ch. III].

Note that if $\nu^x(A)$ is a measurable function of x for every $A \in \mathfrak{A}$ then so is $\overline{\nu^x}(f)$ for every $f \in \mathsf{L}(\mathfrak{A})$.

Lemma 3.1. *Any Borel-measurable image of a compact metrizable space in a separable metric space is analytic, and therefore universally measurable.*

Proof. Let K be compact, S a separable metric space, and $h : K \to S$ Borel. Then the completion $\hat{S}$ of S is analytic and $h : K \to \hat{S}$ is also Borel. Hence by [Fre13, 423G(b)] the image is analytic, hence by [Fre13, 434D(c)] it is universally measurable. $\square$

Lemma 3.2. *Let S and S_0 be separable metric spaces and Σ, Σ_0 their Borel σ-algebras. Let μ and μ_0 be Radon probability measures on Σ and Σ_0, respectively. Let $h\colon S \to S_0$ be a Borel-measurable measure-preserving mapping. Then for every $A \in \Sigma$ there exists $D \in \Sigma_0$ such that $D \subseteq h(A)$ and $\mu_0(D) \geq \mu(A)$.*

Proof. Since μ is inner regular with respect to compact sets, we can find a countable family $\mathcal{K}$ of compact subsets of A such that

$$\mu(A) = \sup\{\mu(K) : K \in \mathcal{K}\}.$$

For every such K the image $h(K)$ is universally measurable by Lemma 3.1. Therefore, there exist $B_K, B'_K \in \Sigma_0$ such that $B_K \subseteq h(K) \subseteq B'_K$ and $\mu_0(B'_K \setminus B_K) = 0$. Since h is measure-preserving, we have

$$\mu_0(B_K) = \mu_0(B'_K) = \mu(h^{-1}(B'_K)).$$

Now $h^{-1}[B'_K] \supseteq K$, and therefore $\mu_0(B_K) = \mu(h^{-1}(B'_K)) \geq \mu(K)$ for each $K \in \mathcal{K}$. Hence we may take $D := \bigcup\{B_K : K \in \mathcal{K}\}$. $\square$

Lemma 3.3 [Fre13, 457C]**.** *Let $\mathfrak{A}$ be an algebra of subsets of a set S, and $\mathfrak{A}_j$, $j = 1, 2$, two of its subalgebras. Let $\nu_j\colon \mathfrak{A}_j \to [0, 1]$, $j = 1, 2$, be finitely additive measures such that $\nu_1(S) = \nu_2(S) = 1$. Assume that $\nu_1(A_1) + \nu_2(A_2) \leq 1$ whenever $A_j \in \mathfrak{A}_j$, $j = 1, 2$, are such that $A_1 \cap A_2 = \emptyset$. Then there exists a finitely additive measure $\nu\colon \mathfrak{A} \to [0, 1]$ that extends both ν_1 and ν_2.* $\square$

We need the following variant of the Hahn-Banach theorem, which preserves measurability.

Lemma 3.4. *Let S be a non-empty set. Let V be a linear space of bounded functions, and W a subspace of V that contains all constant functions on S. Assume also that V has a countable basis as a linear space. Let $\Psi^x\colon W \to \mathbb{R}$, $x \in X$, be a collection of linear functionals on W such that $\Psi^x(f)$ is a measurable function of x for every $f \in W$, $\Psi^x(1) = 1$ and $\Psi^x(f) \geq 0$ for $f \in W^+$, $x \in X$. Then there is a collection of linear functionals $\Phi^x\colon V \to \mathbb{R}$, $x \in X$, that extend Ψ^x and such that $\Phi^x(f)$ is a measurable function of x for every $f \in V$, and $\Phi^x(f) \geq 0$ for $f \in V^+$, $x \in X$.*

Proof. Extend Ψ^x one dimension at a time. Assume that Φ^x has been defined on a linear subspace $U \supseteq W$ and consider $f_0 \in V \setminus U$, We are going to extend Φ^x to $U + \mathbb{R}f_0$ as follows.

$$\begin{aligned} p^x(f) &:= \inf\{\Phi^x(g) : g \in U \text{ and } g \geq f\} \quad \text{for } f \in U + \mathbb{R}f_0 \\ \Phi^x(f_0) &:= \inf\{p^x(g + f_0) - \Phi^x(g) : g \in U\}. \end{aligned}$$

Thus p^x is subadditive and positively homogeneous on $U + \mathbb{R}f_0$. We claim that $\Phi^x(f) \leq p^x(f)$ for every $f \in U + \mathbb{R}f_0$. To prove the claim, write $f = u + rf_0$ where $u \in U$ and $r \in \mathbb{R}$, and distinguish two cases:

For $r > 0$ use $g = u/r$ in the definition of $\Phi^x(f_0)$ to get

$$\begin{aligned}\Phi^x(f) = \Phi^x(u) + r\Phi^x(f_0) &= \Phi^x(u) + \inf\{rp^x(g + f_0) - r\Phi^x(g) : g \in U\}\\ &\le \Phi^x(u) + rp^x((u/r) + f_0) - r\Phi^x(u/r) = p^x(f).\end{aligned}$$

When $r < 0$, for every $g \in U$ we have

$$\Phi^x(g) - \Phi^x(u/r) = p^x(g - (u/r)) \le p^x(g + f_0) + p^x(-(u/r) - f_0)$$

Therefore

$$-p^x(-(u/r) - f_0) - \Phi^x(u/r) \le \inf\{p^x(g + f_0) - \Phi^x(g) : g \in U\} = \Phi^x(f_0),$$

and finally

$$\Phi^x(f) = r\Phi^x(u/r) + r\Phi^x(f_0) \le -rp^x(-(u/r) - f_0) = p^x(u + rf_0) = p^x(f).$$

That proves the claim. It follows that if $f \in (U + \mathbb{R}f_0)^+$ then

$$\Phi^x(f) = -\Phi^x(-f) \ge -p^x(-f) \ge 0.$$

To prove that $\Phi^x(f_0)$ is a measurable function of x, fix a countable basis C of U such that $1 \in C$ and define $\widetilde{U}$ to be the set of finite linear combinations of elements of C with rational coefficients. Then

$$\begin{aligned}p^x(f) &= \inf\{\Phi^x(g) : g \in \widetilde{U} \text{ and } g \ge f\} \quad \text{for } f \in U + \mathbb{R}f_0\\ \Phi^x(f_0) &= \inf\{p^x(g + f_0) - \Phi^x(g) : g \in \widetilde{U}\}\end{aligned}$$

so that $x \mapsto \Phi^x(f_0)$ is the infimum of a countable set of measurable functions. □

The next lemma is a variant of a theorem of Marczewski and Ryll-Nardzewski [MRN53]. When $\mathfrak{B}_j = \Sigma_j$, this is a special case of [Fre13, 454C].

Lemma 3.5. *Let S_1 be a Hausdorff topological space, Σ_1 its Borel σ-algebra and $\mu_1 \colon \Sigma_1 \to [0, 1]$ a Radon probability measure. Let (S_2, Σ_2, μ_2) be any probability space. Denote by $\pi_j \colon S_1 \times S_2 \to S_j$ the natural projections. For $j = 1, 2$, let $\mathfrak{B}_j \subseteq \Sigma_j$ be an algebra of subsets of S_j. Let $\mu \colon \mathfrak{B}_1 \otimes \mathfrak{B}_2 \to [0, 1]$ be a finitely additive measure such that $\mu(\pi_j^{-1}(B_j)) = \mu_j(B_j)$ for $j = 1, 2$ and all $B_j \in \mathfrak{B}_j$. Then μ has an extension to a countably additive measure on the σ-algebra $\mathfrak{B}_1 \mathbin{\widehat{\otimes}} \mathfrak{B}_2$.*

Proof. This is a minor modification of the proof of [Fre13, 454C]. Let $\mathfrak{D}$ be the set of finite unions of sets of the form $C \times B_2$ where C is a compact subset of S_1 and $B_2 \in \mathfrak{B}_2$. As μ_1 is a Radon measure, it follows that for every $\varepsilon > 0$ and $B \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ there are $D \in \mathfrak{D}$ and $E \in \Sigma_1$ such that $D \subseteq B$, $\mu_1(E) < \varepsilon$ and $B \subseteq D \cup (E \times S_2)$.

Now let $\{B_i\}_{i\in\mathbb{N}}$ be a non-increasing sequence of sets in $\mathfrak{B}_1 \otimes \mathfrak{B}_2$ with empty intersection. To prove that $\lim_i \mu(B_i) = 0$, take any $\varepsilon > 0$. There are $D'_i \in \mathfrak{D}$ and $E'_i \in \Sigma_1$ such that $D'_i \subseteq B_i$, $\mu_1(E'_i) < 2^{-i}\varepsilon$ and $B_i \subseteq D'_i \cup (E'_i \times S_2)$. Set $D_n := \bigcap_{i\le n} D'_i$ and $E_n := \bigcup_{i\le n} E'_i$ for each n. Then $\{D_n\}_n$ is a non-increasing sequence of sets in $\mathfrak{D}$, $D_n \subseteq B_n$, $\mu_1(E_n) < 2\varepsilon$ and $B_n \subseteq D_n \cup (E_n \times S_2)$.

For $n \in \mathbb{N}$ and $y \in S_2$ set $D_n^y := \pi_1(D_n \cap \pi_2^{-1}(y))$ and $H_n := \pi_2(D_n)$. Then $\{D_n^y\}_n$ and $\{H_n\}_n$ are non-increasing sequences of subsets of S_1 and S_2, respectively. The sets D_n^y are compact and $H_n \in \mathfrak{B}_2$. Next $\bigcap_n D_n^y = \emptyset$ because $\bigcap_n D_n \subseteq \bigcap_n B_n = \emptyset$. Hence for every $y \in S_2$ there is n such that $D_n^y = \emptyset$, which means that $\bigcap_n H_n = \emptyset$. It follows that

$$\lim_i \mu(B_i) \leq \lim_i \mu(S_1 \times H_i) + \lim_i \mu(B_i \setminus (S_1 \times H_i)) \\ \leq \lim_i \mu_2(H_i) + \lim_i \mu_1(E_i) \leq 2\varepsilon.$$

We have proved that $\lim_i \mu(B_i) = 0$. By [Fre13, 413K] μ has an extension to a countably additive measure on the σ-algebra generated by $\mathfrak{B}_1 \otimes \mathfrak{B}_2$. □

4. Proof of the main theorem

In this section we complete the proof of Theorem 2.2. Recall that S_j, $j = 0, 1, 2$, are separable metric spaces, Σ_j are their Borel σ-algebras, and μ_j are Radon probability kernels from X to S_j.

Our goal is to construct a semipullback (S_3, Σ_3, μ_3) of the cospan $(S_1, \Sigma_1, \mu_1) \to (S_0, \Sigma_0, \mu_0) \leftarrow (S_2, \Sigma_2, \mu_2)$. In Section 2 we have already defined S_3, Σ_3, and the maps k_1 and k_2 closing the diagram. To complete the construction, we will define the measures ν_3^x on $\mathfrak{A}$ by using the results of the previous section.

For $j = 0, 1, 2$, fix countable algebras $\mathfrak{B}_j \subseteq \Sigma_j$ such that

- $\mathfrak{B}_j$ generates Σ_j as a σ-algebra for $j = 0, 1, 2$, and
- $h_j^{-1}(B_0) \in \mathfrak{B}_j$ whenever $B_0 \in \mathfrak{B}_0$, for $j = 1, 2$.

For $j = 1, 2$, define $\mathfrak{A}_j := \{k_j^{-1}(B) : B \in \mathfrak{B}_j\}$. Let $\mathfrak{A}$ be the algebra of subsets of S_3 generated by $\mathfrak{A}_1 \cup \mathfrak{A}_2$. Then $\mathfrak{A}$ is countable and it generates Σ_3 as a σ-algebra.

For $j = 1, 2$, and $B \in \mathfrak{B}_j$, let $\nu_j^x(k_j^{-1}(B)) := \mu_j^x(B)$. As $k_j(S_3) = S_j$, this is well defined and ν_j^x is a finitely additive measure on $\mathfrak{A}_j$.

Take any $A_j \in \mathfrak{A}_j$, $j = 1, 2$, such that $A_1 \cap A_2 = \emptyset$. Then $A_j = k_j^{-1}(B_j)$ for some $B_j \in \mathfrak{B}_j$. By Lemma 3.2 there are $D_j \in \Sigma_0$ such that $D_j \subseteq h_j(B_j)$ and $\mu_0^x(D_j) \geq \mu_j^x(B_j)$. From the definition of S_3 we get $h_1(B_1) \cap h_2(B_2) = \emptyset$, hence $D_1 \cap D_2 = \emptyset$. Therefore

$$\nu_1^x(A_1) + \nu_2^x(A_2) = \mu_1^x(B_1) + \mu_2^x(B_2) \leq \mu_0^x(D_1) + \mu_0^x(D_2) \leq 1.$$

By Lemma 3.3 there is a finitely additive measure ν^x on $\mathfrak{A}$ that extends both ν_1^x and ν_2^x. As the proof of Lemma 3.3 relies on the axiom of choice, $\nu^x(A)$ is not necessarily a measurable function of x for every $A \in \mathfrak{A}$. However, observe that $\overline{\nu^x}(f)$ is a measurable function of x for every $f \in \mathsf{L}(\mathfrak{A}_1) + \mathsf{L}(\mathfrak{A}_2)$. Indeed, if $f = f_1 + f_2$, $f_j \in \mathsf{L}(\mathfrak{A}_j)$ for $j = 1, 2$, then

$$\overline{\nu^x}(f) = \overline{\nu_1^x}(f_1) + \overline{\nu_2^x}(f_2)$$

by the linearity of integral, so that $\overline{\nu^x}(f)$ is a sum of two measurable functions of x.

By Lemma 3.4 with $W = \mathsf{L}(\mathfrak{A}_1) + \mathsf{L}(\mathfrak{A}_2)$ there is a linear functional $\Phi^x \colon \mathsf{L}(\mathfrak{A}) \to \mathbb{R}$ that agrees with $\overline{\nu^x}$ on $\mathsf{L}(\mathfrak{A}_1) + \mathsf{L}(\mathfrak{A}_2)$ and such that $\Phi^x(f)$ is a measurable function of x for every $f \in \mathsf{L}(\mathfrak{A})$, and $\Phi^x(f) \geq 0$ for $f \in \mathsf{L}(\mathfrak{A})^+$, $x \in X$. Now $\nu_3^x(A) := \Phi^x(\chi_A)$ for $A \in \mathfrak{A}$ defines a finitely additive measure $\nu_3^x \geq 0$ on $\mathfrak{A}$ such that $\nu_3^x(S_3) = 1$ and $x \mapsto \nu_3^x(A)$ is measurable for every $A \in \mathfrak{A}$.

We have $\mathfrak{A} = \{B \cap S_3 : B \in \mathfrak{B}_1 \otimes \mathfrak{B}_2\}$. For $B \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ define $\mu^x(B) := \nu_3^x(B \cap S_3)$. By Lemma 3.5 each μ^x extends to a countably additive measure $\widehat{\mu}^x$ on the σ-algebra $\mathfrak{B}_1 \mathbin{\widehat{\otimes}} \mathfrak{B}_2 = \Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2$, which is the Borel σ-algebra of the product topology on $S_1 \times S_2$. By [Fre13, 454A(a)], $\widehat{\mu}^x$ is a Radon measure.

Lemma 4.1. *$S_3 \in \Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2$; moreover, we have*

$$S_1 \times S_2 \setminus S_3 = \bigcup \{ h_1^{-1}(B_0) \times h_2^{-1}(S_0 \setminus B_0) : B_0 \in \mathfrak{B}_0 \} \tag{4.1}$$

and $\widehat{\mu}^x(S_3) = 1$.

Proof. Take any $(x_1, x_2) \in S_1 \times S_2 \setminus S_3$. Then $h_1(x_1) \neq h_2(x_2)$, hence there is $B_0 \in \mathfrak{B}_0$ such that $h_1(x_1) \in B_0$ and $h_2(x_2) \notin B_0$, which means $(x_1, x_2) \in h_1^{-1}(B_0) \times h_2^{-1}(S_0 \setminus B_0)$.

Each set $B := h_1^{-1}(B_0) \times h_2^{-1}(S_0 \setminus B_0)$, where $B_0 \in \mathfrak{B}_0$, is in the algebra $\mathfrak{B}_1 \otimes \mathfrak{B}_2$ and $\widehat{\mu}^x(B) = \mu^x(B) = \nu_3^x(\emptyset) = 0$. It follows that $S_3 \in \Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2$ and $\widehat{\mu}^x(S_3) = 1$. □

By Lemma 4.1, S_3 is a measurable subset of $S_1 \times S_2$. Define μ_3^x to be the restriction of $\widehat{\mu}^x$ to the σ-algebra Σ_3.

It remains to be proved that for every $E \in \Sigma_3$ the function $x \mapsto \mu_3^x$ is measurable. To that end define

$$\mathcal{D} := \{ E \in \Sigma_3 : x \mapsto \mu_3^x(E) \text{ is measurable} \}.$$

Then $\mathfrak{A} \subseteq \mathcal{D}$, and $\mathcal{D}$ is closed under complements and unions of disjoint sequences. By the Monotone Class Theorem [Fre11, 136B] we have $\mathcal{D} = \Sigma_3$.

That completes the proof of Theorem 2.2 for the case of probability kernels. To extend the result to subprobability kernels we work as follows. Let $(S_1, \Sigma_1, \mu_1) \to (S_0, \Sigma_0, \mu_0) \leftarrow (S_2, \Sigma_2, \mu_2)$ be a cospan of Radon subprobability kernels, where S_j are separable metric spaces.

Define $\bar{S}_j := S_j \oplus \{s_j\}$ where $s_j \notin S_j$ for each $j = 0, 1, 2$ and for measurable $E \subseteq \bar{S}_j$, let

$$\bar{\mu}_j^x(E) := \mu_j^x(E \cap S_j) + (1 - \mu_j^x(S_j)) \cdot \chi_E(s_j)$$

Then $\bar{\mu}_j$ are Radon probability kernels. We also extend the maps h_j by stipulating

$$\bar{h}_j(x) := \begin{cases} h_j(x) & x \neq s_j \\ s_0 & x = s_j \end{cases}$$

for $j = 1, 2$. Then $\bar{h}_j$ are kernel morphisms.

By Theorem 2.2 for probability kernels, the cospan $\bar{S}_1 \to \bar{S}_0 \leftarrow \bar{S}_2$ has a semipullback $(\bar{S}_3, \bar{\Sigma}_3, \bar{\mu}_3)$ with kernel morphisms $k_j \colon \bar{S}_3 \to \bar{S}_j$, and $\bar{S}_3 \subseteq \bar{S}_1 \times \bar{S}_2$ is the set pullback. Hence $\bar{S}_3 = S_3 \oplus \{(s_1, s_2)\}$ where S_3 is the set pullback of $S_1 \to S_0 \leftarrow S_2$.

We can take μ_3^x to be the restriction of $\bar{\mu}_3^x$ to $\Sigma_3 := \{ E \cap S_3 : E \in \bar{\Sigma}_3 \}$. It is straightforward to check that the restrictions $k_j \restriction S_3$ are kernel morphisms from S_3 onto S_j for $j = 1, 2$, and we are done.

5. APPLICATION TO THE PROBLEM OF BISIMULATION

5.1. **Labelled Markov processes with Radon measures.** By Theorem 2.2, semipullbacks exist in a certain category of subprobability kernels from a fixed measurable space (X, Ξ).

As a corollary we obtain the following theorem, which asserts the existence of semipullbacks in the corresponding category of LMP and zigzag morphisms:

Theorem 5.1. *Consider the category in which objects are LMP* $(S, \Sigma, \{\tau_a : a \in L\})$ *such that* S *is a separable metric space and* $\tau_a(s, \cdot)$ *are Radon measures, with zigzag morphisms. In this category every cospan has a semipullback.*

Moreover, every cospan

$$(S_1, \Sigma_1, \{\tau_{1a} : a \in L\}) \to (S_0, \Sigma_0, \{\tau_{0a} : a \in L\}) \leftarrow (S_2, \Sigma_2, \{\tau_{2a} : a \in L\})$$

has a semipullback $(S_3, \Sigma_3, \{\tau_{3a} : a \in L\})$ *such that* S_3 *is the set pullback of* $S_1 \to S_0 \leftarrow S_2$ *and* S_3 *is a measurable subset of* $S_1 \times S_2$*.*

Proof. First we deal with the LMP for which the label set L has a single element a, and write $\tau = \tau_a$.

Let $(S_1, \Sigma_1, \tau_1) \to (S_0, \Sigma_0, \tau_0) \leftarrow (S_2, \Sigma_2, \tau_2)$ be a cospan in the given category, with connecting zigzags $h_j \colon S_j \to S_0$, $j = 1, 2$. As in Theorem 2.2, take the measurable pullback (S_3, Σ_3) with the measurable mappings $k_j \colon S_3 \to S_j$, $j = 1, 2$.

Now let $(X, \Xi) := (S_3, \Sigma_3)$ and for $x \in X$, $j = 1, 2$, define

$$\begin{aligned} \mu_j^x &:= \tau_j^{k_j(x)} && j = 1, 2 \\ \mu_0^x &:= \tau_0^{h_1(k_1(x))} = \tau_0^{h_2(k_2(x))}. \end{aligned}$$

Since the maps k_j, h_j and $x \mapsto \tau_j^x$ are measurable, it follows that μ_j are subprobability kernels. By Theorem 2.2 there exists a semipullback μ_3 in the category of Radon subprobability kernels from $X = S_3$. For $A \in \Sigma_j$, $j = 1, 2$, and $x \in S_3$ we have

$$\mu_3^x(k_j^{-1}(A)) = \mu_j^x(A) = \tau_j^{k_j(x)}(A),$$

which means that μ_3 is also a semipullback in the LMP category. That concludes the proof for the case of a singleton label set L.

Now consider an arbitrary label set L. We have just proved that for each $a \in L$ there exists a semipullback (S_3, Σ_3, τ_a) in which S_3 and Σ_3 do not depend on a. But that means that $(S_3, \Sigma_3, \{\tau_a : a \in L\})$ is a semipullback in the category of the LMP labelled by L. ☐

5.2. **Universally measurable labelled Markov processes.** In Theorem 5.1 we assume that each measure $\tau_a(s, \cdot)$ is Radon. It may be more convenient to have instead a single restriction on the underlying space S, as in the next theorem.

Definition 5.2. A measurable space (S, Σ) is a *separable universally measurable space* if it is isomorphic to a universally measurable subset of a separable completely metrizable ("Polish") space with the trace of the Borel σ-algebra.

Theorem 5.3. *Consider the category in which objects are LMP* $(S, \Sigma, \{\tau_a : a \in L\})$ *such that* S *is a separable universally measurable space, with zigzag morphisms. In this category, every cospan has a semipullback.*

Proof. Let

$$(S_1, \Sigma_1, \{\tau_{1a} : a \in L\}) \to (S_0, \Sigma_0, \{\tau_{0a} : a \in L\}) \leftarrow (S_2, \Sigma_2, \{\tau_{2a} : a \in L\})$$

be a cospan of LMP with S_j separable universally measurable spaces. Then each (S_j, Σ_j) is isomorphic to some $(X_j, \mathfrak{B}(Y_j) \restriction X_j)$ where X_j is a universally measurable subset of a Polish space Y_j and $\mathfrak{B}(Y_j)$ is its Borel σ-algebra. Since Y_j is a Radon space, by [Fre13, 434F(c)] we conclude that every Borel measure on X_j is Radon.

Let $(S_3, \Sigma_3, \{\tau_{3a} : a \in L\})$ be a semipullback with the properties from Theorem 5.1. In particular, S_3 is a measurable subset of $S_1 \times S_2$. It remains to prove that S_3 is a separable universally measurable space. There exists a measurable isomorphism

$$f : (S_1 \times S_2, \Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2) \to (X_1 \times X_2, \mathfrak{B}(Y_1 \times Y_2) \restriction X_1 \times X_2),$$

$X_1 \times X_2$ is universally measurable by [Fre13, 434X(c)]. But then $(S_3, \Sigma_3) = (S_3, \Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2 \restriction S_3)$ is isomorphic to $(f(S_3), \mathfrak{B}(Y_1 \times Y_2) \restriction f(S_3))$, where $f(S_3) \in \mathfrak{B}(Y_1 \times Y_2) \restriction X_1 \times X_2$ since $S_3 \in \Sigma_1 \mathbin{\widehat{\otimes}} \Sigma_2$. □

In [ST11] it was asked whether behaviorally equivalent LMP over coanalytic spaces were bisimilar. We can answer this question affirmatively. Recall that a metric space is coanalytic if it is homeomorphic to the complement of an analytic subset of a Polish space. We say that a measurable space is *coanalytic* if it is isomorphic to the Borel space of a coanalytic metric space.

Corollary 5.4. *The category of LMP over coanalytic measurable spaces has semipullbacks.*

Proof. This follows by essentially the same argument for Theorem 5.3, showing that a cospan of coanalytic measurable spaces has a coanalytic pullback, and that coanalytic sets are universally measurable. □

In the same way, every analytic space with its Borel σ-algebra is a separable universally measurable space; hence we obtain Edalat's result [Eda99] as a corollary to Theorem 5.3 as well.

6. COUNTEREXAMPLES

The key assumption in previous sections is that each measure is defined on the Borel σ-algebra. The results no longer hold without that assumption, even for σ-algebras of subsets of $[0, 1]$. The counterexample in [ST11] uses a σ-algebra larger than the Borel σ-algebra on $[0, 1]$; we hint at this construction below. In the opposite direction, the following counterexample uses σ-algebras that are smaller but still large enough to separate the points of $[0, 1]$.

Example 6.1. Consider (S, Σ) to be the interval $[0, 1]$ with the countable-cocountable σ-algebra, and let $\mu_0 : \Sigma \to \{0, 1\}$ be the probability measure such that

$$\mu_0(Q) = 1 \iff Q \text{ has countable complement.}$$

Take $V := [0, \frac{1}{2}]$. It is straightforward to check that for any different $r_1, r_2 \in (0, 1)$, the following maps

$$\mu_i(Q) := \begin{cases} \mu(Q) & Q \in \Sigma \\ r_i & Q \in \Sigma_V \setminus \Sigma \text{ and } Q \setminus V \text{ is countable} \\ 1 - r_i & \text{otherwise,} \end{cases}$$

are probability measures that extend μ_0 to $\Sigma_V := \sigma(\Sigma \cup \{V\})$.

By using these probability spaces we can replicate the idea of [ST11, Thm. 12] to obtain a cospan of LMP that can not be completed to a commutative square. We now sketch the construction. Fix $s_0 \in S$ and define LMP $\mathbf{S}_i := (S, \Sigma_i, \tau_i)$, with $\Sigma_1 = \Sigma_2 := \Sigma_V$ and $\Sigma_0 := \Sigma$, and

$$\tau_i(s, A) := \begin{cases} 1 & s \neq s_0 \text{ and } s_0 \in A \\ \mu_i(A) & s = s_0 \\ 0 & \text{otherwise} \end{cases}$$

for $i = 0, 1, 2$, every $s \in S$, and A in the corresponding σ-algebra.

The identity maps $Id_S : \mathbf{S}_i \to \mathbf{S}_0$ form a cospan of zigzags, and it can be seen that there are no $\mathbf{S}$ and zigzag maps $h_i : \mathbf{S} \to \mathbf{S}_i$ $(i = 1, 2)$ completing that cospan to a semipullback.

Acknowledgment

We would like to thank Zoltán Vidnyánsky for initial suggestions regarding Example 6.1. The second author would also like to thank Prof. S.M. Srivastava and Prof. S. Todorčević for enlightening discussions around this problem.

References

[BDEP97] Richard Blute, Josee Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled markov processes. In *Proceedings of the 12th Annual IEEE Symposium on Logic in Computer Science*, LICS '97, page 149, USA, 1997. IEEE Computer Society.

[DDLP06] Vincent Danos, Josée Desharnais, François Laviolette, and Prakash Panangaden. Bisimulation and cocongruence for probabilistic systems. *Inf. Comput.*, 204(4):503–523, April 2006.

[DEP98] J. Desharnais, A. Edalat, and P. Panangaden. A logical characterization of bisimulation for labeled markov processes. In *Proceedings of the 13th Annual IEEE Symposium on Logic in Computer Science*, LICS '98, page 478, USA, 1998. IEEE Computer Society.

[DEP02] Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. *Inf. Comput.*, 179(2):163–193, 2002.

[Des99] Josée Desharnais. *Labeled Markov Process.* PhD thesis, McGill University, 1999.

[Dob05] Ernst-Erich Doberkat. Semi-pullbacks for stochastic relations over analytic spaces. *Mathematical. Structures in Comp. Sci.*, 15(4):647–670, August 2005.

[Dob09] Ernst-Erich Doberkat. *Stochastic Coalgebraic Logic.* Monographs in theoretical computer science. Springer, 2009.

[DS57] Nelson Dunford and Jacob T. Schwartz. *Linear Operators, 1: General Theory.* Pure and applied mathematics. Academic press, London, New York, Sydney, 1957.

[Eda99] Abbas Edalat. Semi-pullbacks and bisimulation in categories of Markov processes. *Mathematical Structures in Comp. Sci.*, 9(5):523–543, 1999.

[Fre11] David H. Fremlin. *Measure Theory vol. 1.* Torres Fremlin, second edition, 2011.

[Fre13] David H. Fremlin. *Measure Theory vol. 4.* Torres Fremlin, second edition, 2013.

[JNW96] André Joyal, Mogens Nielsen, and Glynn Winskel. Bisimulation from open maps. *Information and Computation*, 127(2):164–185, 1996.

[Kal02] Olav Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[MRN53] Edward Marczewski and Czesław Ryll-Nardzewski. Remarks on the compactness and non direct products of measures. *Fundamenta Mathematicae*, 40(1):165–170, 1953.

[ST11] Pedro Sánchez Terraf. Unprovability of the logical characterization of bisimulation. *Information and Computation*, 209(7):1048–1056, 2011.