Bisimulation on Markov Processes over Arbitrary Measurable Spaces

Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare

Department of Computer Science, Aalborg University, Denmark
{grbacci,giovbacci,kgl,mardare}@cs.aau.dk

Abstract. We introduce a notion of bisimulation on labelled Markov Processes over generic measurable spaces in terms of arbitrary binary relations. Our notion of bisimulation is proven to coincide with the coalgebraic definition of Aczel and Mendler in terms of the Giry functor, which associates with a measurable space its collection of (sub)probability measures. This coalgebraic formulation allows one to relate the concepts of bisimulation and event bisimulation of Danos et al. (i.e., cocongruence) by means of a formal adjunction between the category of bisimulations and a (full sub)category of cocongruences, which gives new insights about the real categorical nature of their results. As a corollary, we obtain sufficient conditions under which state and event bisimilarity coincide.

1 Introduction

The first notion of bisimulation for Markov processes, that are, probabilistic systems with a generic measurable space of states, has been defined categorically by Blute, Desharnais, Edalat, and Panangaden in [1] as a span of zig-zag morphisms, i.e., measurable surjective maps preserving the transition structure of the process. It turned out to be very difficult to prove that the induced notion of bisimilarity is an equivalence and this was only solved under the technical assumption that the state space is analytic. Under this hypothesis in [5] it was proposed a neat logical characterization of bisimilarity, using a much simpler logic than the one previously used for the discrete case.

In [3], Danos, Desharnais, Laviolette, and Panangaden introduced a notion alternative to that of bisimulation, the so called *event bisimulation* (i.e., cocongruence). From a categorical perspective the novelty was that they worked with cospans rather than spans. Remarkably, they were able to give a logical characterization of event bisimilarity without assuming the state space of the Markov processes to be analytic. In addition, they proved that, for analytic spaces, event and state bisimilarity coincide.

It has been always an open question whether the analyticity assumption on the state space can be dropped. In this paper we make a step forward in this direction, providing a notion of state bisimulation on Markov processes over arbitrary measurable spaces. This definition is based on a characterization due to Desharnais et al. [6,7], which mimics the definition probabilistic bisimulation of Larsen and Skou [11] by adding a few measure-theoretic conditions to deal with the fact that some subsets may not be measurable. Their characterization was given assuming that the bisimulation relation is an equivalence, instead, our definition is expressed in terms of arbitrary binary relations. This mild generalization to binary relation allows us to prove that our definition coincides with the coalgebraic notion of bisimulation of Aczel and Mendler in terms of the Giry functor [8], which associates with a measurable space its collection of subprobability measures. A similar result was proven by de Vink and Rutten [4] who studied Markov processes on ultrametric spaces. However, in [4] the characterization was established assuming that bisimulation relations have a Borel decomposition (which is not a mild assumption). Our proof does not need the existence of Borel decompositions and can be used to refine the result in [4].

Our characterization of probabilistic bisimulation is weaker than the original proposal in [1,5] (which requires the relation to be an equivalence) and weaker than the definition in [15] (which only requires the relation to be reflexive). However, we show that our definition is a generalization of both, and we prove that, when we restrict our definition on the case of single Markov processes, all the results continue to hold, in particular that the class of bisimulations is closed under union and that bisimilarity it is an equivalence.

Another contribution of this paper is a formal coalgebraic analysis of the relationships between the notions of bisimulation and cocongruence on Markov processes. This is done by lifting a standard adjunction that occurs between spans and cospans in categories with pushouts and pullbacks. The lifting to the categories of bisimulations and cocongruences is very simple when the behaviour functor weakly-preserves pullbacks. Although Viglizzo [16] proved that the Giry functor does not enjoy this property, we managed to show that the lifting is possible when we restrict to a suitable (full-)subcategory of congruences. This restriction cannot be avoided, since Terraf [13,14] showed that state and event bisimilarity do not coincide on Markov processes over arbitrary measurable spaces. As a corollary, we establish sufficient conditions under which bisimulation and cocongruence coincide, and as an aside result, this adjunction explains at a more abstract categorical level all the results in [3] that relate state and event bisimulations. To the best of our knowledge, this result is new and, together with the counterexample given by Terraf, completes the comparison between these two notions of equivalence between Markov processes over arbitrary measurable spaces.

2 Preliminaries

Binary Relations. For a binary relation $\mathcal{R} \subseteq X \times Y$ we use $\pi_X \colon \mathcal{R} \to X$ and $\pi_Y \colon \mathcal{R} \to Y$ to denote the canonical projections of $\mathcal{R}$ on X and Y, respectively. Given $\mathcal{R} \subseteq X \times Y$ and $\mathcal{S} \subseteq Y \times Z$, we denote by $\mathcal{R}^{-1} = \{(y,x) \mid (x,y) \in \mathcal{R}\}$ the inverse of $\mathcal{R}$, and by $\mathcal{R};\mathcal{S} = \{(x,z) \mid (x,y) \in \mathcal{R} \text{ and } (y,z) \in \mathcal{S} \text{ for some } y \in Y\}$ the composition of $\mathcal{R}$ and $\mathcal{S}$. We say that $\mathcal{R} \subseteq X \times Y$ is *z-closed* if, for all $x, x' \in X$ and $y, y' \in Y$, $(x,y), (x',y), (x',y') \in \mathcal{R}$ implies $(x,y') \in \mathcal{R}$; and we denote by $\mathcal{R}^*$, the *z-closure* of $\mathcal{R}$, i.e., the least z-closed relation that contains $\mathcal{R}$. Note that any equivalence relation is z-closed, indeed one can informally see the z-closure property as a generalization of transitive closure on binary relations.

Measure Theory. A *field* over a set X is a nonempty family of subsets of X closed under complement and union. A σ-algebra over a set X is a field Σ_X such that is closed under countable union. The pair (X, Σ_X) is a *measurable space* and the elements of Σ_X are called *measurable sets*. A *generator* $\mathcal{F}$ for Σ_X is a family of subsets of X such that the smallest σ-algebra containing $\mathcal{F}$, denoted by $\sigma(\mathcal{F})$, is Σ_X.

Let (X, Σ_X), (Y, Σ_Y) be measurable spaces. A function $f : X \to Y$ is called *measurable* if $f^{-1}(E) = \{x \mid f(x) \in E\} \in \Sigma_X$, for all $E \in \Sigma_Y$ (notably, if Σ_Y is generated by $\mathcal{F}$, f is measurable iff $f^{-1}(F) \in \Sigma_X$, for all $F \in \mathcal{F}$). The family $\{E \subseteq Y \mid f^{-1}(E) \in \Sigma_X\}$, called the *final σ-algebra w.r.t. f*, is the largest σ-algebra over Y that renders f measurable. Dually, the family $\{f^{-1}(E) \mid E \in \Sigma_Y\}$ is called *initial σ-algebra w.r.t f*, and it is the smallest σ-algebra over X that makes f measurable. Initial and final σ-algebras generalize to families of maps in the obvious way.

A *measure* on (X, Σ_X) is a σ-additive function $\mu : \Sigma_X \to [0, \infty]$, that is, $\mu(\bigcup_{i \in I} E_i) = \sum_{i \in I} \mu(E_i)$ for all countable collections $\{E_i\}_{i \in I}$ of pairwise disjoint measurable sets. A measure $\mu : \Sigma_X \to [0, \infty]$ is of *(sub)-probability* if $\mu(X) = 1$ (resp. ≤ 1), is *finite* if $\mu(X) < \infty$, and is *σ-finite* if there exists a countable cover $\{E_i\}_{i \in I} \subseteq \Sigma_X$ of X, i.e., $\bigcup_{i \in I} E_i = X$, of measurable sets such that $\mu(E_i) < \infty$, for all $i \in I$. A *pre-measure* on a field $\mathcal{F}$ is a finitely additive function $\mu_0 : \mathcal{F} \to [0, \infty]$ with the additional property that whenever $F_0, F_1, F_2, \ldots$ is a countable disjoint collection sets in $\mathcal{F}$ such that $\bigcup_{n \in \mathbb{N}} F_n \in \mathcal{F}$, then $\mu_0(\bigcup_{n \in \mathbb{N}} F_n) = \sum_{n \in \mathbb{N}} \mu_0(F_n)$.

Coalgebras, Bisimulations, and Cocongruences. Let $F : \mathbf{C} \to \mathbf{C}$ be an endofunctor on a category $\mathbf{C}$. An *F-coalgebra* is a pair (X, α) consisting of an object X, called *carrier*, and an arrow $\alpha : X \to FX$ in $\mathbf{C}$, called *coalgebra structure*. An *homomorphism* between two F-coalgebras (X, α) and (Y, β) is an arrow $f : X \to Y$ in $\mathbf{C}$ such that $\alpha \circ f = Ff \circ \beta$. F-coalgebras and homomorphisms between them form a category, denoted by F-**coalg**.

An *F-bisimulation* (R, f, g) between two F-coalgebras (X, α) and (Y, β) is a span $X \xleftarrow{f} R \xrightarrow{g} Y$ with jointly monic legs (a *monic span*) such that there exists a coalgebra structure $\gamma : R \to FR$ making f and g homomorphisms of F-coalgebras. Two F-coalgebras are *bisimilar* if there is a bisimulation between them. A notion alternative to bisimilarity that has been proven very useful in reasoning about probabilistic systems [3], is *cocongruence*. An *F-cocongruence* (K, f, g) between F-coalgebras (X, α) and (Y, β) is an cospan $X \xrightarrow{f} K \xleftarrow{g} Y$ with jointly epic legs (an *epic cospan*) such that there exists a (unique) coalgebra structure $\kappa : K \to FK$ on K such that f and g are F-homomorphisms.

3 Labelled Markov Processes and Bisimulation

In this section we recall the notions of labelled Markov kernels and processes, using a notation similar to [2], then we propose a general definition of (state) bisimulation between pairs of Markov kernels.

Let (X, Σ) be a measurable space. We denote by $\Delta(X, \Sigma)$ the set of all subprobability measures over (X, Σ). For each $E \in \Sigma$ there is a canonical evaluation function $ev_E \colon \Delta(X, \Sigma) \to [0,1]$, defined by $ev_E(\mu) = \mu(E)$, for all $\mu \in \Delta(X, \Sigma)$, and called *evaluation at* E. By means of these evaluation maps, $\Delta(X, \Sigma)$ can be organized into a measurable space $(\Delta(X, \Sigma), \Sigma_{\Delta(X,\Sigma)})$, where $\Sigma_{\Delta(X,\Sigma)}$ the initial σ-algebra with respect to $\{ev_E \mid E \in \Sigma\}$, i.e., the smallest σ-algebra making ev_E measurable w.r.t. the Borel σ-algebra on $[0,1]$, for all $E \in \Sigma$. Equivalently, $\Sigma_{\Delta(X,\Sigma)}$ can be also generated by the sets $L_q(E) = \{\mu \in \Delta(X, \Sigma) \mid \mu(E) \geq q\}$, for $q \in [0,1] \cap \mathbb{Q}$ and $E \in \Sigma$ (see [9]).

Definition 1. *Let (X, Σ) be a measurable space and L a set of action labels. An* L-labelled Markov kernel *is a tuple $\mathcal{M} = (X, \Sigma, \{\theta_a\}_{a \in L})$ where, for all $a \in L$*

$$\theta_a \colon X \to \Delta(X, \Sigma)$$

is a measurable function, called Markov a-transition function*. An L-labelled Markov kernel $\mathcal{M}$ with a distinguished initial state $x \in X$, is said* Markov process*, and it is denoted by $(\mathcal{M}, x)$.*

The labels in L constitute all possible interactions of the processes with the environment: for an action $a \in L$, a current state $x \in X$, and a measurable set $E \in \Sigma_X$, the value $\theta_a(x)(E)$ represents the probability of taking an a-transition from x to arbitrary elements in E.

Before presenting the notions of bisimulation and bisimilarity between labelled Markov kernels, we introduce some preliminary notation.

Definition 2 ($\mathcal{R}$-closed pair). *Let $\mathcal{R} \subseteq X \times Y$ be a relation on the sets X and Y, and $E \subseteq X$, $F \subseteq Y$. A pair (E, F) is* $\mathcal{R}$-closed *if $\mathcal{R} \cap (E \times Y) = \mathcal{R} \cap (X \times F)$.*

A pair (E, F) is $\mathcal{R}$-closed iff $\pi_X^{-1}(E) = \pi_Y^{-1}(F)$, where π_X, π_Y are the canonical projections on X and Y, respectively. The following lemmas will be useful later in the paper, and are direct consequences of the definition.

Lemma 3. *Let $\mathcal{R}' \subseteq \mathcal{R} \in X \times Y$. If (E, F) is $\mathcal{R}$-closed, then (E, F) is $\mathcal{R}'$-closed.*

Lemma 4. *Let $\mathcal{R} \subseteq X \times X$ be an equivalence relation. If (E, F) is $\mathcal{R}$-closed then $E = F$, moreover E is an union of $\mathcal{R}$-equivalence classes.*

Definition 5 (Bisimulation and Bisimilarity). *Let $\mathcal{M} = (X, \Sigma_X, \{\alpha_a\}_{a \in L})$ and $\mathcal{N} = (Y, \Sigma_Y, \{\beta_a\}_{a \in L})$ be two L-labelled Markov kernels. A binary relation $\mathcal{R} \subseteq X \times Y$ is a* (state) bisimulation *between $\mathcal{M}$ and $\mathcal{N}$ if, for all $(x, y) \in \mathcal{R}$, $a \in L$, and any pair $E \in \Sigma_X$ and $F \in \Sigma_Y$ such that (E, F) is $\mathcal{R}$-closed*

$$\alpha_a(x)(E) = \beta_a(y)(F)\,.$$

Two L-labelled Markov processes $(\mathcal{M}, x)$ *and* $(\mathcal{N}, y)$ *are* (state) bisimilar*, written* $x \sim_{(\mathcal{M},\mathcal{N})} y$*, if the initial states* x *and* y *are related by some bisimulation* $\mathcal{R}$ *between* $\mathcal{M}$ *and* $\mathcal{N}$*.*

Originally, the definition of bisimulation between labelled Markov processes was given in terms of spans of zig-zag morphisms [1]. Later, Danos et al. [3] gave a more direct relational definition, called *state bisimulation*, characterizing the original zig-zag definition in the case of equivalence relations. Their definition differs from Definition 5 only on how $\mathcal{R}$-closed subsets are characterized and in that they require bisimulations to be equivalence relations[1]. Later, in [15], van Breugel et al. proposed a weaker definition, where bisimulation relations are only required to be reflexive. Definition 5 subsumes the definitions of state bisimulation given in [3] and [15] (this is a direct consequence of Lemma 4).

Proposition 6. *Let* $\mathcal{F}$ *be a family of bisimulation relations between the L-labelled Markov kernels* $\mathcal{M} = (X, \Sigma_X, \{\alpha_a\}_{a \in L})$ *and* $\mathcal{N} = (Y, \Sigma_Y, \{\beta_a\}_{a \in L})$*. Then* $\bigcup \mathcal{F}$ *is a bisimulation.*

Proof. Let $(x, y) \in \bigcup \mathcal{F}$, $a \in L$, and $E \in \Sigma_X$ and $F \in \Sigma_Y$ such that (E, F) is $\bigcup \mathcal{F}$-closed. By $(x, y) \in \bigcup \mathcal{F}$, there exists a bisimulation $\mathcal{R} \subseteq X \times Y$ such that $(x, y) \in \mathcal{R}$. Obviously $\mathcal{R} \subseteq \bigcup \mathcal{F}$, thus, by Lemma 3, (E, F) is $\mathcal{R}$-closed. Since $(x, y) \in \mathcal{R}$ and $\mathcal{R}$ is a bisimulation, we have $\alpha_a(x)(E) = \beta_a(y)(F)$. □

Corollary 7. $\sim_{(\mathcal{M},\mathcal{N})}$ *is the largest bisimulation between* $\mathcal{M}$ *and* $\mathcal{N}$*.*

Proof. By definition $\sim_{(\mathcal{M},\mathcal{N})} = \bigcup \{\mathcal{R} \mid \mathcal{R}$ bisimulation between $\mathcal{M}$ and $\mathcal{N}\}$, thus, by Lemma 6 it is a bisimulation and in particular it is the largest one. □

The following results proves that, if bisimilarity is restricted to single labelled Markov kernels, then it is an equivalence.

Theorem 8 (Equivalence). *Let* $\mathcal{M} = (X, \Sigma, \{\theta_a\}_{a \in L})$ *be an L-labelled Markov kernel. Then, the bisimilarity relation* $\sim_{\mathcal{M}} \subseteq X \times X$ *on* $\mathcal{M}$ *is an equivalence.*

Proof. Symmetry: if $\mathcal{R}$ is a bisimulation, so is $\mathcal{R}^{-1}$. Reflexivity: we prove that the identity relation Id_X is a bisimulation, i.e., for all $x \in X$, $a \in A$, and $E, F \in \Sigma$ such that (E, F) is Δ_X-closed, $\theta_a(x)(E) = \theta_a(x)(F)$. This holds trivially by Lemma 4, since Id_X is an equivalence. Transitivity: it suffices to show that, given $\mathcal{R}_1$ and $\mathcal{R}_2$ bisimulations on $\mathcal{M}$, there exists a bisimulation $\mathcal{R}$ on $\mathcal{M}$ that *contains* $\mathcal{R}_1; \mathcal{R}_2$. Let $\mathcal{R}$ be the (unique) smallest equivalence relation containing $\mathcal{R}_1 \cup \mathcal{R}_2$. $\mathcal{R}$ can be defined as $\mathcal{R} = Id_X \cup \bigcup_{n \in \mathbb{N}} \mathcal{S}_n$, where

$$\mathcal{S}_0 \triangleq \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_1^{-1} \cup \mathcal{R}_2^{-1} \qquad\qquad \mathcal{S}_{n+1} \triangleq \mathcal{S}_n; \mathcal{S}_n .$$

[1] Actually, in [3], the definition of state bisimulation is given without mentioning that the relation must be an equivalence, but without that requirement many subsequent results do not hold (e.g. Lemmas 4.1, 4.6, 4.8, Proposition 4.7, and Corollary 4.9). However, looking at the proofs it seems that they were imposing this condition.

It is easy to see that $\mathcal{R}_1;\mathcal{R}_2 \subseteq \mathcal{R}$. We are left to show that $\mathcal{R}$ is a bisimulation. By Lemma 4, it suffices to prove that for all $a \in L$, and measurable set $E \in \Sigma$ such that (E, E) is $\mathcal{R}$-closed, the following implication holds

$$(x, y) \in \mathcal{R} \implies \theta_a(x)(E) = \theta_a(y)(E)\,. \tag{1}$$

If $(x, y) \in \mathcal{R}$, then $(x, y) \in Id_X$ or $(x, y) \in \mathcal{S}_n$ for some $n \geq 0$. If $(x, y) \in Id_X$ then $x = y$, hence (1) trivially holds. Now we show, by induction on $n \geq 0$, that

$$(x, y) \in \mathcal{S}_n \implies \theta_a(x)(E) = \theta_a(y)(E)\,. \tag{2}$$

Base case ($n = 0$): let $j \in \{1, 2\}$. If $(x, y) \in \mathcal{R}_j$, (2) holds since, by hypothesis that $\mathcal{R}_j$ is a bisimulation and by Lemma 3, (E, E) is $\mathcal{R}_j$-closed. If $(x, y) \in \mathcal{R}_j^{-1}$ we have $(y, x) \in \mathcal{R}_j$, hence (2) holds again.
Inductive case ($n + 1$): if $(x, y) \in \mathcal{S}_{n+1}$, then there exists some $z \in X$ such that $(x, z) \in \mathcal{S}_n$ and $(z, y) \in \mathcal{S}_n$. Then, applying the inductive hypothesis twice, we have $\theta_a(x)(E) = \theta_a(z)(E)$ and $\theta_a(z)(E) = \theta_a(y)(E)$, from which (2) follows. □

Remark 9. Theorem 8 has been already stated by van Breugel et al. [15] considering a more restrictive definition of bisimulation than the one given in Definition 1. Although the result is not new, we put the proof here to show that it can be carried out in a much simpler way.

Moreover, notice that, in the proof of Theorem 8 transitivity is verified using a strategy that avoids to prove that bisimulation relations are closed under composition. Indeed, this would have required that (semi-)pullbacks of relations in **Meas** are weakly preserved by the Giry functor. Recently, in [13,14] Terraf showed that this is not the case. The proof is based on the existence of a non-Lebesgue-measurable set V in the open unit interval $(0, 1)$, which is used to define two measures on the σ-algebra extended with V such that they differ in this set. In the light of this, the simplicity of the proof of Theorem 8 is even more remarkable. ■

By Corollary 7 we have the following direct characterization for bisimilarity.

Proposition 10. *Let $\mathcal{M} = (X, \Sigma, \{\theta_a\}_{a \in L})$ be an L-labelled Markov kernel, then, for $x, y \in X$, $x \sim_{\mathcal{M}} y$ iff for all $a \in L$ and $E \in \Sigma$ such that (E, E) is $\sim_{\mathcal{M}}$-closed, $\theta_a(x)(E) = \theta_a(y)(E)$.*

Interestingly, Theorem 8 can be alternatively proven as a corollary of this result; indeed the above characterization implies that $\sim_{\mathcal{M}}$ is an equivalence relation.

Another important property, which will be used later in the paper, is that the class of bisimulations is closed by z-closure.

Lemma 11 ([4]). *If R is a state bisimulation between $(X, \Sigma_X, \{\alpha_a\}_{a \in L})$ and $(Y, \Sigma_Y, \{\beta_a\}_{a \in L})$, then so is R^*, the z-closure of R.*

4 Characterization of the Coalgebraic Bisimulation

In this section, we prove that the notion of state bisimulation (Definition 5) coincides with the abstract coalgebraic definition of Aczel and Mendler.

In order to model Markov processes as coalgebras the most natural choice for a category is **Meas**, the category of measurable spaces and measurable maps. This category is complete and cocomplete: limits and colimits are obtained as in **Set** and endowed, respectively, with initial and final σ-algebra w.r.t. their cone and cocone maps. Hereafter, for the sake of readability, we adopt a notation that makes no distinction between measurable spaces and objects in **Meas**: by the boldface symbol $\mathbf{X}$ we denote the measurable space (X, Σ_X) (the subscript is used accordingly).

Let L be a set of action labels. L-labelled Markov kernels are standardly modeled as Δ^L-coalgebras, where $(\cdot)^L \colon \mathbf{Meas} \to \mathbf{Meas}$ is the exponential functor and $\Delta \colon \mathbf{Meas} \to \mathbf{Meas}$ is the Giry functor acting on objects $\mathbf{X}$ and arrows $f \colon \mathbf{X} \to \mathbf{Y}$ as follows, for $\mu \in \Delta(X, \Sigma_X)$

$$\Delta\mathbf{X} = (\Delta(X, \Sigma_X), \Sigma_{\Delta(X,\Sigma_X)}) \qquad\qquad \Delta(f)(\mu) = \mu \circ f^{-1} \,.$$

It is folklore that L-labelled Markov kernels and Δ^L-coalgebras coincide. For $\mathbf{X} = (X, \Sigma_X)$ a measurable space, the correspondence is given by

$$(X, \Sigma_X, \{\theta_a\}_{a\in L}) \mapsto (\mathbf{X}, \alpha \colon \mathbf{X} \to \Delta^L\mathbf{X}) \,, \qquad \text{where } \alpha(x)(a) = \theta_a(x) \,, \tag{3}$$

$$(\mathbf{X}, \alpha \colon \mathbf{X} \to \Delta^L\mathbf{X}) \mapsto (X, \Sigma_X, \{ev_a \circ \alpha\}_{a\in L}) \,. \tag{4}$$

Where, in (3) α is measurable since $ev_E \circ ev_a \circ \alpha$ is measurable, for all $a \in L$ and $E \in \Sigma$, and in (4), $ev_a \circ \alpha$ is measurable since is the composite of measurable functions. Hereafter, we will make no distinction between Δ^L-coalgebras and L-labelled Markov kernels, and the correspondence above will be used without reference.

Next we relate the notion of Δ^L-bisimulation to the notion of state bisimulation. Recall that in categories with binary products monic spans (R, f, g) are in one-to-one correspondence with monic arrows $R \to X \times Y$. Thus, without loss of generality, we restrict our attention only to relations $R \subseteq X \times Y$ with measurable canonical projections $\pi_X \colon \mathbf{R} \to \mathbf{X}$ and $\pi_Y \colon \mathbf{R} \to \mathbf{Y}$.

Proposition 12. *Let $(\mathbf{X}, \alpha)$ and $(\mathbf{Y}, \beta)$ be Δ^L-coalgebras and $(\mathbf{R}, \pi_X, \pi_Y)$ be a Δ^L-bisimulation between them. Then, $R \subseteq X \times Y$ is a state bisimulation between $(X, \Sigma_X, \{ev_a \circ \alpha\}_{a\in L})$ and $(Y, \Sigma_Y, \{ev_a \circ \beta\}_{a\in L})$.*

Proof. We have to show that for all $(x, y) \in R$, $a \in L$ and $E \in \Sigma_X$, $F \in \Sigma_Y$ such that (E, F) is R-closed, $\alpha(x)(a)(E) = \beta(y)(a)(F)$. Since $(\mathbf{R}, \pi_X, \pi_Y)$ is a Δ^L-bisimulation, there exists a coalgebraic structure $\gamma \colon \mathbf{R} \to \Delta^L\mathbf{R}$ on $\mathbf{R}$ such that π_X and π_Y are Δ^L-homomorphisms.

$$\begin{aligned}
\alpha(x)(a)(E) &= (\alpha \circ \pi_X)(x,y)(a)(E) && \text{(by def. } \pi_X\text{)} \\
&= (\Delta^L \pi_X \circ \gamma)(x,y)(a)(E) && \text{(by } \Delta^L\text{-homomorphism)} \\
&= \Delta \pi_X(\gamma(x,y)(a))(E) && \text{(by def. } Id^L\text{)} \\
&= \gamma(x,y)(a) \circ \pi_X^{-1}(E) && \text{(by def. } \Delta\text{)} \\
&= \gamma(x,y)(a) \circ \pi_Y^{-1}(F) && \text{(by } (E,F)\ R\text{-closed)} \\
&= \beta(y)(a)(F) && \text{(by reversing the previous steps)}
\end{aligned}$$

□

The converse of Proposition 12 is more intricate, and we need some preliminary work involving techniques from measure theory in order to formally define a suitable mediating measurable Δ^L-coalgebra structure.

Proposition 13. *Let (X, Σ_X) and (Y, Σ_Y) be measurable spaces, $r_1 \colon R \to X$, $r_2 \colon R \to Y$ be surjective maps, and Σ_R be the initial σ-algebra on R w.r.t. r_1 and r_2, i.e., $\Sigma_R = \sigma(\{r_1^{-1}(E) \mid E \in \Sigma_X\} \cup \{r_2^{-1}(F) \mid F \in \Sigma_Y\})$.*

Then, for any pair of measures μ on (X, Σ_X) and ν on (Y, Σ_Y), such that, for all $E \in \Sigma_X$ and $F \in \Sigma_Y$,

$$\textit{if } r_1^{-1}(E) = r_2^{-1}(F)\,, \ \textit{then } \mu(E) = \nu(F)\,,$$

there exists a measure $\mu \wedge \nu$ on (R, Σ_R) such that, for all $E \in \Sigma_X$ and $F \in \Sigma_Y$

$$(\mu \wedge \nu)(r_1^{-1}(E)) = \mu(E) \qquad \textit{and} \qquad (\mu \wedge \nu)(r_2^{-1}(F)) = \mu(F)\,.$$

Moreover, if μ and ν are σ-finite, $\mu \wedge \nu$ is unique.

The existence and uniqueness of $\mu \wedge \nu$ is guaranteed by the Hahn-Kolmogorov extension theorem. Note that, the conditions imposed on μ and ν are necessary for $\mu \wedge \nu$ to be well-defined (see the appendix for a detailed proof).

Thank to Proposition 13 we can prove the following result, which concludes the correspondence between state bisimulation and Δ^L-bisimulation.

Proposition 14. *Let $R \subseteq X \times Y$ be a state bisimulation between the L-labelled Markov kernels $(X, \Sigma_X, \{\alpha_a\}_{a \in L})$ and $(Y, \Sigma_Y, \{\beta_a\}_{a \in L})$. Then $(\mathbf{R}, \pi_X, \pi_Y)$ is a Δ^L-bisimulation between $(\mathbf{X}, \alpha)$ and $(\mathbf{Y}, \beta)$, where Σ_R is initial w.r.t. π_X and π_Y, and for all $a \in L$, $x \in X$, and $y \in Y$, $\alpha(x)(a) = \alpha_a(x)$ and $\beta(y)(a) = \beta_a(y)$.*

Proof. We have to provide a measurable coalgebra structure $\gamma \colon \mathbf{R} \to \Delta^L \mathbf{R}$ making $\pi_X \colon \mathbf{R} \to \mathbf{X}$ and $\pi_Y \colon \mathbf{R} \to \mathbf{Y}$ Δ^L-homomorphisms.

First notice that, without loss of generality we can assume π_X and π_Y to be surjective. Indeed, if it is not so, we can factorize π_X and π_Y as $\pi_X = m_X \circ e_X$, $\pi_Y = m_Y \circ e_Y$ such that e_X, e_Y are surjective (epic) and m_X, m_Y are injective (monic), to obtain the following commuting diagrams

$$\begin{array}{ccccccccc}
\mathbf{X} & \xleftarrow{m_X} & \mathbf{X}' & \xleftarrow{e_X} & \mathbf{R} & \xrightarrow{e_Y} & \mathbf{Y}' & \xrightarrow{m_Y} & \mathbf{Y} \\
\downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \alpha'} & & & & \downarrow{\scriptstyle \beta'} & & \downarrow{\scriptstyle \beta} \\
\Delta^L\mathbf{X} & \xleftarrow{\Delta m_X} & \Delta^L\mathbf{X}' & \xleftarrow{\Delta e_X} & \Delta^L\mathbf{R} & \xrightarrow{\Delta e_Y} & \Delta^L\mathbf{Y}' & \xrightarrow{\Delta m_Y} & \Delta^L\mathbf{Y}
\end{array}$$

where $\mathbf{X}' = (X', \Sigma_{X'})$, X' is the image $\pi_X(R)$, $\Sigma_{X'}$ is the initial σ-algebra w.r.t. m_X (i.e., $\{m_X^{-1}(E) \mid E \in \Sigma_X\}$), and $\alpha'(x')(a)(m_X^{-1}(E)) = \alpha(m_X(i))(a)(E)$, for all $x' \in X'$ and $E \in \Sigma_X$; (similarly for $\mathbf{Y}$ and β'). Therefore, to find the coalgebra structure γ for making e_X and e_Y Δ^L-homomorphisms, it solves the problem for π_X and π_Y as well.

Recall that (E, F) is R-closed iff $\pi_X^{-1}(E) = \pi_Y^{-1}(F)$. By hypothesis, R is a state bisimulation, so that, for all $(x, y) \in R$, $E \in \Sigma_X$, $F \in \Sigma_Y$, and $a \in L$,

$$\text{if } \pi_X^{-1}(E) = \pi_Y^{-1}(F)\,, \qquad \text{then} \qquad \alpha(a)(x)(E) = \beta(a)(y)(F)\,.$$

For all $(x, y) \in R$ and $a \in L$, by Proposition 13, we define $\gamma((x, y))(a)$ as the unique (sub)probability measure on (R, Σ_R), s.t., for all $E \in \Sigma_X$ and $F \in \Sigma_Y$,

$$\gamma((x,y))(a)(\pi_X^{-1}(E)) = \alpha(x)(a)(E)\,, \quad \text{and} \quad \gamma((x,y))(a)(\pi_Y^{-1}(F)) = \beta(y)(a)(F)\,.$$

By definition of γ, both π_X and π_Y are Δ^L-homomorphisms, indeed,

$$\begin{aligned}
(\Delta^L\pi_X \circ \gamma)((x,y))(a)(E) &= \gamma((x,y))(a)(\pi_X^{-1}(E)) && \text{(by def. } \Delta^L) \\
&= \alpha(x)(a)(E) && \text{(by def. } \gamma) \\
&= (\alpha \circ \pi_X)((x,y))(a)(E)\,. && \text{(by def. } \pi_X)
\end{aligned}$$

for all $(x, y) \in R$, $a \in A$, and $E \in \Sigma_X$. The proof for π_Y is similar.

To prove that γ is measurable, by [9, Lemma 4.5] it suffices to show that for any finite union of the form $S = \bigcup_{i=0}^{k}(\pi_X^{-1}(E_i) \cap \pi_Y^{-1}(F_i))$, where $E_i \in \Sigma_X$ and $F_i \in \Sigma_Y$, for $0 \le i \le k$, $(ev_a \circ \gamma)^{-1}(L_q(S)) \in \Sigma_R$. We may assume, without loss of generality, that S is given as a disjoint union (otherwise we may represent it taking a disjoint refinement), and by finite additivity, it suffices to consider only the case $S = \pi_X^{-1}(E') \cap \pi_Y^{-1}(F')$, for some $E' \in \Sigma_X$ and $F' \in \Sigma_Y$. According to the definition of γ (see Proposition 13), we have to consider three cases:

– if $\exists E \in \Sigma_X$ such that $S = \pi_X^{-1}(E)$, then

$$\begin{aligned}
(ev_a \circ \gamma)^{-1}(L_q(S)) &= \{(x,y) \in R \mid \gamma((x,y))(a) \in L_q(S)\} && \text{(by inverse image)} \\
&= \{(x,y) \in R \mid \gamma((x,y))(a)(S) \ge q\} && \text{(by def. } L_q(\cdot)) \\
&= \{(x,y) \in R \mid \alpha(x)(a)(E) \ge q\} && \text{(by def. } \gamma) \\
&= (\pi_X \circ ev_a \circ \alpha)^{-1}(L_q(E)) && \text{(by inverse image)}
\end{aligned}$$

which is a measurable set in Σ_R, since π_X, ev_a, α are measurable;

– if $\exists F \in \Sigma_Y$, $S = \pi_Y^{-1}(F)$, and $\forall E \in \Sigma_X$, $S \neq \pi_X^{-1}(E)$ one can proceed similarly, replacing in the above derivation E, π_X, and $\alpha(x)$ by F, π_Y and $\beta(y)$, respectively.
– if $\forall E \in \Sigma_X, F \in \Sigma_Y$, $S \neq \pi_X^{-1}(E)$ and $S \neq \pi_Y^{-1}(F)$, then

$$\begin{aligned}
(ev_a \circ \gamma)^{-1}(L_q(S)) &= \{(x,y) \in R \mid \gamma((x,y))(a) \in L_q(S)\} && \text{(by inverse image)} \\
&= \{(x,y) \in R \mid \gamma((x,y))(a)(S) \geq q\} && \text{(by def. } L_q(\cdot)) \\
&= \{(x,y) \in R \mid 0 \geq q\} && \text{(by def. } \gamma) \\
&= R && \text{(by } q \in [0,1] \cap \mathbb{Q})
\end{aligned}$$

□

Remark 15 (Ultrametric spaces). A similar result was proven by de Vink and Rutten [4] in the setting of ultrametric spaces and non-expansive maps. A characterization for the coalgebraic definition of bisimulation in the continuous case was established under the assumption that the bisimulation relation has a Borel decomposition. Proposition 13 does not need such an extra assumption and it holds for Borel measures as well, so that the proof-strategy of Proposition 14 can be used to drop the assumption in [4, Theorem 5.8]. ■

Theorem 16. *State bisimulation and Δ^L-bisimulation coincide.*

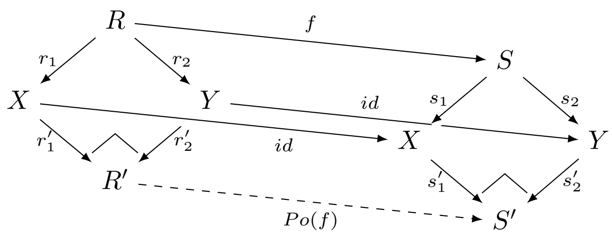
Proof. Direct consequence of Propositions 12, and 14. □

5 Relating Bisimulations and Cocongruences

In this section, we compare the notions of Δ^L-bisimulation and Δ^L-cocongruence. We do this establishing an adjunction between the category of Δ^L-bisimulations and a (suitable) subcategory of Δ^L-cocongruences. As a result of this adjunction, we obtain also sufficient conditions under which the notions of Δ^L-bisimulation and Δ^L-cocongruence coincide.

The correspondence between Δ^L-bisimulation and Δ^L-cocongruence is based on a standard adjunction between span and cospans in categories with pushouts and pullbacks, which we briefly recall. The category $\mathbf{MSp}_{\mathbf{C}}(X,Y)$ has as objects monic spans (R,f,g) between X, Y in $\mathbf{C}$, and arrows $f\colon (R,r_1,r_2) \to (S,s_1,s_2)$ which are morphisms $f\colon R \to S$ in $\mathbf{C}$ such that $s_i \circ f = r_i$, for all $i \in \{1,2\}$. The category $\mathbf{ECoSp}_{\mathbf{C}}(X,Y)$, of epic cospans between X, Y in $\mathbf{C}$, is defined analogously. In the following, we will omit the subscript $\mathbf{C}$ when the category of reference is understood.

If the category $\mathbf{C}$ has pullbacks and pushouts, we can define two functors: $Pb_{(X,Y)}\colon \mathbf{ECoSp}(X,Y) \to \mathbf{MSp}(X,Y)$ mapping each epic cospan to its pullback, and $Po_{(X,Y)}\colon \mathbf{MSp}(X,Y) \to \mathbf{ECoSp}(X,Y)$ mapping each monic span to its pushout. As for morphisms, let $f\colon (R,r_1,r_2) \to (S,s_1,s_2)$ be a morphism in $\mathbf{MSp}(X,Y)$, then $Po_{(X,Y)}(f)$ is defined as the unique arrow, given by the universal property of pushout, making the following diagram commute

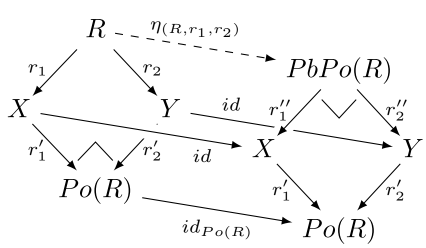


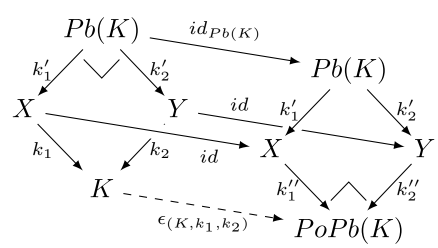
where $Po_{(X,Y)}(R, r_1, r_2) = (R', r'_1, r'_2)$ and $Po_{(X,Y)}(S, s_1, s_2) = (S', s'_1, s'_2)$. The action on arrows for $Pb_{(X,Y)}$ is defined similarly, using the universal property of pullbacks. When the domain (X, Y) of the spans and cospans is clear, the subscript in $Po_{(X,Y)}$ and $Pb_{(X,Y)}$ will be omitted. The following is standard:

Lemma 17. *Let* $\mathbf{C}$ *be a category with pushouts and pullbacks, then*

(i) $Po \dashv Pb$*;*
(ii) $PoPbPo \cong Po$ *and* $PbPoPb \cong Pb$*.*

The unit $\eta \colon Id \Rightarrow PbPo$ and counit $\epsilon \colon PoPb \Rightarrow Id$ of the adjunction $Po \dashv Pb$, are given component-wise as follows, for (R, r_1, r_2) in $\mathbf{MSp}(X, Y)$ and (K, k_1, k_2) in $\mathbf{ECoSp}(X, Y)$.





The adjunction $Po \dashv Pb$ induces a monad $(PbPo, \eta, Pb\epsilon Po)$ in $\mathbf{MSp}(X, Y)$ and a comonad $(PoPb, \epsilon, Po\eta Pb)$ in $\mathbf{ECoSp}(X, Y)$, which, by Lemma 17, are idempotent.

Since **Meas** has both pushouts and pullbacks, the construction above can be instantiated in this category. Note that, **Meas** has binary products and co-products, hence we can identify the categories $\mathbf{MSp}(\mathbf{X}, \mathbf{Y})$ and $\mathbf{ECoSp}(\mathbf{X}, \mathbf{Y})$, respectively, as the categories of relations $R \subseteq X \times Y$ (with measurable canonical projections) and quotients $(X + Y)/E$ (with measurable canonical injections), where E is an equivalence relation on $(X + Y)$.

Moreover, in **Meas** it holds that $PoPb \cong Id$ and the composite functor $PbPo$ has the following explicit description. Let $R \subseteq X \times Y$ be an object in $\mathbf{MSp}(\mathbf{X}, \mathbf{Y})$, then $PbPo(R) = R^*$ where $R^* \subseteq X \times Y$ is the *z-closure of* R, with σ-algebra Σ_{R^*} given as the initial σ-algebra w.r.t. the canonical projections;

arrows $f\colon R \to S$ between objects $R, S \subseteq X \times Y$ in $\mathbf{MSp}(\mathbf{X}, \mathbf{Y})$ are mapped to $f^*\colon R^* \to S^*$, the z-closure extension of f, defined in the obvious way.

The monad $PbPo$ in $\mathbf{MSp}(\mathbf{X}, \mathbf{Y})$ has unit $\eta\colon Id \Rightarrow PbPo$ given by the natural inclusion and multiplication $Pb\epsilon Po\colon PbPoPbPo \Rightarrow PbPo$ defined as the "identity" natural transformation. The comonad $PoPb$ in $\mathbf{ECoSp}(\mathbf{X}, \mathbf{Y})$ has counit $\epsilon\colon PoPb \Rightarrow Id$ and comultiplication $Po\eta Pb\colon PoPb \Rightarrow PoPbPoPb$ given as the "identity" natural transformations.

5.1 Adjunction between Bisimulations and Cocongruences

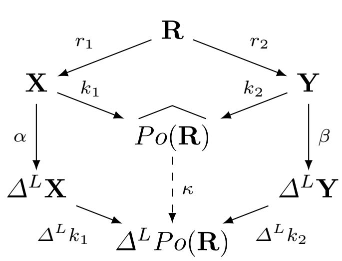
The adjunction $Po \dashv Pb$ over monic span and epic cospans in **Meas** can be partially lifted to an adjunction $\overline{Po} \dashv \overline{Pb}$ between the categories of Δ^L-bisimulations and Δ^L-cocongruences. The term "partially" is used since the adjunction can be established only restricting the category of Δ^L-cocongruences to the image given by the lifting $\overline{Po}$. Moreover, we show that the subcategories given by the images of $\overline{Pb}\,\overline{Po}$ and $\overline{Po}\,\overline{Pb}$ are equivalent. This provides sufficient conditions under which the notions of bisimulation and cocongruence coincide.

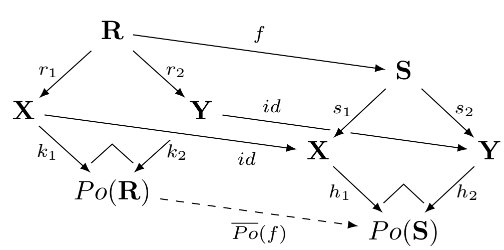
We denote by $\mathbf{Bisim}((\mathbf{X}, \alpha), (\mathbf{Y}, \beta))$ the category with Δ^L-bisimulations $((\mathbf{R}, \gamma_R), f, g)$ between Δ^L-coalgebras $(\mathbf{X}, \alpha)$ and $(\mathbf{Y}, \beta)$ as objects and arrows $f\colon ((\mathbf{R}, \gamma_R), r_1, r_2) \to ((\mathbf{S}, \gamma_S), s_1, s_2)$ which are morphisms $f\colon \mathbf{R} \to \mathbf{S}$ in **Meas** such that $s_i \circ f = r_i$, for all $i \in \{1, 2\}$, and $\gamma_S \circ f = \Delta^L f \circ \gamma_R$, i.e, f is a morphism both in Δ^L-**coalg** and $\mathbf{MSp}(\mathbf{X}, \mathbf{Y})$. The category $\mathbf{Cocong}((\mathbf{X}, \alpha), (\mathbf{Y}, \beta))$ of Δ^L-cocongruences between Δ^L-coalgebras $(\mathbf{X}, \alpha)$ and $(\mathbf{Y}, \beta)$ is defined similarly.

The functor $Po\colon \mathbf{MSp}(\mathbf{X}, \mathbf{Y}) \to \mathbf{Cospan}(\mathbf{X}, \mathbf{Y})$ is lifted to the categories of Δ^L-bisimulations and Δ^L-cocongruences by the functor

$$\overline{Po}\colon \mathbf{Bisim}((\mathbf{X}, \alpha), (\mathbf{Y}, \beta)) \to \mathbf{Cocong}((\mathbf{X}, \alpha), (\mathbf{Y}, \beta))\,,$$

acting on Δ^L-bisimulations as $\overline{Po}(((\mathbf{R}, \gamma_R), r_1, r_2)) = ((Po(\mathbf{R}), \kappa), k_1, k_2)$, where $(Po(\mathbf{R}), k_1, k_2)$ is the pushout of $(\mathbf{R}, r_1, r_2)$ and $\kappa\colon Po(\mathbf{R}) \to \Delta^L Po(\mathbf{R})$ is the unique measurable map given by the universal property of pushouts, making the diagram below commute; and on arrows $f\colon ((\mathbf{R}, \gamma_R), r_1, r_2) \to ((\mathbf{S}, \gamma_S), s_1, s_2)$ as $\overline{Po}(f)$, defined as the unique arrow, given by the universal property of pushouts, making the diagram on the right commute:





The arrow $\overline{Po}(f)$ is obviously a morphism between the cospans $(Po(\mathbf{R}), k_1, k_2)$ and $(Po(\mathbf{S}), h_1, h_2)$, and can be proved to be also a Δ^L-homomorphism exploiting the universal property of pushouts. Functoriality follows similarly.

Remark 18. The above construction is standard and applies in any category with pushouts independently of the choice of the behaviour functor.

If the behavior functor preserves weak pullbacks, cocongruences give rise to bisimulations via pullbacks (see [10, Prop. 1.2.2]). However, although Δ^L does not preserves weak pullbacks [16], if we restrict our attention only to the full subcategory of Δ^L-cocongruences that are $\overline{Po}$-images of some Δ^L-bisimulation, namely, $\overline{Po}\big(\mathbf{Bisim}((\mathbf{X},\alpha),(\mathbf{Y},\beta))\big)$, the functor Pb can be lifted as follows.

For $((\mathbf{R},\gamma_R),r_1,r_2)$ and f objects and arrows in $\mathbf{Bisim}((\mathbf{X},\alpha)(\mathbf{Y},\beta))$, respectively, the lifting $\overline{Pb}$ of Pb is defined by

$$\overline{Pb}\colon \overline{Po}\big(\mathbf{Bisim}(((\mathbf{X},\alpha)(\mathbf{Y},\beta)))\big) \to \mathbf{Bisim}(((\mathbf{X},\alpha)(\mathbf{Y},\beta)))$$

$$\overline{Pb}(\overline{Po}((\mathbf{R},\gamma_R),r_1,r_2)) = ((\mathbf{R}^*,\gamma_R^*),r_1^*,r_2^*)\,, \qquad \overline{Pb}(\overline{Po}(f)) = f^*\,,$$

where $PbPo(\mathbf{R},r_1,r_2) = (\mathbf{R}^*,r_1^*,r_2^*)$, $PbPo(f) = f^*$, and $\gamma_R^*\colon \mathbf{R}^* \to \Delta^L\mathbf{R}^*$ is the unique (sub)probability measure on $\mathbf{R}^*$ (given as in Proposition 13) such that, for all $r \in R^*$, $a \in L$, $E \in \Sigma_X$ and $F \in \Sigma_Y$

$$\gamma_R^*(r)(a)((r_1^*)^{-1}(E)) = \alpha(r_1^*(r))(a)(E)\,,$$

and

$$\gamma_R^*(r)(a)((r_2^*)^{-1}(F)) = \beta(r_2^*(r))(a)(F)\,,$$

Note that the well definition of the measure $\gamma_R^*(r)(a)$ is guaranteed by Lemma 11 and Proposition 14.

Remark 19. In the definition of γ_R^* above, we applied Proposition 13 without ensuring that r_1^* and r_2^* are surjective maps. This is not an issue and it can be solved easily as in the proof of Proposition 14. ■

The next lemma ensures that the functor $\overline{Pb}$ is well defined.

Lemma 20. *Let* $((\mathbf{R},\gamma_R),r_1,r_2)$ *and* f *be, respectively, an object and an arrow in the category* $\mathbf{Bisim}((\mathbf{X},\alpha)(\mathbf{Y},\beta))$*. Then, the following hold:*

(i) $\overline{Pb}(\overline{Po}((\mathbf{R},\gamma_R),r_1,r_2))$ *is a* Δ^L*-bisimulation between* $(\mathbf{X},\alpha)$ *and* $(\mathbf{Y},\beta)$*;*
(ii) $\overline{Pb}(\overline{Po}(f))$ *is a monic span-morphism and a* Δ^L*-homomorphism.*

The functors $\overline{Po}$ and $\overline{Pb}$ are actual liftings of Po and Pb, respectively, i.e., they commute w.r.t. to the forgetful functors mapping Δ^L-bisimulations to their monic spans and Δ^L-cocongruences to their epic cospans. This is reflected also in the following result:

Theorem 21. *Let* $(\mathbf{X},\alpha)$ *and* $(\mathbf{Y},\beta)$ *be* Δ^L*-coalgebras, then*

(i) $\overline{Po} \dashv \overline{Pb}$*;*
(ii) $\overline{Pb}\,\overline{Po}\big(\mathbf{Bisim}((\mathbf{X},\alpha),(\mathbf{Y},\beta))\big) \cong \overline{Po}\,\overline{Pb}\,\overline{Po}\big(\mathbf{Bisim}((\mathbf{X},\alpha),(\mathbf{Y},\beta))\big)$.

Proof. (i) By the universal properties of pushouts, for any pair of bisimulations $((\mathbf{R}, \gamma_R), r_1, r_2)$ and $((\mathbf{S}, \gamma_S), s_1, s_2)$ in $\mathbf{Bisim}((\mathbf{X}, \alpha), (\mathbf{Y}, \beta))$ it holds that,

$$Hom(\overline{Po}((\mathbf{R}, \gamma_R), r_1, r_2), \overline{Po}((\mathbf{S}, \gamma_S), s_1, s_2) \cong \\ Hom(((\mathbf{R}, \gamma_R), r_1, r_2), \overline{Pb}\,\overline{Po}((\mathbf{S}, \gamma_S), s_1, s_2)\,,$$

i.e., $\overline{Po}$ is left adjoint to $\overline{Pb}$. (ii) The equivalence is given by the functors $\overline{Po}$ and $\overline{Pb}$ and follows by Proposition 17(ii). ☐

Theorem 21 provides sufficient conditions under which the notions of bisimulation and cocongruence coincide: (ii) states that the Δ^L-bisimulation formed applying the (closure) operator $\overline{Pb}\,\overline{Po}$ to a Δ^L-bisimulation $((\mathbf{R}, \gamma_R), r_1, r_2)$ is equivalent to the Δ^L-cocongruences obtained applying the operator $\overline{Po}\,\overline{Pb}\,\overline{Po}$.

Related Work. In [3], Danos et al. proposed a notion alternative to bisimulations, the so called *event bisimulation*, being aware that it coincides with cocongruence. Here we recall its definition and try to make a comparison between the results in [3] in connection to Theorem 21.

Definition 22 (Event bisimulation). *Let $\mathcal{M} = (X, \Sigma_X, \{\theta_a\}_{a\in L})$ be an L-labelled Markov kernel. A sub-σ-algebra $\Lambda \subseteq \Sigma_X$ is an* event bisimulation *on $\mathcal{M}$ if, for all $a \in L$, $q \in \mathbb{Q} \cap [0, 1]$, and $E \in \Lambda$, $\theta_a^{-1}(L_q(E)) \in \Lambda$.*

Any σ-algebra Σ on X induces a notion of separability in the form of an relation $\Re(\Sigma) \subseteq X \times X$ defined by $\Re(\Sigma) = \{(x, y) \mid \forall E \in \Sigma. [x \in E \text{ iff } y \in E]\}$. Moreover, considering only equivalence relations $\mathcal{R} \subseteq X \times X$, they denoted by $\Sigma(\mathcal{R}) = \{E \in \Sigma \mid (E, E) \text{ is } \mathcal{R}\text{-closed}\}$ the set of measurable $\mathcal{R}$-closed sets, which is readily seen to be a σ-algebra on X. The "operator" $\Re(\cdot)$ maps σ-algebras to equivalence relations and, conversely, $\Sigma(\cdot)$ maps equivalence relations to σ-algebras. Moreover, as the next lemma states, under certain circumstances, they can also be thought of as maps between event bisimulations and state bisimulations.

Lemma 23 ([3, Lemma 4.1]). *Let $\mathcal{M} = (X, \Sigma_X, \{\theta_a\}_{a\in L})$ be an L-labelled Markov kernel. Then, $\mathcal{R}$ is a state bisimulation iff $\Sigma(\mathcal{R})$ is an event bisimulation.*

Proof. Assume $\mathcal{R}$ is a state bisimulation on $\mathcal{M}$. It is easy to show that $\Sigma(\mathcal{R})$ is a sub-σ-algebra of Σ. It remains to prove that for all $a \in L$ and $q \in \mathbb{Q} \cap [0, 1]$, and $E \in \Sigma(\mathcal{R})$, $\theta_a^{-1}(L_q(E)) \in \Sigma(\mathcal{R})$. By denoting the canonical projections of $\mathcal{R}$, by π_1 and π_2, we have

$$\begin{aligned} \pi_1^{-1}(\theta_a^{-1}(L_q(E))) &= \{(x, y) \in \mathcal{R} \mid \theta_a(x)(E) \geq q\} && \text{(by pre-image)} \\ &= \{(x, y) \in \mathcal{R} \mid \theta_a(y)(E) \geq q\} && \text{(by } \mathcal{R} \text{ state bisim.)} \\ &= \pi_2^{-1}(\theta_a^{-1}(L_q(E)))\,. && \text{(by pre-image)} \end{aligned}$$

This proves that $(\theta_a^{-1}(L_q(E)), \theta_a^{-1}(L_q(E)))$ is $\mathcal{R}$-closed, so that $\Sigma(\mathcal{R})$ is an event bisimulation. Conversely, assume that $\Sigma(\mathcal{R})$ is an event bisimulation on $\mathcal{M}$. Then, for every $(x, y) \in \mathcal{R}$, $a \in L$, and $E \in \Sigma(\mathcal{R})$, we have that

$$\text{for all } q \in \mathbb{Q} \cap [0,1]\,, \quad [\theta_a(x)(E) \geq q \quad \text{iff} \quad \theta_a(y)(E) \geq q]\,.$$

Since $\mathbb{Q} \cap [0,1]$ is dense in $[0,1]$, the above implies $\theta_a(x)(E) = \theta_a(y)(E)$. So $\mathcal{R}$ is a state bisimulation. □

Note that, given that Λ is an event bisimulation, it is not always the case that its induced separability relation $\Re(\Lambda)$ is a state bisimulation. This, somehow, seems in accordance with our restriction to a well-behaved subcategory of cocongruences in the definition of the functor $\overline{Pb}$. Indeed, when one restricts the attention only to state bisimulations $\mathcal{R}$ that are assumed to be equivalence relations, the results in [3] are related to our adjunction as follows:

$$Po(\mathcal{R}) = X/\Re(\Sigma(\mathcal{R}))\,, \qquad\qquad PbPo(\mathcal{R}) = R^* = R\,.$$

In particular, many of the lemmas and propositions in [3, §4] are consequences of Theorem 21.

6 Conclusions and Future Work

We have proposed a genuinely new characterization of bisimulation in plain mathematical terms, which is proven to be in one-to-one correspondence with the coalgebraic definition of Aczel and Mendler.

Then, the notions of bisimulation and cocongruence (equivalently, event bisimulation) are formally compared establishing an adjunction between the category of coalgebraic bisimulations and a suitable subcategory of cocongruences. By means of this adjunction we provided sufficient conditions under which the notions of bisimulation and cocongruence coincide.

A comparison between bisimulations and cocongruences by means of an adjunction between their categories is interesting not just for Markov processes but, more in general, for any F-coalgebra. Usually, the final bisimulation (i.e., bisimilarity) between two coalgebras is said to be "well behaved" if it coincides with the pullback of the final cocongruence between the same pair of coalgebras. When a final coalgebra exists, the final cocongruence is given by the pair of final homomorphisms and its pullback is called behavioral equivalence (or final semantics). A sufficient condition to ensure that bisimilarity coincides with behavioral equivalence is to require that the behaviour functor (weakly) preserves pullbacks or semi-pullbacks (for instance see [12, Theorem 9.3]). When the behavior functor does not (weakly) preserves pullbacks or semi-pullbacks, one may use cocongruences instead of bisimulations. Another way (the one we have explored in Section 5) is to use adjunctions, which allows one to focus on the well behaved bisimulations by considering some suitable full-subcategories. In this light, Theorem 21 generalizes the results in [3], which are consequences of the existence of the adjunction $\overline{Po} \dashv \overline{Pb}$.

More generally, it would be interesting to study to what extent the weak-pullback preservation assumption on the functor could be removed. For example, in [17] Worrell proved that the category of F-coalgebras over **Set** is complete, provided that the functor F weakly-preserves pullbacks and is bounded. It would be nice to extend this result to general categories, replacing the former assumption with the existence of an adjunction between F-bisimulations and F-cocongruences.

References

1. R. Blute, J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. In *LICS*, pages 149–158. IEEE Computer Society, 1997.
2. L. Cardelli and R. Mardare. The measurable space of stochastic processes. In *Proc. QEST*, pages 171–180. IEEE Computer Society, 2010.
3. V. Danos, J. Desharnais, F. Laviolette, and P. Panangaden. Bisimulation and cocongruence for probabilistic systems. *Inf. Comput.*, 204(4):503–523, 2006.
4. E. P. de Vink and J. J. M. M. Rutten. Bisimulation for probabilistic transition systems: A coalgebraic approach. *Theor. Comput. Sci.*, 221(1-2):271–293, 1999.
5. J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. *Inf. Comput.*, 179(2):163–193, 2002.
6. J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labeled markov processes. In *Proc. LICS*, pages 95–106, 2000.
7. J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labelled markov processes. *Inf. Comput.*, 184(1):160–200, 2003.
8. M. Giry. A categorical approach to probability theory. In Banaschewski, editor, *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer Berlin/Heidelberg, 1982. 10.1007/BFb0092872.
9. A. Heifetz and D. Samet. Topology-Free Typology of Beliefs. *Journal of Economic Theory*, 82(2):324–341, 1998.
10. A. Kurz. *Logics for Coalgebras and Applications to Computer Science*. PhD thesis, Ludwig-Maximilians-Universität München, 2000.
11. K. G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.
12. J. J. M. M. Rutten. Universal coalgebra: a theory of systems. *TCS*, 249(1):3–80, 2000.
13. P. S. Terraf. Unprovability of the logical characterization of bisimulation. *CoRR*, abs/1005.5142, 2010.
14. P. S. Terraf. Bisimilarity is not borel. *CoRR*, abs/1211.0967, 2012.
15. F. van Breugel, M. W. Mislove, J. Ouaknine, and J. Worrell. Domain theory, testing and simulation for labelled Markov processes. *Theoretical Computer Science*, 333(1-2):171–197, 2005.
16. I. D. Viglizzo. *Coalgebras on Measurable Spaces*. PhD thesis, Department of Mathematics, IndianaUniversity, 2005.
17. J. Worrell. Toposes of Coalgebras and Hidden Algebras. *Electronic Notes in Theoretical Computer Science*, 11(0):212 – 230, 1998.

A Technical proofs

In this appendix we provide the proofs of the technical results used in the paper.

Proof (of Lemma 3). We have to prove that, given $\mathcal{R} \cap (E \times Y) = \mathcal{R} \cap (X \times F)$ and $\mathcal{R}' \subseteq \mathcal{R}$, we have $\mathcal{R}' \cap (E \times Y) = \mathcal{R}' \cap (X \times F)$.

($\subseteq$) Let $(x, y) \in \mathcal{R}'$ and $x \in E$. By $\mathcal{R}' \subseteq \mathcal{R}$, $(x, y) \in \mathcal{R}$. By (E, F) $\mathcal{R}$-closed, we have $y \in F$.

($\supseteq$) Let $(x, y) \in \mathcal{R}'$ and $y \in F$. By $\mathcal{R}' \subseteq \mathcal{R}$, $(x, y) \in \mathcal{R}$. By (E, F) $\mathcal{R}$-closed, we have $x \in E$. □

Proof (of Lemma 4). We prove only the inclusion $E \subseteq F$, the reverse is similar. Assume $x \in E$. By reflexivity of $\mathcal{R}$, $(x, x) \in \mathcal{R}$. Since (E, F) is $\mathcal{R}$-closed, we have $x \in F$. To prove that E is an union of $\mathcal{R}$-equivalence classes, it suffices to show that if $x \in E$ and $(x, y) \in \mathcal{R}$, then $y \in E$. This easily follows since $E = F$. □

Proof (of Proposition 10). ($\Rightarrow$) Immediate from Corollary 7. ($\Leftarrow$) Let $x, y \in X$ be such that

$$\forall a \in L, \forall E \in \Sigma \text{ such that } (E, E) \sim\text{-closed}\,, \quad \theta_a(x)(E) = \theta_a(y)(E)\,. \tag{5}$$

We prove $x \sim y$ showing a bisimulation $\mathcal{R}$ such that $(x, y) \in \mathcal{R}$. Let $\mathcal{R}$ be the smallest equivalence containing $\{(x, y)\}$ and $\sim$, hence $\mathcal{R} = Id_X \cup \bigcup_{n \in \mathbb{N}} \mathcal{S}_n$, where

$$\mathcal{S}_0 \triangleq \{(x, y), (y, x)\} \cup \sim \qquad\qquad \mathcal{S}_{n+1} \triangleq \mathcal{S}_n ; \mathcal{S}_n\,.$$

By Lemma 4, it suffices to prove that, for all $a \in L$ and $E' \in \Sigma$ such that (E', E') is $\mathcal{R}$-closed, the following holds:

$$(x', y') \in \mathcal{R} \implies \theta_a(x')(E') = \theta_a(y')(E')\,. \tag{6}$$

If $(x', y') \in \mathcal{R}$, then $(x', y') \in Id_X$ or $(x', y') \in \mathcal{S}_n$ for some $n \geq 0$. If $(x', y') \in Id_X$ (6) holds trivially. Now we show, by induction on $n \geq 0$, that

$$(x', y') \in \mathcal{S}_n \implies \theta_a(x')(E') = \theta_a(y')(E')\,. \tag{7}$$

Base case ($n = 0$): let $(x', y') \in \sim$. By $\sim \subseteq \mathcal{R}$ and Lemma 3, (E', E') is $\sim$-closed. Thus, by Corollary 7, (7) holds. If $(x', y') \in \{(x, y), (y, x)\}$, then property (5) holds. Again, by Lemma 3, (E', E') is $\sim$-closed, thus (7) holds.
Inductive case ($n+1$): if $(x', y') \in \mathcal{S}_{n+1}$, then there exists some $z \in X$ such that $(x', z) \in \mathcal{S}_n$ and $(z, y') \in \mathcal{S}_n$. Then, applying the inductive hypothesis twice, we have $\theta_a(x')(E) = \theta_a(z)(E) = \theta_a(y')(E)$. □

Proof (of Lemma 11). Note that $R^* = \bigcup_{n \in \mathbb{N}} R_n$, for R_n and $n \in \mathbb{N}$ defined as

$$R_0 = R\,, \qquad\qquad R_{n+1} = R; R_n^{-1}; R\,.$$

Therefore, that R^* is a bisimulation between the two Markov kernels, directly follows if we can prove that, whenever $(x, y) \in R_n$, then $\alpha_a(E) = \beta_a(F)$, for all $a \in L$, and any pair (E, F) of R_n-closed measurable sets $E \in \Sigma_X$ and $F \in \Sigma_Y$. We proceed by induction on $n \geq 0$.

Base case ($n = 0$): Clearly, $R_0 = R$ and R is a state bisimulation.
Inductive step ($n > 0$): Suppose $(x, y) \in R_{n+1}$, then there exist $x' \in X$ and $y' \in Y$ such that $(x, y') \in R$, $(x', y') \in R_n$, $(x', y) \in R$. By Lemma 4, any pair (E, F) which is R_{n+1}-closed is also R_n-closed and R-closed. So that, by hypothesis on R and the inductive hypothesis on R_n, for every $a \in L$, $E \in \Sigma_Y$ and $F \in \Sigma_Y$, we have

$$\alpha_a(x)(E) = \beta_a(y')(F)\,, \quad \alpha_a(x')(E) = \beta_a(y')(F)\,, \quad \alpha_a(x')(E) = \beta_a(y)(F)\,.$$

Therefore, $\alpha_a(x)(E) = \beta_a(y)(F)$. □

Proof (of Proposition 13). We define $\mu \wedge \nu$ has the Hahn-Kolmogorov extension of a pre-measure defined on a suitable field $\mathcal{F}$ such that $\sigma(\mathcal{F}) = \Sigma_R$. Let $\mathcal{F}$ be the collection of all *finite* unions $\bigcup_{i=0}^{k} G_i$, where $k \in \mathbb{N}$, and for all $i = 0..k$, $G_i = r_1^{-1}(E_i) \cap r_2^{-1}(F_i)$, for some $E_i \in \Sigma_X$ and $F_i \in \Sigma_Y$. Clearly, $\sigma(\mathcal{F}) = \Sigma_R$. To prove that $\mathcal{F}$ is a field we need to show that it is closed under finite intersection and complement (it is already closed under finite union). This is immediate by De Morgan laws and the following equalities:

$$\Big(r_1^{-1}(E_i) \cap r_2^{-1}(F_i)\Big) \cap \Big(r_1^{-1}(E_j) \cap r_2^{-1}(F_j)\Big) = r_1^{-1}(E_i \cap E_j) \cap r_2^{-1}(F_i \cap F_j)\,,$$
$$R \setminus \Big(r_1^{-1}(E_i) \cap r_2^{-1}(F_i)\Big) = r_1^{-1}(X \setminus E_i) \cup r_2^{-1}(Y \setminus F_i)\,.$$

Now we define $\mu \wedge \nu \colon \mathcal{F} \to [0, \infty]$. Note that any element in $S \in \mathcal{F}$ can always be decomposed into a finite union $S = \bigcup_{i=0}^{k} G_i$ of *pair-wise disjoint* sets of the form $G_i = r_1^{-1}(E_i) \cap r_2^{-1}(F_i)$, where $E_i \in \Sigma_X$ and $F_i \in \Sigma_Y$ (hereafter, called G-sets). Then for any G-set we define

$$(\mu \wedge \nu)(G_i) = \begin{cases} \mu(E) & \text{if } \exists E \in \Sigma_X .\, G_i = r_1^{-1}(E) \text{ and } \forall F \in \Sigma_Y .\, G_i \neq r_2^{-1}(F) \\ \nu(E) & \text{if } \forall E \in \Sigma_X .\, G_i \neq r_1^{-1}(E) \text{ and } \exists F \in \Sigma_Y .\, G_i = r_2^{-1}(F) \\ \mu(E) & \text{if } \exists E \in \Sigma_X\,, F \in \Sigma_Y .\, G_i = r_1^{-1}(E) = r_2^{-1}(F) \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

and we define

$$(\mu \wedge \nu)(S) = \sum_{i=0}^{k} (\mu \wedge \nu)(G_i)\,. \tag{9}$$

Note that, in the definition of $\mu \wedge \nu$ on G-sets, the surjectivity of r_1 and r_2 guarantees that if $G_i = r_1^{-1}(E)$ or $G_i = r_2^{-1}(F)$, then E and F are unique. So that, (8) is well-defined. Moreover, the definition on S does not depend on how S is decomposed into a disjoint union. To see this, note that any two representations $\bigcup_{i=0}^{k} G_i$ and $\bigcup_{i=0}^{k'} G'_i$ for S can be decomposed into a common refinement $\bigcup_{i=0}^{k''} G''_i$, so that, by the well definition of $\mu \wedge \nu$ on G-sets they must agree on it. Therefore, $\mu \wedge \nu$ is well-defined on all $\mathcal{F}$, and by construction is finitely additive.

It remains to show that, if $S \in \mathcal{F}$ is the countable disjoint union of sets $S_0, S_1, S_2, \ldots \in \mathcal{F}$, then

$$(\mu \wedge \nu)(S) = \textstyle\sum_{n \in \mathbb{N}} (\mu \wedge \nu)(S_n)\,.$$

Splitting up S into disjoint G-sets, and restricting S_n to each of these G-sets sets in turn, for all $n \in \mathbb{N}$, by finite additivity of $\mu \wedge \nu$, we may assume without loss of generality that $S = r_1^{-1}(E) \cap r_2^{-1}(F)$, for some $E \in \Sigma_X$ and $F \in \Sigma_Y$. In the same way, by breaking up each S_n into a G-set and using finite additivity of $\mu \wedge \nu$ again, we may assume without loss of generality that each S_n takes the form $S_n = r_1^{-1}(E_n) \cap r_2^{-1}(F_n)$, for some $E_n \in \Sigma_X$ and $F_n \in \Sigma_Y$. By definition of $\mu \wedge \nu$ and σ-additivity of μ and ν, (9) is rewritten as follows

$$(\mu \wedge \nu)(S) = \textstyle\sum_{n \in \mathbb{N}} (\mu \wedge \nu)(S_n) = (\mu \wedge \nu)(r_1^{-1}(E) \cap r_2^{-1}(F)) \,.$$

This proves that $\mu \wedge \nu \colon \mathcal{F} \to [0, \infty]$ is a pre-measure. By Hahn-Kolmogorov theorem, $\mu \wedge \nu$ can be extended to Σ_R, and since, for all $E \in \Sigma_X$ and $F \in \Sigma_Y$,

$$\begin{aligned} r_1^{-1}(E) &= r_1^{-1}(E) \cap R = r_1^{-1}(E) \cap r_2^{-1}(Y) \in \mathcal{F} \,, \\ r_2^{-1}(F) &= R \cap r_2^{-1}(F) = r_1^{-1}(X) \cap r_2^{-1}(F) \in \mathcal{F} \,, \end{aligned}$$

together with the hypothesis made on μ and ν, i.e.,

$$\text{if } r_1^{-1}(E) = r_2^{-1}(F) \,, \text{ then } \mu(E) = \nu(F) \,,$$

by (8) we have

$$(\mu \wedge \nu)(r_1^{-1}(E)) = \mu(E) \qquad \text{and} \qquad (\mu \wedge \nu)(r_2^{-1}(F)) = \nu(F) \,,$$

therefore, the required conditions are satisfied. If both μ and ν are σ-finite, so is the pre-measure $\mu \wedge \nu \colon \mathcal{F} \to [0, \infty]$, hence its extension on Σ_R is unique. □

Proof (of Lemma 20).

(i) Immediate from Lemma 11 and the correspondence between state bisimulation and Δ^L-bisimulation (Theorem 16).

(ii) Let $f \colon ((\mathbf{R}, \gamma_R), r_1, r_2) \to ((\mathbf{S}, \gamma_S), s_1, s_2)$ in $\mathbf{Bisim}((\mathbf{X}, \alpha), (\mathbf{Y}, \beta))$. By definition $\overline{Pb}(\overline{Po})(f) = PbPo(f) = f^*$, hence it is a monic span-morphism from $(\mathbf{R}^*, r_1^*, r_2^*)$ to $(\mathbf{S}^*, s_1^*, s_2^*)$. To prove that f^* is a Δ^L-homomorphism between the coalgebras $(\mathbf{R}^*, \gamma_R^*)$ and $(\mathbf{S}^*, \gamma_S^*)$, we have to show $\gamma_S^* \circ f^* = \Delta^L f^* \circ \gamma_R^*$. By the unicity of the definition of the (sub)probability measures $\gamma_S^*(f^*(r))(a)$ and $\gamma_R^*(r)(a)$, to prove the equality it suffices to show that for arbitrary $r \in R^*$, $a \in L$, $E \in \Sigma_X$, and $F \in \Sigma_Y$,

$$(\gamma_S^* \circ f^*)(r)(a)((s_1^*)^{-1}(E)) = (\Delta^L f^* \circ \gamma_R^*)(r)(a)((s_1^*)^{-1}(E)) \,,$$

and

$$(\gamma_S^* \circ f^*)(r)(a)((s_2^*)^{-1}(F)) = (\Delta^L f^* \circ \gamma_R^*)(r)(a)((s_2^*)^{-1}(F)) \,.$$

We prove only the first equality, the other follows similarly.

$$
\begin{aligned}
(\gamma_S^* \circ f^*)(r)(a)((s_1^*)^{-1}(E)) = & \\
&= \gamma_S^*(f^*(r))(a)((s_1^*)^{-1}(E)) && \text{(composition)}\\
&= \alpha(s_1^* \circ f^*(r))(a)(E) && \text{(by def. } \gamma_S^*\text{)}\\
&= \alpha(r_1(r))(a)(E) && \text{(by } f \text{ span-morphism)}\\
&= \gamma_R^*(r)(a)((r_1^*)^{-1}(E)) && \text{(by def. } \gamma_R^*\text{)}\\
&= \gamma_R^*(r)(a)((s_1^* \circ f^*)^{-1}(E)) && \text{(by } f^* \text{ span-morphism)}\\
&= \gamma_R^*(r)(a)((f^*)^{-1} \circ (s_1^*)^{-1}(E)) && \text{(by comp. inverses)}\\
&= (\varDelta^L f^* \circ \gamma_R^*)(a)((s_1^*)^{-1}(E)) && \text{(by def. } \varDelta^L\text{)}
\end{aligned}
$$

□