Composition of stochastic transition systems based on spans and couplings *

Extended version - July 8, 2016

Daniel Gburek, Christel Baier, and Sascha Klüppelholz

Faculty of Computer Science,
Technische Universität Dresden, Germany
{daniel.gburek,christel.baier,sascha.klueppelholz}@tu-dresden.de

Abstract

Conventional approaches for parallel composition of stochastic systems relate probability measures of the individual components in terms of product measures. Such approaches rely on the assumption that components interact stochastically independent, which might be too rigid for modeling real world systems. In this paper, we introduce a parallel-composition operator for stochastic transition systems that is based on couplings of probability measures and does not impose any stochastic assumptions. When composing systems within our framework, the intended dependencies between components can be determined by providing so-called spans and span couplings. We present a congruence result for our operator with respect to a standard notion of bisimilarity and develop a general theory for spans, exploiting deep results from descriptive set theory. As an application of our general approach, we propose a model for stochastic hybrid systems called stochastic hybrid motion automata.

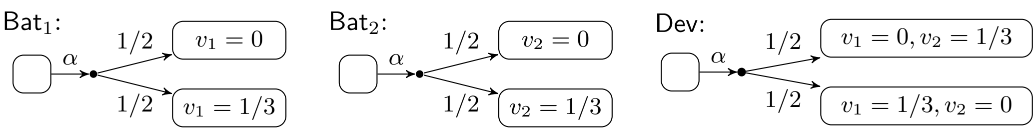
1998 ACM Subject Classification [F.1.1] Models of Computation

Keywords and phrases Stochastic Transition System, Composition, Stochastic Hybrid Motion Automata, Stochastically Independent, Coupling, Span, Bisimulation, Congruence, Polish Space

Digital Object Identifier 10.4230/LIPIcs.ICALP.2016.XXX

1 Introduction

When modeling complex systems, compositional approaches enjoy many favorable properties compared to their monolithic counterparts. They allow for a systematic system design, facilitate the interchangeability and reusability of components, and thus also ease the maintainability. A major objective in defining compositional frameworks is to separate concerns into components – specifying the operational behavior – and composition operators – addressing the communication and interaction of the components. Within conventional approaches for stochastic systems, the composition operator relates probability distributions of the individual components in terms of product distributions. Therefore, such operators are based on the assumption that the components interact stochastically independent, which is often not adequate. For instance, let us regard a systems composed of a device Dev and two batteries Bat_1 and Bat_2 providing the energy for Dev as detailed below:



* The authors are supported by the DFG through the Collaborative Research Center SFB 912 – HAEC, the Excellence Initiative by the German Federal and State Governments (cluster of excellence cfAED and Institutional Strategy), the Research Training Groups QuantLA (GRK1763) and RoSI (GRK1907).

43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016).
Editors: Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi; Article No. XXX; pp. XXX:1–XXX:47
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In this example, the device provides the environmental context in which Bat_1 and Bat_2 are operating. Hence, Dev may, e.g., be the reason for common cause failures arising in the system. Let the variables v_1 and v_2 capture the amount of energy stored within Bat_1 and Bat_2, respectively. The action label α stands for the occurrence of a failure after which all the components will crash. As a consequence, the level of the stored energy of the batteries instantaneously drops to either 0 or 1/3 with probability 1/2, respectively. When considering the batteries in isolation, Bat_1 and Bat_2 appear stochastically independent in the first place and thus, product distributions in the parallel composition $\mathsf{Bat}_1 \parallel \mathsf{Bat}_2$ seem to be adequate. However, additional dependencies can be imposed by Dev, influencing the interplay between the batteries. The assumption that Bat_1 and Bat_2 are stochastically independent is hence not adequate. Assume, e.g., that Dev uses Bat_1 as the default power supply and Bat_2 as a backup. Then, within a failure situation, Bat_1 is more likely to be affected than Bat_2. The most likely case is that Bat_1 drops to 0 whereas Bat_2 drops to 1/3. Hence, v_1 and v_2 might be not independent in the composite system $(\mathsf{Bat}_1 \parallel \mathsf{Bat}_2) \parallel \mathsf{Dev}$.

Motivated by this example, we consider hybrid systems that combine discrete behaviors and continuous dynamics. In this setting, the most prominent modeling formalism are hybrid automata, which comprise a control graph with discrete jumps between (control) locations and flows that model the evolution of continuous variables over time. When time passes in a hybrid system, a flow starting from the current variable evaluation is selected non-deterministically and then the variables evolve according to the chosen flow. Besides the stochastic independence, additional aspects are relevant for the composition of hybrid systems. Let us assume that α_1 is a (local) action of Bat_1 which cannot be observed by Bat_2 or Dev. Particularly, α_1 does not affect the value of variable v_2. The hybrid automaton $\mathsf{Bat}_1 \parallel \mathsf{Bat}_2$ has states of the form $\langle s^1, s^2 \rangle$. Suppose $\langle s^1_1, s^2_1 \rangle \to^{t_1} \langle s^1_2, s^2_2 \rangle \to^{\alpha_1} \langle s^1_3, s^2_3 \rangle \to^{t_2} \langle s^1_4, s^2_4 \rangle$ is a finite path in $\mathsf{Bat}_1 \parallel \mathsf{Bat}_2$, comprising two timed transitions with time passages t_1 and t_2 and one jump transition involving action α_1. As α_1 cannot be observed by Bat_2, we expect $s^2_1 \to^{t_1+t_2} s^2_4$ in Bat_2. In particular, a faithful model for the composite system would allow for selecting a flow for v_1 within time passage t_1, which is continued within the subsequent time passage. Thus, the adaption of the flow for variable v_2 should only be possible when executing an action involving Bat_2 or Dev. This aspect is also crucial in the context of modeling controller strategies for hybrid systems. Typically, control decisions are made at distinct points and fixed until a next control decision is enabled. For instance, when considering a traffic alert and collision avoidance systems on aircraft, the advise of a corrective maneuver is determined when a critical situation occurs and fixed until sensor values exceed a threshold that indicates changes of the situation. A crucial point is to identify exactly those situations where adaptation of flows is allowed and required, as from a practical point of view it is important to minimize costs of adaptation and to keep the complexity of controllers manageable.

Contribution. We introduce a generic composition operator for stochastic transition systems (STSs) [18] based on spans and span couplings. Our operator does not rely on the assumption that the STSs to be composed are stochastically independent and covers standard composition operators by dealing with specific spans. Spans provide a formal approach for introducing a universal notion of coupling probability measures. We develop an extensive theory for spans exploiting profound results known from descriptive set theory [39]. Based on a standard notion of bisimulation, we provide a congruence result with respect to our span composition. In the second part of the paper, we instantiate our general approach and introduce stochastic hybrid motion automata (SHMA) in which the progressing flow is recorded within states. We present a compositional framework for SHMA including an

STS-semantics, a composition operator that does not rely on the assumption of stochastic independence, and where the adjustment of flows is always accompanied with an action. We show that the congruence result for STSs transfers to our SHMA framework.

Additional material and detailed proofs can be found in the appendix.

Related Work. We are not aware of a compositional modeling approach of stochastic systems which does not rely on the assumption that the components to be composed are stochastically independent. Our work thus addresses a fundamental challenge in the context of probabilistic operational models. The recent work [33] gives a comprehensive overview on compositional probabilistic modeling formalisms regarding expressive power and available analysis techniques. The concept of compositionality has its roots in the theory of process calculi [45, 37] and there are many fundamental contributions in the field of stochastic extensions of process calculi and probabilistic automata [49, 1, 2, 20, 15, 21, 22]. Results on discrete systems have been extended to formalisms with continuous state spaces [18, 41]. The theory on non-deterministic labeled Markov processes (NLMPs) provide elegant notions and results on bisimulation and its logical characterization [24, 25, 19, 23, 6, 36]. Unfortunately, NLMPs are a priori not appropriate for our purposes as the class of NLMPs is not closed under the composition of stochastic transition systems (cf. appendix): Given two NLMPs, the transition function of their composition does not need to be measurable.

When considering real-time systems, an important distinguishing aspect is the notion of residence time, which is the time spend in a state before moving to a successor state. In prominent compositional frameworks, timing behavior is modeled by clocks (timed automata) [4, 11, 46, 13] or one has exponential-distributed holding times (Markov automata and interactive Markov chains) [35, 29]. A general theory on compositionality and behavioral equivalences has been also achieved for probabilistic real-time systems modeled by interactive generalized semi-Markov processes [16, 14]. When adding flows to specify the evolution of continuous variables between jumps, one enters the field of hybrid systems [3, 34, 12]. The spirit of our work concerning hybrid systems is closest to the compositional frameworks developed for hybrid extensions of I/O-automata [44] and reactive modules [5] in the non-stochastic case. [44] studies parallel composition, simulation relations, and the receptiveness property and deals with prefix-, suffix- and concatenation-closed sets of flows on a syntactic level to obtain time-transitivity. Probabilistic hybrid automata [51, 31] extend classical hybrid automata by discrete probabilistic updates for the jumps. In [30, 32, 31], stochastic hybrid automata are considered where variables can be updated according to continuous distributions. Different from these hybrid automata, the change of flows in SHMA is only possible when some action is executed. Stochastic flows, i.e., where stochastic choices can be made continuously over time, are considered in [17, 38]. Our framework does not incorporate this kind of flows so far.

2 Preliminaries

We suppose the reader is familiar with standard concepts from measure and probability theory [9, 10]. We briefly summarize our notations used throughout this paper.

Couplings. Within our work we understand couplings as a "modeling tool". Intuitively, couplings relate given measures in a product space by a measure with corresponding marginals. $\mathsf{Prob}(X)$ denotes the set of all probability measures on the measurable space X. Let X_1 and X_2 be measurable spaces. Given $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$, $\mu \in \mathsf{Prob}(X_1 \times X_2)$ is called a *coupling of* (μ_1, μ_2) if $\mu(M_1 \times X_2) = \mu_1(M_1)$ and $\mu(X_1 \times M_2) = \mu_2(M_2)$ for all measurable $M_1 \subseteq X_1$ and $M_2 \subseteq X_2$. The *independent coupling of* (μ_1, μ_2) is the *product* *measure* of μ_1 and μ_2 denoted by $\mu_1 \otimes \mu_2$. If $\mu_1 = \mathsf{Dirac}[x_1]$ for some $x_1 \in X_1$, then there is exactly one coupling of (μ_1, μ_2), namely the independent one. Here, $\mathsf{Dirac}[x_1]$ denotes the probability measure where for all measurable $M_1 \subseteq X_1$, $\mathsf{Dirac}[x_1](M_1) = 1$ iff $x_1 \in M_1$.

Polish spaces. A separable and completely metrizable topological space is called a *Polish space* [39]. If X is a Polish space, then $\mathsf{Prob}(X)$ is well equipped with the topology induced by the weak convergence of probability measures. To obtain a measurable space, Polish spaces are equipped with the Borel sigma algebra, i.e., the coarsest sigma-algebra where all open sets are measurable. We call a measurable space X *standard Borel* if there exists a Polish topology on X where the induced Borel sigma-algebra coincides with the given one. The Polish topology is in general not uniquely determined. We refer to measurable subsets of standard Borel spaces as *Borel sets.* Of course, every Polish space is standard Borel.

Functions for probability measures. Given a measurable function $f\colon X_1 \to X_2$ between measurable spaces X_1 and X_2, the *pushforward of* f is defined by $f_\sharp\colon \mathsf{Prob}(X_1) \to \mathsf{Prob}(X_2)$, $f_\sharp(\mu)(M_2) = \mu(f^{-1}(M_1))$. Assuming Polish spaces X_1 and X_2, a *Markov kernel* is a Borel function $k\colon X_1 \to \mathsf{Prob}(X_2)$. Here, for every $\mu_1 \in \mathsf{Prob}(X_1)$ we define *semi-product measure* $\mu_1 \rtimes k \in \mathsf{Prob}(X_1 \times X_2)$, $\mu_1 \rtimes k(M_1 \times M_2) = \int_{M_1} k(x_1)(M_2)\, d\mu_1(x_1)$.

Relations. Let $R \subseteq X_1 \times X_2$ be a binary relation over some sets X_1 and X_2. We usually write $x_1 \, R \, x_2$ instead of $\langle x_1, x_2\rangle \in R$. Then, R is called *lr-total in* $X_1 \times X_2$ if for all $x_1 \in X_1$ there exists $x_2 \in X_2$ such that $x_1 \, R \, x_2$ and vice versa, i.e., also for all $x_2 \in X_2$ there exists $x_1 \in X_1$ where $x_1 \, R \, x_2$. Assume X_1 and X_2 constitute measurable spaces and let $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$. A *weight function for* (μ_1, R, μ_2) is a coupling W of (μ_1, μ_2) such that $x_1 \, R \, x_2$ for W-almost all $\langle x_1, x_2\rangle \in X_1 \times X_2$. We write $\mu_1 \, R^w \, \mu_2$ if there exists a weight function for (μ_1, R, μ_2). Notice, R^w constitutes a relation in $\mathsf{Prob}(X_1) \times \mathsf{Prob}(X_2)$. Notice, weight functions are also well-established in the discrete setting [49].

Variables. Let Var denote a countable set of variables and $V \subseteq \mathsf{Var}$. We denote by $\mathsf{Ev}(V)$ the set of all *variable evaluations for* V, i.e., functions from V to $\mathbb{R}$. As the countable product of Polish spaces equipped with the product topology again yields a Polish space, $\mathsf{Ev}(V)$ constitutes a Polish space. Let $e \in \mathsf{Ev}(\mathsf{Var})$ and $\eta \in \mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}))$. The projection $e_{|V} \in \mathsf{Ev}(V)$ is given by $e_{|V}(v) = e(v)$ for all $v \in V$. As $f\colon \mathsf{Ev}(\mathsf{Var}) \to \mathsf{Ev}(V)$, $f(e) = e_{|V}$ is measurable, we can safely define $\eta_{|V} = f_\sharp(\eta)$. $\mathsf{Cond}(\mathsf{Var})$ denotes the set of all Boolean conditions over Var and we write $e \models c$ if the variable evaluation e satisfies condition c. For instance, $e \models (v \leq 3.14159) \wedge (v \geq 2.71828)$ iff $e(v) \leq 3.14159$ and $e(v) \geq 2.71828$.

Stochastic transition systems. An *STS* is a triple $\mathcal{T} = (S, \Gamma, \to)$ comprising a measurable space S of *states*, a set Γ of *labels*, and a relation $\to\, \subseteq S \times \Gamma \times \mathsf{Prob}(S)$ of *transitions.* If S is a standard Borel space, then $\mathcal{T}$ is called *standard Borel.* Let $\mathcal{T}_a = (S_a, \Gamma, \to_a)$ and $\mathcal{T}_b = (S_b, \Gamma, \to_b)$ be STSs with the same sets of labels. A relation $R \subseteq S_a \times S_b$ is a *bisimulation for* $(\mathcal{T}_a, \mathcal{T}_b)$ if R is lr-total in $S_a \times S_b$ and for all $s_a \, R \, s_b$ and $\gamma \in \Gamma$ it holds: Given $\mu_a \in \mathsf{Prob}(S_a)$ where $s_a \to_a^\gamma \mu_a$, then there exists $\mu_b \in \mathsf{Prob}(S_b)$ such that $s_b \to_b^\gamma \mu_b$ and $\mu_a \, R^w \, \mu_b$. Vice versa, given $\mu_b \in \mathsf{Prob}(S_b)$ with $s_b \to_b^\gamma \mu_b$, then there is $\mu_a \in \mathsf{Prob}(S_a)$ where $s_a \to_a^\gamma \mu_a$ and $\mu_a \, R^w \, \mu_b$. We emphasize that a bisimulation is not required to be measurable. In the context of bisimulation an important question is how to lift a relation $R \subseteq S_a \times S_b$ to probability measures. However, there are other approaches using R-stable pairs instead [25], closely related to the weight lifting [53, 49]. Given STSs $\mathcal{T}_1 = (S_1, \Gamma_1, \to_1)$ and $\mathcal{T}_2 = (S_2, \Gamma_2, \to_2)$ and a set of synchronization labels $\mathsf{Sync} \subseteq \Gamma_1 \cap \Gamma_2$, their composition is the STS $\mathcal{T}_1 \parallel^{\otimes}_{\mathsf{Sync}} \mathcal{T}_2 = (S_1 \times S_2, \Gamma_1 \cup \Gamma_2, \to)$ with $\langle s_1, s_2\rangle \to^\gamma \mu_1 \otimes \mu_2$ iff the following holds [18]: If $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$, then $s_1 \to^\gamma \mu_1$ and $\mu_2 = \mathsf{Dirac}[s_2]$. If $\gamma \in \Gamma_2 \setminus \mathsf{Sync}$, then $\mu_1 = \mathsf{Dirac}[s_1]$ and $s_2 \to^\gamma \mu_2$. If $\gamma \in \mathsf{Sync}$, then $s_1 \to^\gamma \mu_1$ and $s_2 \to^\gamma \mu_2$.

Flows. By $\mathbb{T} = \mathbb{R}_{\geq 0}$ we denote the *time axis.* A *flow* is a function $\vartheta\colon \mathbb{T} \to \mathsf{Ev}(\mathsf{Var})$ that has the càdlàg property, i.e., ϑ is right continuous and has left limits everywhere. $\mathsf{Flow}(\mathsf{Var})$ denotes the set of all flows. Let $\vartheta \oplus T(t) = \vartheta(T+t)$ denote the *shift of* ϑ *at time* $T \in \mathbb{T}$ *by time* $t \in \mathbb{T}$. A subset F of $\mathsf{Flow}(\mathsf{Var})$ is *shift invariant* if $\vartheta \oplus T \in F$ for every $\vartheta \in F$ and $T \in \mathbb{T}$. In the theory of stochastic processes, the càdlàg property is well established as, amongst others, there is a topology on $\mathsf{Flow}(\mathsf{Var})$ such that $\mathsf{Flow}(\mathsf{Var})$ becomes a Polish space [8]. The exact definition of this topology is not relevant for our purposes. If $V \subseteq \mathsf{Var}$ and $\vartheta \in \mathsf{Flow}(\mathsf{Var})$, then $\vartheta_{|V} \in \mathsf{Flow}(V)$ is given by $\vartheta_{|V}(t) = \vartheta(t)_{|V}$ for all $t \in \mathbb{T}$. Given $V_1, V_2 \subseteq \mathsf{Var}$ where $V_1 \cap V_2 = \varnothing$ and $\vartheta_1 \in \mathsf{Flow}(V_1)$ and $\vartheta_2 \in \mathsf{Flow}(\mathsf{Var}_2)$, then $\vartheta_1 \uplus \vartheta_2 \in \mathsf{Flow}(V_1 \cup V_2)$ is the flow obtained by merging ϑ_1 and ϑ_2.

3 Composition of stochastic transition systems

We develop our approach towards the composition of STSs. As a preparation, we introduce spans first and give some insights on our mathematical theory for those. After that, we present the main contribution of the paper, namely our composition operator for STSs. We then give a congruence theorem having a quite challenging proof. Section 4 presents an application of our framework in the context of stochastic hybrid systems.

3.1 Spans

We will formalize dependencies for the composition of STSs using spans and span couplings, which is a generic and flexible formalism our framework benefits from in many occasions. The idea is to allow for arbitrary Polish spaces together with continuous functions that specify the relationships between the components. Various properties of spans then transfer to their probabilistic version, e.g., properness or the existence of inverses. This is an essential point in the context of stochastic models and hence also for STSs. We will then use spans within the definition of our composition in STS and later on also in the context of stochastic hybrid systems as a mathematical tool for our argumentation.

▶ **Definition 1.** A *span* is a tuple $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ consisting of Polish spaces X, X_1, and X_2 and continuous functions $\iota_1\colon X \to X_1$ and $\iota_2\colon X \to X_2$. We call $\mathcal{X}$ *proper*, if $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is compact in X for all compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$.

Intuitively, X denotes the joint state space of X_1 and X_2, where ι_1 and ι_2 are projective functions from X to X_1 and X_2, respectively. Properness connects topological aspects of the involved spaces. The following examples are natural instances of proper spans:

- $\mathcal{X}$ is a *Cartesian span* if $X = X_1 \times X_2$ and ι_1 and ι_2 are the natural projections.
- $\mathcal{X}$ is a *variable span* if $X_1 = \mathsf{Ev}(\mathsf{Var}_1)$, $X_2 = \mathsf{Ev}(\mathsf{Var}_2)$, and $X = \mathsf{Ev}(\mathsf{Var}_1 \cup \mathsf{Var}_2)$ for some sets of variables Var_1 and Var_2, and ι_1 and ι_2 are the natural projections.
- $\mathcal{X}$ is a *identity span* if $X = X_1 = X_2$ and $\iota_1(x) = x$ and $\iota_2(x) = x$ for all $x \in X$.

Span couplings are a crucial notion for our approach towards a composition operator in the next section. Given $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$, we call $\mu \in \mathsf{Prob}(X)$ a $\mathcal{X}$*-coupling of* (μ_1, μ_2) if $(\iota_1)_\sharp(\mu) = \mu_1$ and $(\iota_2)_\sharp(\mu) = \mu_2$. Recall that $(\iota_1)_\sharp$ and $(\iota_2)_\sharp$ denote the pushforwards of ι_1 and ι_2, respectively. A span coupling places two probability measures in the same probabilistic space specified by the span by exhibiting an adequate witness measure over pairs. Thus, the ordinary notion for couplings is generalized. For all x and μ we use $x_{|1}$, $x_{|2}$, $\mu_{|1}$, and $\mu_{|2}$ as shorthand notations for $\iota_1(x)$, $\iota_2(x)$, $(\iota_1)_\sharp(\mu)$, and $(\iota_2)_\sharp(\mu)$

respectively. Given $x_1 \in X_1$ and $x_2 \in X_2$, we then write $x_1 \, \mathcal{X} \, x_2$ if there exists $x \in X$ where $x_{|1} = x_1$ and $x_{|2} = x_2$. Similarly, we write $\mu_1 \, \mathcal{X}^c \, \mu_2$ if there is a $\mathcal{X}$-coupling of (μ_1, μ_2). We sometimes drop the projection functions from the tuple and refer to (X, X_1, X_2) as a span.

Probabilistic version. There are various operations for spans that yield complex spans out of some given basic spans. The question whether the operation preserves properness is important for practical purposes. For instance, a countable product of proper spans yields a proper span again. Within stochastic models, the following operation is important: For a span $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ its *probabilistic version* is given by the tuple $\mathsf{Prob}(\mathcal{X}) = (\mathsf{Prob}(X), \mathsf{Prob}(X_1), \mathsf{Prob}(X_2), (\iota_1)_\sharp, (\iota_2)_\sharp)$. Notice, $\mathsf{Prob}(\mathcal{X})$ involves all $\mathcal{X}$-couplings and $\mu_1 \, \mathcal{X}^c \, \mu_2$ iff $\mu_1 \, \mathsf{Prob}(\mathcal{X}) \, \mu_2$ for all $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$.

▶ **Proposition 2.** *The probabilistic version of a span is a span. Moreover, the probabilistic version of a proper span is proper as well.*

The claim regarding properness follows from Prokhorov's theorem [8], which characterizes relatively compact subsets of $\mathsf{Prob}(X)$: If $P \subseteq \mathsf{Prob}(X)$ is a set of probability measures, then P is relatively compact in $\mathsf{Prob}(X)$ iff P is tight in $\mathsf{Prob}(X)$, i.e., for every $\varepsilon \in \mathbb{R}_{>0}$ there is a compact set $K \subseteq X$ where $\mu(K) > 1 - \varepsilon$ for all $\mu \in P$.

Span inverse. In a compositional setting, the states of the components determine the states of the composed system. Within our approach, a state as an element of X in the composed system is not required to be uniquely determined: Given a span $\mathcal{X} = (X, X_1, X_2)$, $x_1 \in X_1$, and $x_2 \in X_2$, every $x \in X$ where $x_{|1} = x_1$ and $x_{|2} = x_2$ stands for a state in the composed system resulting from the states x_1 and x_2 of the components. However, in applications later it is important to have a mapping with additional properties: Given a span $\mathcal{X} = (X, X_1, X_2)$, a Borel function $f \colon X_1 \times X_2 \to X$ is called an $\mathcal{X}$*-inverse*, if for all $x_1 \in X_1$ and $x_2 \in X_2$, if $x_1 \, \mathcal{X} \, x_2$, then $f(x_1, x_2)_{|1} = x_1$ and $f(x_1, x_2)_{|2} = x_2$.

▶ **Theorem 3.** *Every proper span $\mathcal{X}$ has an $\mathcal{X}$-inverse.*

It follows $\mu_1 \, \mathcal{X}^c \, \mu_2$ iff $\mu_1 \, \mathsf{Rel}(\mathcal{X})^w \, \mu_2$ for all $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$, where $\mathsf{Rel}(\mathcal{X}) = \{\langle x_{|1}, x_{|2} \rangle \; ; \, x \in X\}$. Our proof of Theorem 3 is an application of a *measurable selection theorem* [10]: Take some $\hat{x} \in X$ and define $\Phi \colon X_1 \times X_2 \to 2^X$, $\Phi(x_1, x_2) = \{x \in X \; ; \, x_{|1} = x_1 \text{ and } x_{|2} = x_2\}$, if the set on the right-hand side is non-empty, and $\Phi(x_1, x_2) = \{\hat{x}\}$, otherwise. It suffices to argue that Φ admits a measurable selection, i.e., there is a measurable function $f \colon X_1 \times X_2 \to X$ where $f(x_1, x_2) \in \Phi(x_1, x_2)$ for all $x_1 \in X_1$ and $x_2 \in X_2$. To do so, we rely on results from descriptive set theory. Notice, together with Proposition 2, Theorem 3 yields an $\mathsf{Prob}(\mathcal{X})$-inverse if $\mathcal{X}$ is proper, which is an important observation for our discussions later. This is not obvious even for simple spans considering for instance the probabilistic version of a variable span. We remark that there are spans $\mathcal{X}$ that have no $\mathcal{X}$-inverses and thus, the properness assumption is important (cf. appendix).

3.2 Composition

A major objective in defining compositional frameworks is to separate the concerns of components specifying the operational behavior and composition operators addressing their interaction or coordination. We start with two STSs $\mathcal{T}_1 = (S_1, \Gamma_1, \to_1)$ and $\mathcal{T}_2 = (S_2, \Gamma_2, \to_2)$, where we assume S_1 and S_2 are Polish spaces. To declare the interactions between $\mathcal{T}_1$ and $\mathcal{T}_2$, we specify a set of synchronization labels $\mathsf{Sync} \subseteq \Gamma_1 \cap \Gamma_2$, a span $\mathcal{S} = (S, S_1, S_2)$ to characterizes the state space of the composition, and an so-called agreement $\mathcal{G} = (LC_1, LC_2)$ between $\mathcal{T}_1$ and $\mathcal{T}_2$. Here, LC_1 and LC_2 are so-called local constraints and for the moment, to present the central definition of this paper, it suffices to require $LC_1, LC_2 \subseteq S \times \mathsf{Prob}(S)$. Intuitively, we use local constraints to specify the behavior of local variables within local transitions (see below).

▶ **Definition 4.** We define the STS $\mathcal{T}_1 \parallel_{\mathcal{S},\mathcal{G},\mathsf{Sync}} \mathcal{T}_2 = (S, \Gamma_1 \cup \Gamma_2, \to)$, where for all $s \in S$, $\gamma \in \Gamma$, and $\mu \in \mathsf{Prob}(S)$ it holds $s \to^\gamma \mu$ iff one of the following three conditions hold:

- $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$ and $s_{|1} \to_1^\gamma \mu_{|1}$ and $s\, LC_2\, \mu$.
- $\gamma \in \Gamma_2 \setminus \mathsf{Sync}$ and $s\, LC_1\, \mu$ and $s_{|2} \to_2^\gamma \mu_{|2}$.
- $\gamma \in \mathsf{Sync}$ and $s_{|1} \to_1^\gamma \mu_{|1}$ and $s_{|2} \to_2^\gamma \mu_{|2}$.

To illustrate the crux of our composition operator, we regard the case where $\mathcal{S}$ is a Cartesian span, i.e., $S = S_1 \times S_2$. Former approaches [49, 18] assume that $\mathcal{T}_1$ and $\mathcal{T}_2$ behave stochastically independent in a synchronizing step, i.e., if $s_{|1} \to_1^\gamma \mu_1$ and $s_{|2} \to_2^\gamma \mu_2$, then $s \hookrightarrow^\gamma \mu_1 \otimes \mu_2$ in $\mathcal{T}_1 \parallel_H^\otimes \mathcal{T}_2$. Our operator does not rely on any stochastic assumptions: Instead of considering only the independent coupling we take all the couplings into account, i.e., if $s_{|1} \to_1^\gamma \mu_1$ and $s_{|2} \to_2^\gamma \mu_2$, then $s \to^\gamma \mu$ for all couplings μ of (μ_1, μ_2). During a discussion about the example from the introduction and SHMAs, we will see how additional stochastic information between the components can be incorporated within our general framework.

Local constraints. Our composition operator is indexed by a span, which determines the dependencies between the states of $\mathcal{T}_1$ and $\mathcal{T}_2$. For instance, one can specify shared and local variables using the variable span. When composing STSs, one has to ensure that local transitions and variables of the components behave in a compatible way. Let us illustrate this and regard again the case where $\mathcal{S}$ is a Cartesian span. If $\mathcal{T}_1$ performs a local transition, i.e., a transition that is labeled by some $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$, then the current state of $\mathcal{T}_2$ must not change. The properties of a local constraint should hence guarantee $s\, LC_2\, \mu$ iff $\mu_{|2} = \mathsf{Dirac}[s_{|2}]$. It then follows that $\langle s_1, s_2 \rangle \to^\gamma \mu_1 \otimes \mathsf{Dirac}[s_2]$ for all $s_2 \in S_2$ and $s_1 \to_1^\gamma \mu_1$ where $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$. Of course, the same discussion applies for $\mathcal{T}_1$ and the local constraint LC_1. This leads to the following requirements for a local constraint $LC_2 \subseteq S \times \mathsf{Prob}(S)$:

- For all $s \in S$ and $\mu \in \mathsf{Prob}(S)$, if $\mu_{|2} = \mathsf{Dirac}[s_{|2}]$, then $s\, LC_2\, \mu$.
- For all $s\, LC_2\, \mu$ and $\mu' \in \mathsf{Prob}(S)$, if $\mu_{|1} = \mu'_{|1}$ and $\mu_{|2} = \mu'_{|2}$, then $s\, LC_2\, \mu'$.
- For all $s\, LC_2\, \mu$, if $\mu_{|1}\, \mathcal{S}^c\, \mathsf{Dirac}[s_{|2}]$, then μ is a $\mathcal{S}$-coupling of $(\mu_{|1}, \mathsf{Dirac}[s_{|2}])$.

The requirements for LC_1 are similar. Intuitively, the first requirement for LC_2 ensures that the STS $\mathcal{T}_2$ cannot block a local transition of $\mathcal{T}_1$ which is not critical from the view of $\mathcal{T}_2$, i.e., variables of $\mathcal{T}_2$ are not affected within the transition of $\mathcal{T}_1$. Thus, such local transition of $\mathcal{T}_1$ are independent of $\mathcal{T}_2$ and can happen autonomously. Different couplings of given probability measures cannot be distinguished within local constraints imposed by the second property. The third requirement intuitively demands that whenever $\mathcal{T}_1$ performs a local transition where no local variables of $\mathcal{T}_2$ are modified, the state of $\mathcal{T}_2$ must not change. In case of a Cartesian span the above requirements yield

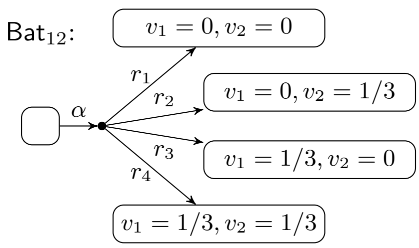
$$LC_2 = \{\langle \langle s_1, s_2 \rangle, \mu_1 \otimes \mathsf{Dirac}[s_2] \rangle \;;\; s_1 \in S_1 \text{ and } s_2 \in S_2 \text{ and } \mu_1 \in \mathsf{Prob}(S_1)\} \quad \text{and}$$
$$LC_1 = \{\langle \langle s_1, s_2 \rangle, \mathsf{Dirac}[s_1] \otimes \mu_2 \rangle \;;\; s_1 \in S_1 \text{ and } s_2 \in S_2 \text{ and } \mu_2 \in \mathsf{Prob}(S_2)\}.$$

Hence, the agreement $\mathcal{G}$ is uniquely determined by STSs $\mathcal{T}_1$ and $\mathcal{T}_2$. We thus simply write $\mathcal{T}_1 \parallel_{\times,\mathsf{Sync}} \mathcal{T}_2$ instead of $\mathcal{T}_1 \parallel_{\mathcal{S},\mathcal{G},\mathsf{Sync}} \mathcal{T}_2$. Observe that $\mathcal{T}_1 \parallel_{\times,\mathsf{Sync}} \mathcal{T}_2$ and $\mathcal{T}_1 \parallel_{\mathsf{Sync}}^\otimes \mathcal{T}_2$ are not bisimilar in general. This is due to the fact that our composition operator does not incorporate any stochastic assumptions concerning the interaction of $\mathcal{T}_1$ and $\mathcal{T}_2$. In case where $\mathcal{S}$ is a variable span, i.e., $S_1 = \mathsf{Ev}(\mathsf{Var}_1)$, $S_2 = \mathsf{Ev}(\mathsf{Var}_2)$, and $S = \mathsf{Ev}(\mathsf{Var}_1 \cup \mathsf{Var}_2)$ for some sets of variables Var_1 and Var_2, there are more possible local constraints:

$$LC_2' = \{\langle e, \eta \rangle \in S \times \mathsf{Prob}(S) \; ; \; \eta_{|\mathsf{SVar}} = \mathsf{Dirac}[e_{|\mathsf{SVar}}] \text{ implies } \eta_{|\mathsf{Var}_2} = \mathsf{Dirac}[e_{|\mathsf{Var}_2}]\},$$
$$LC_2'' = \{\langle e, \eta \rangle \in S \times \mathsf{Prob}(S) \; ; \; \eta_{|\mathsf{LVar}_2} = \mathsf{Dirac}[e_{|\mathsf{LVar}_2}]\}, \quad \text{and}$$
$$LC_2''' = \{\langle e, \eta \rangle \in S \times \mathsf{Prob}(S) \; ; \; \eta_{|\mathsf{Var}_2} = \mathsf{Dirac}[e_{|\mathsf{Var}_2}]\}$$

are local constraints and $LC_2' \supseteq LC_2'' \supseteq LC_2'''$. Here, $\mathsf{LVar}_2 = \mathsf{Var}_2 \setminus \mathsf{Var}_1$ and $\mathsf{SVar} = \mathsf{Var}_1 \cap \mathsf{Var}_2$. Considering for instance LC_2'', all the variables in LVar_2 cannot be modified within a local transition of $\mathcal{T}_1$. Constraint LC_2''' is more restrictive: Here, all the variables in Var_2 are controlled by $\mathcal{T}_2$ and cannot be modified in a local transition of $\mathcal{T}_1$, i.e., variables in SVar can be observed by $\mathcal{T}_1$ only. It turns out $LC_2' \supseteq LC_2$ for every local constraint LC_2. Every local constraint hence enjoys the property that variables in Var_2 must not be adapted within a local transition of $\mathcal{T}_1$ if the evaluations of the variables in SVar remain the same.

Example from the introduction. We return to the introductory stochastic systems illustrated in Section 1. Of course, Bat_1, Bat_2, and Dev can be seen as STSs with sets of states $\mathsf{Ev}(\{v_1\})$, $\mathsf{Ev}(\{v_2\})$, and $\mathsf{Ev}(\{v_1, v_2\})$, respectively. When composing them, we need not to worry about local constraints as there is only one synchronization action α. In what follows, we rely on the obvious variable spans. The composition of Bat_1 and Bat_2 yields the STS Bat_{12} depicted on the right. There are infinitely many transitions: Every solution of the linear equation system $r_1 + r_2 = r_1 + r_3 = r_2 + r_4 = r_3 + r_4 = 1/2$ where $r_1, r_2, r_3, r_4 \in [0, 1]$ represents a coupling of the involved measures. When composing Bat_{12} and Dev, the set of all couplings is refined. We are moreover able to handle more complex stochastic information that depend on the operational behavior of the components. To illustrate this, assume systems which result from Bat_1 and Bat_2 such that α can be executed repeatedly (e.g., add some local transitions back to the blank state). An additional component might encode that, if the system has crashed repeatedly in the past, the event that the stored energy drops to 0 in both batteries at the same time becomes more likely within an execution of α. We emphasize that the ordinary composition of STSs can be expressed within our framework using an additional component (cf. appendix).



3.3 Congruence

In the context of process calculi, an important issue of bisimulation is the compatibility with syntactic operators in the process calculus, such as parallel composition. We show that bisimulation is a congruence for our composition operator under reasonable side-constraints, i.e., our composition operator enjoys the substitution property with respect to bisimulation. Suppose STSs $\mathcal{T}_{a1} = (S_{a1}, \Gamma_1, \rightarrow_{a1})$, $\mathcal{T}_{a2} = (S_{a2}, \Gamma_2, \rightarrow_{a2})$, $\mathcal{T}_{b1} = (S_{b1}, \Gamma_1, \rightarrow_{b1})$, and $\mathcal{T}_{b2} = (S_{b2}, \Gamma_2, \rightarrow_{b2})$ such that $\mathcal{T}_{a1} \sim \mathcal{T}_{b1}$ and $\mathcal{T}_{a2} \sim \mathcal{T}_{b2}$. Define

$$\mathcal{T}_a = \mathcal{T}_{a1} \parallel_{\mathcal{S}_a, \mathcal{G}_a, \mathsf{Sync}} \mathcal{T}_{a2} \quad \text{and} \quad \mathcal{T}_b = \mathcal{T}_{b1} \parallel_{\mathcal{S}_b, \mathcal{G}_b, \mathsf{Sync}} \mathcal{T}_{b2},$$

where $\mathsf{Sync} \subseteq \Gamma_1 \cap \Gamma_2$, $\mathcal{S}_a = (S_a, S_{a1}, S_{a2})$ and $\mathcal{S}_b = (S_b, S_{b1}, S_{b2})$ are proper spans, and $\mathcal{G}_a = (LC_{a1}, LC_{a2})$ and $\mathcal{G}_b = (LC_{b1}, LC_{b2})$ are agreements. Assume R_1 is a bisimulation for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and R_2 is a bisimulation for $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$ and define

$$R_1 \wedge R_2 = \{\langle s_a, s_b \rangle \in S_a \times S_b \; ; \; s_{a|1} \, R_1 \, s_{b|1} \text{ and } s_{a|2} \, R_2 \, s_{b|2}\}.$$

We aim to show that $R_1 \wedge R_2$ is a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$ and hence $\mathcal{T}_a \sim \mathcal{T}_b$. However, we cannot expect this result without any compatibility requirements for the involved spans and agreements, since important relationships concerning the communication of the components are determined within our composition operator. This motivates the following notions: We refer to the tuple $\mathcal{C} = (\mathcal{S}_a, \mathcal{S}_b, R_1, R_2)$ as *span connection* and call $\mathcal{C}$ *adequate* if for all $\mu_{a1}\, R_1^w\, \mu_{b1}$ and $\mu_{a2}\, R_2^w\, \mu_{b2}$ it holds $\mu_{a1}\, \mathcal{S}_a^c\, \mu_{a2}$ iff $\mu_{b1}\, \mathcal{S}_b^c\, \mu_{b2}$. Intuitively, adequacy requires that the existence of span couplings is preserved by the relations R_1 and R_2. Observe, if $\mathcal{S}_a$ and $\mathcal{S}_b$ are Cartesian spans, then $\mathcal{C}$ is always adequate. The local constraints LC_{a2} and LC_{b2} are called $\mathcal{C}$*-bisimilar* if for all $s_a\,(R_1 \wedge R_2)\, s_b$ holds:

- For all $\mu_a \in \mathsf{Prob}(S_a)$ and $\mu_{b1} \in \mathsf{Prob}(S_{b1})$, if $s_a\, LC_{a2}\, \mu_a$ and $\mu_{a|1}\, R_1^w\, \mu_{b1}$, then there is $\mu_b \in \mathsf{Prob}(S_b)$ where $s_b\, LC_{b2}\, \mu_b$, $\mu_{b|1} = \mu_{b1}$, and $\mu_{a|2}\, R_2^w\, \mu_{b|2}$.
- For all $\mu_b \in \mathsf{Prob}(S_b)$ and $\mu_{a1} \in \mathsf{Prob}(S_{a1})$, if $s_b\, LC_{b2}\, \mu_b$ and $\mu_{a1}\, R_1^w\, \mu_{b|1}$, then there is $\mu_a \in \mathsf{Prob}(S_a)$ where $s_a\, LC_{a2}\, \mu_a$, $\mu_{a|1} = \mu_{a1}$, and $\mu_{a|2}\, R_2^w\, \mu_{b|2}$.

LC_{a1} and LC_{b1} are called $\mathcal{C}$*-bisimilar* if analogous properties are fulfilled. Observe that the stated requirement is motivated by the definition of bisimulation in the sense that each element of a local constraint LC_{a2} can be mimicked by LC_{b2} regarding the relations R_1 and R_2. If LC_{a2} and LC_{b2} as well as LC_{a1} and LC_{b1} are $\mathcal{C}$-bisimilar, respectively, then we refer to $\mathcal{G}_a$ and $\mathcal{G}_b$ as $\mathcal{C}$*-bisimilar*.

▶ **Theorem 5.** *If the span connection $\mathcal{C}$ is adequate and the agreements $\mathcal{G}_a$ and $\mathcal{G}_b$ are $\mathcal{C}$-bisimilar, then $R_1 \wedge R_2$ is a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$.*

The challenging part of the proof can be summarized by the following claim (cf. appendix): Let $\mu_a \in \mathsf{Prob}(S_a)$, $\mu_{b1} \in \mathsf{Prob}(S_{b1})$, and $\mu_{b2} \in \mathsf{Prob}(S_{b2})$ where $\mu_{a|1}\, R_1^w\, \mu_{b1}$ and $\mu_{a|2}\, R_2^w\, \mu_{b2}$. Then there is an $\mathcal{S}_b$-coupling μ_b of (μ_{b1}, μ_{b2}) such that $\mu_a\,(R_1 \wedge R_2)^w\, \mu_b$. Our proof of this claim proceeds as follows. Assume W_1 is a weight function for $(\mu_{a|1}, R_1, \mu_{b1})$ and W_2 is a weight function for $(\mu_{a|2}, R_2, \mu_{b2})$. Using disintegration of measures [39], there are Markov kernels $k_1 \colon S_{a1} \to \mathsf{Prob}(S_{a2})$ and $k_2 \colon S_{b1} \to \mathsf{Prob}(S_{b2})$ such that $W_1 = \mu_{a|1} \rtimes k_1$ and $W_2 = \mu_{a|2} \rtimes k_2$. The crucial point is now to argue that there is a Markov kernel $k \colon S_a \to \mathsf{Prob}(S_b)$ where $k(s_a)$ is an $\mathcal{S}_b$-coupling of $(k_1(s_{a|1}), k_2(s_{a|2}))$ for μ_a-almost all $s_a \in S_a$. Here, we make use of an $\mathcal{S}_b$-inverse (cf. Theorem 3). With this Markov kernel at hand, we define $W \in \mathsf{Prob}(S_a \times S_b)$ by $W = \mu \rtimes k$ and $\mu_b \in \mathsf{Prob}(S_b)$ by $\mu_b(M_b) = W(S_a \times M_b)$. It turns out that μ_b is an appropriate $\mathcal{S}_b$-coupling. To summarize, we defined a potential weight function W out of the weight functions W_1 and W_2 and then introduced the measure μ_b via W.

Path measures. When resolving the non-determinism in STSs using schedulers, one obtains a probability measure – the path measure – on the set of all infinite paths of the STS [18]. Besides our congruence result, we expect compatibility of path measures induced by schedulers in our compositional framework. To provide an intuition, assume STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ and let $\mathcal{T}$ be an STSs obtained by a composition involving $\mathcal{T}_1$ and $\mathcal{T}_2$. Assume that $\mathfrak{S}_1$ and $\mathfrak{S}_2$ are schedulers for $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively, and $\mathfrak{S}$ is a scheduler for $\mathcal{T}$. If $\mathfrak{S}$ satisfies certain compatibility requirements regarding $\mathfrak{S}_1$ and $\mathfrak{S}_2$, one can show that the induced path measure for $\mathcal{T}$ is a coupling of the corresponding path measures for $\mathcal{T}_1$ and $\mathcal{T}_2$. Here, we consider a natural span that connects the sets of all infinite paths of $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}$.

4 Stochastic hybrid motion automata

We apply our general results of the preceding sections and develop a compositional modeling framework for stochastic hybrid systems. The formal definition of our model relies on a standard schema of hybrid automata [3, 34, 31], i.e., there are a discrete control structure consisting of locations and jumps in-between, and continuous variables whose values evolve according to a flow formalized by a motion function. Within a jump, the variables can be updated instantaneously. The novelty of our approach is that every jump is indexed by a set of those variables that are not affected in the corresponding discrete step. As a consequence, the adjustment of flows is always accompanied by a specific command.

Syntax. Every jump in our hybrid-automaton model is labeled by a command: Given a set Var of variables and a set Act of actions, a *command on* $(\mathsf{Var}, \mathsf{Act})$ is a tuple $\langle c, \alpha, V, \mathsf{upd}\rangle$ consisting of a *guard* $c \in \mathsf{Cond}(\mathsf{Var})$, an *action* $\alpha \in \mathsf{Act}$, a set of *disabled* variables $V \subseteq \mathsf{Var}$, and an *(non-deterministic) update* $\mathsf{upd}\colon \mathsf{Ev}(\mathsf{Var}) \to 2^{\mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}))}$ where $\eta_{|V} = \mathsf{Dirac}[e_{|V}]$ for all $\eta \in \mathsf{upd}(e)$ and $e \in \mathsf{Ev}(\mathsf{Var})$. $\mathsf{Cmd}(\mathsf{Var}, \mathsf{Act})$ denotes the set of all commands on $(\mathsf{Var}, \mathsf{Act})$. Intuitively, a jump is enabled if the current variable evaluation satisfies the guard. The action name indicates whether the jump is an internal location switch or subject to an interaction with another component. The set of disabled variables specifies those variables which are not affected within the jump. This also clarifies the additional requirement for updates.

▶ **Definition 6.** An *SHMA* is a tuple $(\mathsf{Loc}, \mathsf{Var}, \mathsf{Act}, \mathsf{Inv}, \mathsf{Mot}, \twoheadrightarrow)$ where Loc is a finite set of *locations*, Var is a set of variables, Act is a set of actions, $\mathsf{Inv}\colon \mathsf{Loc} \to \mathsf{Cond}(\mathsf{Var})$ is an *invariant function*, $\mathsf{Mot}\colon \mathsf{Loc} \to 2^{\mathsf{Flow}(\mathsf{Var})}$ is a *motion function* which assigns a shift-invariant set of flows to every location, and $\twoheadrightarrow\ \subseteq \mathsf{Loc} \times \mathsf{Cmd}(\mathsf{Var}, \mathsf{Act}) \times \mathsf{Prob}(\mathsf{Loc})$ is a *jump relation.*

We write $l \dashv\!\![\, cmd \,]\!\!\twoheadrightarrow \lambda$ instead of $\langle l, cmd, \lambda\rangle \in \twoheadrightarrow$. The behavior in a location l depends on the current variable evaluation e. In a discrete step, a jump $l \dashv\!\![\, c, \alpha, V, \mathsf{upd} \,]\!\!\twoheadrightarrow \lambda$ where $e \models c$ is chosen non-deterministically. Then, action α is executed and a successor location is sampled according to λ. The evaluation of the variables changes according to a non-deterministically chosen probability measure contained in $\mathsf{upd}(e)$. Entering a location l', a flow in $\mathsf{Mot}(l')$ is also chosen non-deterministically and the variables then evolve according to this flow.

Semantics. Every SHMA $\mathcal{H} = (\mathsf{Loc}, \mathsf{Var}, \mathsf{Act}, \mathsf{Inv}, \mathsf{Mot}, \twoheadrightarrow)$ can be interpreted as an STS resulting from unfolding. In what follows, $S = \mathsf{Loc} \times \mathsf{Flow}(\mathsf{Var})$ denotes the set of states. Notice, S constitutes a Polish space as $\mathsf{Flow}(\mathsf{Var})$ is known to be a Polish space [8]. Intuitively, a state $\langle l, \vartheta\rangle$ represents the actual location l and the current active flow ϑ, i.e., ϑ corresponds to the flow chosen in the preceding jump. Moreover, $\vartheta(0)$ stands for the present variable evaluation. We call $\langle l, \vartheta\rangle$ *well-formed* if $\vartheta \in \mathsf{Mot}(l)$ and $\vartheta(0) \models \mathsf{Inv}(l)$.

There are two kinds of transitions within our STS for $\mathcal{H}$, namely transitions where time passes and transitions corresponding to a jump. Time can pass in a location l as long as the flow does not violate the invariant $\mathsf{Inv}(l)$. Transitions for jumps are more intricate. Assume $l \dashv\!\![\, c, \alpha, V, \mathsf{upd} \,]\!\!\twoheadrightarrow \lambda$ is enabled in state $\langle l, \vartheta\rangle$, i.e., $e \models c$ where $e = \vartheta(0)$. Basically, jumps in SHMAs proceed in two phases: First, a successor location and a variable evaluation are sampled according to λ and some $\eta \in \mathsf{upd}(e)$, respectively. In the second phase, a flow is chosen non-deterministically for those variables which are not disabled, i.e., the variables in $\mathsf{Var} \setminus V$. This is formalized as follows: A *flow adapter for* (ϑ, V) is a Borel function $\chi\colon \mathsf{Loc} \times \mathsf{Ev}(\mathsf{Var}) \to \mathsf{Flow}(\mathsf{Var})$ such that for all $l' \in \mathsf{Loc}$ and $e', \tilde{e}' \in \mathsf{Ev}(\mathsf{Var})$:

$$\chi(l', e')_{|V} = \vartheta_{|V} \quad \text{and} \quad e'_{|\mathsf{Var}\setminus V} = \tilde{e}'_{|\mathsf{Var}\setminus V} \text{ implies } \chi(l', e')_{|\mathsf{Var}\setminus V} = \chi(l', \tilde{e}')_{|\mathsf{Var}\setminus V}.$$

Intuitively, if state $\langle l', e'\rangle$ is sampled within the first phase of a jump, then $\chi(l', e')$ represents the new flow, i.e., the flow which determines the evolution of variables in a subsequent time passage. The first condition for a flow adapter requires that the flow for disabled variables is not allowed to change. The required implication ensures that a flow is chosen independently of the disabled variables. This is important for our compositional approach, as we want to make sure that the choice of a new flow in an SHMA obtained by composition does not depend on the local variables of the respective communication partners. If χ is a flow adapter, then we define the auxiliary function $\hat{\chi}\colon \mathsf{Loc} \times \mathsf{Ev}(\mathsf{Var}) \to S$, $\hat{\chi}(l,e) = \langle l, \chi(l,e)\rangle$.

▶ **Definition 7.** The *semantics of* $\mathcal{H}$ is given by the STS $[\![\mathcal{H}]\!] = (S, \mathbb{T} \cup \mathsf{Act}, \to)$, where $\to$ is the smallest relation satisfying the following requirements for all well-formed states $s = \langle l, \vartheta\rangle$:

- For all $T \in \mathbb{T}$, if $\vartheta(t) \models \mathsf{Inv}(l)$ for every $t \in [0,T]$, then $s \to^t \mathsf{Dirac}[\langle l, \vartheta \oplus T\rangle]$.
- For all $l \mathrel{-\!\![} c, \alpha, V, \mathsf{upd} \mathrel{]\!\!\twoheadrightarrow} \lambda$, $\eta \in \mathsf{upd}(e)$, couplings ν of (λ, η), and flow adapter χ for (ϑ, V), if $e \models c$ and for ν-almost all $\langle l', e'\rangle \in \mathsf{Loc} \times \mathsf{Ev}(\mathsf{Var})$ the state $\hat{\chi}(l', e')$ is well-formed, then $s \to^\alpha \hat{\chi}_\sharp(\nu)$. Here, we abbreviate $e = \vartheta(0)$.

An SHMA almost surely enters a well-formed state, i.e., if $s \to^\gamma \mu$ where $\gamma \in \mathbb{T} \cup \mathsf{Act}$, then s' is well-formed for μ-almost all $s' \in S$. We emphasize that for our approach concerning the adaption of flows it is crucial that the current flow is part of a state. Otherwise, it would be not possible to ensure that the flow for disabled variables is not allowed to change.

Composition. We now introduce a composition operator for SHMAs. For $i \in \{1,2\}$ let $\mathcal{H}_i = (\mathsf{Loc}_i, \mathsf{Var}_i, \mathsf{Act}_i, \mathsf{Inv}_i, \mathsf{Mot}_i, \twoheadrightarrow_i)$ be SHMAs. When running $\mathcal{H}_1$ and $\mathcal{H}_2$ in parallel, $\mathcal{H}_1$ and $\mathcal{H}_2$ synchronize on all actions contained in $\mathsf{Act}_1 \cap \mathsf{Act}_2$ and the variables in $\mathsf{Var}_1 \cap \mathsf{Var}_2$ are shared, i.e., $\mathsf{Var}_1 \setminus \mathsf{Var}_2$ and $\mathsf{Var}_2 \setminus \mathsf{Var}_1$ represent the sets of the respective local variables. Abbreviate $\mathsf{Loc} = \mathsf{Loc}_1 \times \mathsf{Loc}_2$, $\mathsf{Var} = \mathsf{Var}_1 \cup \mathsf{Var}_2$, and $\mathsf{Act} = \mathsf{Act}_1 \cup \mathsf{Act}_2$. Let upd_1 and upd_2 be updates for Var_1 and Var_2, respectively. The Var-*lifting of* $(\mathsf{upd}_1, \mathsf{upd}_2)$ is the update upd for Var such that for all $e \in \mathsf{Ev}(\mathsf{Var})$, $\mathsf{upd}(e)$ consists of all $\eta \in \mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}))$ where $\eta_{|\mathsf{Var}_1} = \eta_1$ and $\eta_{|\mathsf{Var}_2} = \eta_2$ for some $\eta_1 \in \mathsf{upd}(e_{|\mathsf{Var}_1})$ and $\eta_2 \in \mathsf{upd}(e_{|\mathsf{Var}_2})$. We define Var-liftings with respect to an update accordingly, i.e., upd is a Var-lifting of upd_1 if for all $e \in \mathsf{Ev}(\mathsf{Var})$, $\mathsf{upd}(e)$ consists of all $\eta \in \mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}))$ where $\eta_{|\mathsf{Var}_1} = \eta_1$ for some $\eta_1 \in \mathsf{upd}(e_{|\mathsf{Var}_1})$ and $\eta_{|\mathsf{Var}\setminus\mathsf{Var}_1} = \mathsf{Dirac}[e_{|\mathsf{Var}\setminus\mathsf{Var}_1}]$. Notice, the definition of Var-liftings involves couplings concerning a variable span, which provides a connection to the preceding sections.

▶ **Definition 8.** $\mathcal{H}_1 \parallel \mathcal{H}_2 = (\mathsf{Loc}, \mathsf{Var}, \mathsf{Act}, \mathsf{Inv}, \mathsf{Mot}, \twoheadrightarrow)$ is the SHMA with $\mathsf{Inv}(l_1, l_2) = \mathsf{Inv}_1(l_1) \wedge \mathsf{Inv}_2(l_2)$ and $\mathsf{Mot}(l_1, l_2) = \{\vartheta \in \mathsf{Flow}(\mathsf{Var})\ ;\ \vartheta_{|\mathsf{Var}_1} \in \mathsf{Mot}_1(l_1) \text{ and } \vartheta_{|\mathsf{Var}_2} \in \mathsf{Mot}_2(l_2)\}$ for all $\langle l_1, l_2\rangle \in \mathsf{Loc}$ and $\twoheadrightarrow$ is the smallest relation such that $\langle l_1, l_2\rangle \mathrel{-\!\![} c, \alpha, V, \mathsf{upd} \mathrel{]\!\!\twoheadrightarrow} \lambda$, if λ is a coupling of $\lambda_1 \in \mathsf{Prob}(\mathsf{Loc}_1)$ and $\lambda_2 \in \mathsf{Prob}(\mathsf{Loc}_2)$ and one of the following statements hold:

- $\alpha \in \mathsf{Act}_1 \setminus \mathsf{Act}_2$, $\lambda_2 = \mathsf{Dirac}[l_2]$, and there is $l_1 \mathrel{-\!\![} c_1, \alpha, V_1, \mathsf{upd}_1 \mathrel{]\!\!\twoheadrightarrow_1} \lambda_1$ such that $c = c_1$, $V = V_1 \cup (\mathsf{Var}_2 \setminus \mathsf{Var}_1)$, and upd is the Var-lifting of upd_1.
- $\alpha \in \mathsf{Act}_2 \setminus \mathsf{Act}_1$, $\lambda_1 = \mathsf{Dirac}[l_1]$, and there is $l_2 \mathrel{-\!\![} c_2, \alpha, V_2, \mathsf{upd}_2 \mathrel{]\!\!\twoheadrightarrow_2} \lambda_2$ such that $c = c_2$, $V = V_2 \cup (\mathsf{Var}_1 \setminus \mathsf{Var}_2)$, and upd is the Var-lifting of upd_2.
- $\alpha \in \mathsf{Act}_1 \cap \mathsf{Act}_2$ and there are $l_1 \mathrel{-\!\![} c_1, \alpha, V_1, \mathsf{upd}_1 \mathrel{]\!\!\twoheadrightarrow_1} \lambda_1$ and $l_2 \mathrel{-\!\![} c_2, \alpha, V_2, \mathsf{upd}_2 \mathrel{]\!\!\twoheadrightarrow_2} \lambda_2$ where $c = c_1 \wedge c_2$, $V = V_1 \cup V_2$, and upd is the Var-lifting of $(\mathsf{upd}_1, \mathsf{upd}_2)$.

When composing SHMAs, local variables of participating SHMAs become disabled for corresponding internal jumps. Within our semantics, flow adapters thus ensure that the adaption of flows in internal jumps in $\mathcal{H}_1 \parallel \mathcal{H}_2$ are independent of the local variables of the respective communication partners. Moreover, flows for local variables of $\mathcal{H}_2$ cannot be adapted within an internal jump of $\mathcal{H}_1$ and vice versa. It is easy to see that the composition operator for SHMAs is commutative and associative.

Congruence. We aim for a congruence theorem for SHMAs relying on Theorem 5. For this, we relate the composition of SHMAs with our general approach towards a composition of STSs, i.e., we represent the STS $[\![\mathcal{H}_1 \parallel \mathcal{H}_1]\!]$ as a composition involving the components $[\![\mathcal{H}_1]\!]$ and $[\![\mathcal{H}_2]\!]$. Notice that sampling a successor location in $\mathcal{H}_1 \parallel \mathcal{H}_2$ happens according to a coupling measure. This observation also applies when combining measures for locations and variable evaluations within our semantics of SHMAs. To this end, it is easy to define the corresponding span $\mathcal{S}$ and agreement $\mathcal{G}$ such that

$$[\![\mathcal{H}_1 \parallel \mathcal{H}_2]\!] = [\![\mathcal{H}_1]\!] \parallel_{\mathcal{S},\mathcal{G},\mathsf{Act}_1\cap\mathsf{Act}_2} [\![\mathcal{H}_2]\!].$$

More precisely, $\mathcal{S}$ is a span arising from a Cartesian span for the locations and a span for the sets of flows. For the agreement $\mathcal{G}$, we regard local constraints where the shared variables can be modified by both involved systems $\mathcal{H}_1$ and $\mathcal{H}_2$. The obtained representation of $[\![\mathcal{H}_1 \parallel \mathcal{H}_2]\!]$ underpins again the flexibility of our composition operator for STS.

We rephrase Theorem 5 in the context SHMAs. Two SHMAs are bisimilar if their semantics in terms of STSs are bisimilar. Let $\mathcal{H}_{a1}$ and $\mathcal{H}_{b1}$ be SHMAs with the same sets of variables Var_1 and actions Act_1 and similar, let $\mathcal{H}_{a2}$ and $\mathcal{H}_{b2}$ be SHMAs with variables Var_2 and actions Act_2. Abbreviate $\mathsf{LVar}_1 = \mathsf{Var}_1 \setminus \mathsf{Var}_2$, $\mathsf{LVar}_2 = \mathsf{Var}_2 \setminus \mathsf{Var}_1$, and $\mathsf{SVar} = \mathsf{Var}_1 \cap \mathsf{Var}_2$.

▶ **Theorem 9.** *Let R_1 and R_2 be bisimulations for $(\mathcal{H}_{a1}, \mathcal{H}_{b1})$ and $(\mathcal{H}_{a2}, \mathcal{H}_{b2})$, respectively. $\mathcal{H}_{a1} \parallel \mathcal{H}_{a2}$ and $\mathcal{H}_{b1} \parallel \mathcal{H}_{b2}$ are bisimilar if R_1 and R_2 do not involve shared variables, i.e.,*

$$\begin{aligned} R_1 = \{\langle\langle l_{a1}, \vartheta_{a1|\mathsf{LVar}_1} \uplus \vartheta^S\rangle, \langle l_{b1}, \vartheta_{b1|\mathsf{LVar}_1} \uplus \vartheta^S\rangle\rangle \ ; \\ \langle l_{a1}, \vartheta_{a1}\rangle\, R_1\, \langle l_{b1}, \vartheta_{b1}\rangle \text{ and } \vartheta^S \in \mathsf{Flow}(\mathsf{SVar})\}, \\ R_2 = \{\langle\langle l_{a2}, \vartheta_{a2|\mathsf{LVar}_2} \uplus \vartheta^S\rangle, \langle l_{b2}, \vartheta_{b2|\mathsf{LVar}_2} \uplus \vartheta^S\rangle\rangle \ ; \\ \langle l_{a2}, \vartheta_{a2}\rangle\, R_2\, \langle l_{b2}, \vartheta_{b2}\rangle \text{ and } \vartheta^S \in \mathsf{Flow}(\mathsf{SVar})\}. \end{aligned}$$

Our requirement that R_1 and R_2 do not distinguish between shared variables yields the compatibility assumption required for Theorem 5. Our proof then simply exploits the representation of $[\![\mathcal{H}_{a1} \parallel \mathcal{H}_{a2}]\!]$ and $[\![\mathcal{H}_{b1} \parallel \mathcal{H}_{b2}]\!]$ in terms of a composition of STSs.

5 Concluding remarks

In this paper, we introduced a generic parallel-composition operator for STSs and SHMAs. The essential new feature that distinguishes the novel composition from previous ones is that it uses the mathematical concepts of spans and couplings to model the effect of composing (potentially dependent) stochastic behaviors. The latter is crucial for systems where the components communicate via shared variables. A further feature of the novel stochastic-hybrid-system model (SHMA) is that the adaption of flows depends on commands rather happening on arbitrary occasions. We proved important algebraic properties in the context of composition, e.g., congruence with respect to bisimulation. This shows that even within our generic operator one does not have to forgo desired properties of compositional systems. There is plenty room for further elaborations. Firstly, we are going to develop a mathematical theory for SHMA that also involves stochastic flows. Furthermore, we will work on a modeling language for couplings and spans in order to obtain a theoretical basis for practical tools. Also other kinds of models, where spans yield a powerful approach for compositional modeling, could be investigated. Moreover, our approach concerning couplings as a modeling formalism enables many new verification questions, e.g., for directly reasoning about the coordination between components.

References

1 Luca de Alfaro. *Formal Verification of Probabilistic Systems.* PhD thesis, University of Stanford, 1997.

2 Luca de Alfaro. Stochastic transition systems. In *Proceedings of the 9th International Conference on Concurrency Theory (CONCUR)*, volume 1446 of *Lecture Notes in Computer Science*, pages 423–438. Springer, 1998.

3 Rajeev Alur, Costas Courcoubetis, Nicolas Halbwachs, Thomas A. Henzinger, Pei-Hsin Ho, Xavier Nicollin, Alfredo Olivero, Joseph Sifakis, and Sergio Yovine. The algorithmic analysis of hybrid systems. *Theoretical Computer Science*, 138(1):3–34, 1995.

4 Rajeev Alur and David L. Dill. A theory of timed automata. *Theoretical Computer Science*, 126:183–235, 1994.

5 Rajeev Alur and Thomas A. Henzinger. Modularity for timed and hybrid systems. In *Proceedings of the 8th International Conference on Concurrency Theory (CONCUR)*, volume 1243 of *Lecture Notes in Computer Science*, pages 74–88. Springer, 1997.

6 Giorgio Bacci, Giovanni Bacci, Kim G. Larsen, and Radu Mardare. Bisimulation on Markov processes over arbitrary measurable spaces. In *Horizons of the Mind. A Tribute to Prakash Panangaden*, volume 8464 of *Lecture Notes in Computer Science*, pages 76–95. Springer, 2014.

7 Christel Baier, Bettina Engelen, and Mila Majster-Cederbaum. Deciding bisimilarity and similarity for probabilistic processes. *Journal of Computer and System Sciences*, 60(1):187–231, 2000.

8 Patrick Billingsley. *Convergence of Probability Measures.* Wiley-Interscience, 2 edition, 1999.

9 Vladimir Igorevich Bogachev. *Measure Theory*, volume 1. Springer, 2007.

10 Vladimir Igorevich Bogachev. *Measure Theory*, volume 2. Springer, 2007.

11 Henrik Bohnenkamp, Pedro R. D'Argenio, Holger Hermanns, and Joost-Pieter Katoen. MoDeST: A compositional modeling formalism for real-time and stochastic systems. *IEEE Transactions on Software Engineering*, 32(10):812–830, 2006.

12 Sébastien Bornot and Joseph Sifakis. On the composition of hybrid systems. In *Proceedings of the 1st International Workshop on Hybrid Systems: Computation and Control (HSCC)*, volume 1386 of *Lecture Notes in Computer Science*, pages 49–63. Springer, 1998. URL: `http://dx.doi.org/10.1007/3-540-64358-3_31`, `doi:10.1007/3-540-64358-3_31`.

13 Patricia Bouyer, Thomas Brihaye, Pierre Carlier, and Quentin Menet. Compositional design of stochastic timed automata. In *Proceedings of the 11th International Computer Science Symposium in Russia (CSR)*, Lecture Notes in Computer Science, 2016 (to appear).

14 Mario Bravetti. Real time and stochastic time. In *Formal Methods for the Design of Real-Time Systems, International School on Formal Methods for the Design of Computer, Communication and Software Systems (SFM-RT)*, volume 3185 of *Lecture Notes in Computer Science*, pages 132–180. Springer, 2004. URL: `http://dx.doi.org/10.1007/978-3-540-30080-9_5`, `doi:10.1007/978-3-540-30080-9_5`.

15 Mario Bravetti and Pedro R. D'Argenio. Tutte le algebre insieme: Concepts, discussions and relations of stochastic process algebras with general distributions. In *Validation of Stochastic Systems - A Guide to Current Research*, volume 2925 of *Lecture Notes in Computer Science*, pages 44–88. Springer, 2004. `doi:10.1007/978-3-540-24611-4_2`.

16 Mario Bravetti and Roberto Gorrieri. The theory of interactive generalized semi-Markov processes. *Theoretical Computer Science*, 282(1):5–32, 2002. URL: `http://www.sciencedirect.com/science/article/pii/S0304397501000433`, `doi:http://dx.doi.org/10.1016/S0304-3975(01)00043-3`.

17 Manuela L. Bujorianu and John Lygeros. Toward a general theory of stochastic hybrid systems. In *Stochastic Hybrid Systems*, volume 337 of *Lecture Notes in Control and Information Science*, pages 3–30. Springer, 2006.

18 Stefano Cattani, Roberto Segala, Marta Z. Kwiatkowska, and Gethin Norman. Stochastic transition systems for continuous state spaces and non-determinism. In *Proceedings of the 8th International Conference on Foundations of Software Science and Computational Structures (FOSSACS)*, volume 3441 of *Lecture Notes in Computer Science*, pages 125–139. Springer, 2005.

19 Vincent Danos, Josée Desharnais, François Laviolette, and Prakash Panangaden. Bisimulation and cocongruence for probabilistic systems. *Information and Computation*, 204(4):503–523, 2006.

20 Pedro R. D'Argenio. *Algebras and Automata for Timed and Stochastic Systems*. PhD thesis, University of Twente, 1999.

21 Pedro R. D'Argenio and Joost-Pieter Katoen. A theory of stochastic systems part i: Stochastic automata. *Information and Computation*, 203(1):1–38, 2005.

22 Pedro R. D'Argenio and Joost-Pieter Katoen. A theory of stochastic systems part ii: Process algebra. *Information and Computation*, 203(1):39–74, 2005.

23 Pedro R. D'Argenio, Pedro Sánchez Terraf, and Nicolás Wolovick. Bisimulations for nondeterministic labelled Markov processes. *Mathematical Structures in Computer Science*, 22:43–68, 2012.

24 Erik P. de Vink and Jan J.M.M. Rutten. Bisimulation for probabilistic transition systems: a coalgebraic approach. *Theoretical Computer Science*, 221:271–293, 1999.

25 Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. *Information and Computation*, 179(2):163–193, 2002. URL: `http://dx.doi.org/10.1006/inco.2001.2962`, `doi:10.1006/inco.2001.2962`.

26 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximating labelled markov processes. *Information and Computation*, 184(1):160–200, 2003. URL: `http://www.sciencedirect.com/science/article/pii/S0890540103000518`, `doi:http://dx.doi.org/10.1016/S0890-5401(03)00051-8`.

27 Ernst-Erich Doberkat. Semi-pullbacks and bisimulations in categories of stochastic relations. pages 996–1007, 2003. URL: `http://dx.doi.org/10.1007/3-540-45061-0_77`, `doi:10.1007/3-540-45061-0_77`.

28 David Alan Edwards. On the existence of probability measures with given marginals. *Annales de l'institut Fourier*, 28(4):53–78, 1978. URL: `http://eudml.org/doc/74386`.

29 Christian Eisentraut, Holger Hermanns, and Lijun Zhang. On probabilistic automata in continuous time. In *Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 342–351. IEEE Computer Society, 2010. URL: `http://doi.org/10.1109/LICS.2010.41`, `doi:10.1109/LICS.2010.41`.

30 Martin Fränzle, Ernst Moritz Hahn, Holger Hermanns, Nicolás Wolovick, and Lijun Zhang. Measurability and safety verification for stochastic hybrid systems. In *Proceedings of the 14th International Conference on Hybrid Systems: Computation and Control (HSCC)*, pages 43–52. ACM, 2011. URL: `http://doi.acm.org/10.1145/1967701.1967710`, `doi:10.1145/1967701.1967710`.

31 Ernst Moritz Hahn. *Model checking stochastic hybrid systems*. PhD thesis, Universität des Saarlandes, 2013.

32 Ernst Moritz Hahn, Arnd Hartmanns, Holger Hermanns, and Joost-Pieter Katoen. A compositional modelling and analysis framework for stochastic hybrid systems. *Formal Methods in System Design*, 2012.

33 Arnd Hartmanns and Holger Hermanns. In the quantitative automata zoo. *Science of Computer Programming*, 112:3–23, 2015.

34 Thomas A. Henzinger. The theory of hybrid automata. In *Proceedings of the 11th Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 278–292. IEEE Computer Society, 1996. URL: `http://dx.doi.org/10.1109/LICS.1996.561342`, `doi:10.1109/LICS.1996.561342`.

35 Holger Hermanns. *Interactive Markov Chains: And the Quest for Quantified Quality*. Springer, 2002.

36 Holger Hermanns, Jan Krcál, and Jan Kretínský. Probabilistic bisimulation: Naturally on distributions. In *Proceedings of the 25th International Conference on Concurrency Theory (CONCUR)*, volume 8704 of *Lecture Notes in Computer Science*, pages 249–265. Springer, 2014. URL: `http://dx.doi.org/10.1007/978-3-662-44584-6_18`, `doi:10.1007/978-3-662-44584-6_18`.

37 Charles Antony Richard Hoare. *Communicating Sequential Processes*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1985.

38 Jianghai Hu, John Lygeros, and Shankar Sastry. Towards a theory of stochastic hybrid systems. In *Proceedings of the 12th International Conference on Hybrid Systems: Computation and Control (HSCC)*, volume 1790 of *Lecture Notes in Computer Science*, pages 160–173. Springer, 2000. URL: `http://dx.doi.org/10.1007/3-540-46430-1_16`, `doi:10.1007/3-540-46430-1_16`.

39 Alexander S. Kechris. *Classical Descriptive Set Theory*, volume 156 of *Graduate Texts in Mathematics*. Springer, 1995. `doi:10.1007/978-1-4612-4190-4`.

40 Hans G. Kellerer. Duality theorems for marginal problems. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(4):399–432, 1984. URL: `http://dx.doi.org/10.1007/BF00532047`, `doi:10.1007/BF00532047`.

41 Henning Kerstan and Barbara König. Coalgebraic trace semantics for probabilistic transition systems based on measure theory. In *Proceedings of the 23th International Conference on Concurrency Theory (CONCUR)*, volume 7454 of *Lecture Notes in Computer Science*, pages 410–424. Springer, 2012. URL: `http://dx.doi.org/10.1007/978-3-642-32940-1_29`.

42 László Lovász. Coupling measure concentrated on a given set. Unpublished notes on [43], 2012. URL: `http://www.cs.elte.hu/~lovasz/book/homnotes-A-3-4b.pdf`.

43 László Lovász. *Large Networks and Graph Limits*. American Mathematical Society colloquium publications. American Mathematical Society, 2012. URL: `https://books.google.de/books?id=FsFqHLid8sAC`.

44 Nancy Lynch, Roberto Segala, and Frits Vaandrager. Hybrid I/O automata. *Information and Computation*, 185(1):105–157, 2003.

45 Robin Milner. *A Calculus of Communicating Systems*. Springer, 1982.

46 Sayan Mitra and Nancy Lynch. Trace-based semantics for probabilistic timed I/O automata. In *Proceedings of the 12th International Conference on Hybrid Systems: Computation and Control (HSCC)*, volume 4416 of *Lecture Notes in Computer Science*, pages 718–722. Springer, 2007. URL: `http://dx.doi.org/10.1007/978-3-540-71493-4_69`, `doi:10.1007/978-3-540-71493-4_69`.

47 Prakash Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.

48 Klaus D. Schmidt. *Maß und Wahrscheinlichkeit*. Springer, 2008.

49 Roberto Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, Massachusetts Institute of Technology, 1995.

50 Heinz J. Skala. The existence of probability measures with given marginals. *The Annals of Probability*, 21(1):136–142, 1993. URL: `http://www.jstor.org/stable/2244753`.

51 Jeremy Sproston. *Model Checking for Probabilistic Timed and Hybrid Systems*. PhD thesis, University of Birmingham, 2001.

52 Shashi M. Srivastava. *A Course on Borel Sets*, volume 180 of *Graduate texts in mathematics*. Springer, 2008.

53 Volker Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, 1965. URL: `http://dx.doi.org/10.1214/aoms/1177700153`, `doi:10.1214/aoms/1177700153`.

54 Daniel H. Wagner. *Measure Theory Oberwolfach 1979: Proceedings of the Conference Held at Oberwolfach, Germany, July 1–7, 1979*, chapter Survey of measurable selection theorems: An update, pages 176–219. Springer, 1980. URL: `http://dx.doi.org/10.1007/BFb0088224`, `doi:10.1007/BFb0088224`.

55 Nicolás Wolovick. *Continuous Probability and Nondeterminism in Labeled Transition Systems*. PhD thesis, FaMAF - Universidad Nacional de Córdoba, 2012.

A Descriptive set theory

A.1 In the zoo of spaces

Polish spaces. A *Polish space* is a separable completely metrizable topological space. We endow a Polish space X with its Borel sigma-algebra, i.e., the sigma-algebra generated by the open (or equivalently closed) subsets of X. Moreover, we refer to a measurable subset of X as *Borel set*.

▶ **Remark 10** (Proposition 3.3 and Theorem 17.23 in [39]). The following statements hold:

1. If X is a Polish space and $A \subseteq X$ is an open or a closed subset of X, then A is a Polish space when equipped with the induced topology concerning X.
2. If X is a Polish space, then $\mathsf{Prob}(X)$ is a Polish space when equipped with the topology of weak convergence.
3. If $(X_n)_{n\in\mathbb{N}}$ is a sequence of Polish spaces, then $\prod_{n\in\mathbb{N}} X_n$ is a Polish space when equipped with the product product topology.

Standard Borel spaces. A measurable space $(X, \mathcal{M})$ is called a *standard Borel space* if there exists a Polish topology $\mathcal{O}$ on $\mathcal{X}$ such that $\mathsf{Borel}(\mathcal{O}) = \mathcal{M}$ (cf. Definition 12.5 in [39]). Here, we refer to a topology $\mathcal{O}$ on X where $(X, \mathcal{O})$ is a Polish space and $\mathsf{Borel}(\mathcal{O}) = \mathcal{M}$ as a *Polish topology for* X. The Polish topology for X is not uniquely determined, which turns to be beneficial, as the next remarks illustrate. As for Polish spaces we refer to measurable subsets of standard Borel spaces as *Borel sets*.

▶ **Remark 11** (Theorem 13.1 in [39]). Let $(X, \mathcal{O})$ be a Polish space and $M \subseteq X$ be a Borel set. Then, there is a Polish topology $\mathcal{O}_M$ on X such that $\mathcal{O} \subseteq \mathcal{O}_M$, $\mathsf{Borel}(\mathcal{O}) = \mathsf{Borel}(\mathcal{O}_M)$, and M is clopen in $\mathcal{O}_M$.

▶ **Remark 12** (Corollary 13.4 in [39]). Let X be a standard Borel space and $M \subseteq X$ be a Borel set. Then, M equipped with the induced sigma-algebra from X is also a standard Borel space. Notice, the claim is basically a corollary from Remarks 10 and 11

▶ **Remark 13** (Theorem 13.11 in [39]). Let $(X_1, \mathcal{O}_1)$ be a Polish space, $(X_2, \mathcal{O}_2)$ be a second-countable space, and $f : X_1 \to X_2$ be a Borel function. Then there is a Polish topology $\mathcal{O}_f$ on X_1 such that $\mathcal{O}_1 \subseteq \mathcal{O}_f$, $\mathsf{Borel}(\mathcal{O}_1) = \mathsf{Borel}(\mathcal{O}_f)$, and f is $(\mathcal{O}_f, \mathcal{O}_2)$-continuous. In particular, the claim holds if X_2 is a Polish space as every Polish space is second countable.

▶ **Remark 14** (Corollary 4.5 in [55]). Let X be a standard Borel space and $M \subseteq X$ be a Borel set. Then, $\{\mathsf{Dirac}[m] \,;\, m \in M\}$ is Borel in $\mathsf{Prob}(X)$.

Disintegration of measures.

▶ **Remark 15** (Disintegration, Exercise 17.35 in [39]). Let X_1 and X_2 be Borel spaces and $\mu \in \mathsf{Prob}(X)$. Define $\mu_1 \in \mathsf{Prob}(X_1)$ by $\mu_1 = f_\sharp(\mu)$ where $f_1 : X_1 \times X_2 \to X_1$, $f(x_1, x_2) = x_1$. Then, there exists a Markov kernel $k : X_1 \to \mathsf{Prob}(X_2)$ such that $\mu = \mu_1 \rtimes k$, i.e., for all Borel sets $M \subseteq X_1 \times X_2$, the function $g : X_1 \to [0, 1]$, $g(x_1) = k(x_1)([x_1]_{M,-})$ is Borel and

$$\mu(M) = \int g(x_1)\, d\mu_1(x_1).$$

Borel functions. A Borel function is a measurable function between standard Borel spaces.

▶ **Remark 16** (cf. Proposition 3.1.21 in [52]). Given standard Borel spaces X and Y and a Borel function $f\colon X \to Y$, then the graph of f is Borel in $X \times Y$.

▶ **Remark 17.** Let X_1 and X_2 be Polish spaces and let $f\colon X \to Y$ be a function. If f is Borel function, then $f_\sharp$ is a Borel function (cf. Theorem 17.24 in [39]). If f is continuous, then $f_\sharp$ is continuous. Let us give recall the argument for the second part and assume that f is continuous. Suppose a sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ in $\mathsf{Prob}(X)$ such that $(\mu_n)_{n\in\mathbb{N}}$ converges to $\mu \in \mathsf{Prob}(X)$. We justify that $(f_\sharp(\mu_n))_{n\in\mathbb{N}}$ converges to $f_\sharp(\mu)$. Let $g\colon Y \to [0,1]$ be a bounded and continuous function. Then their composition $g \circ f\colon X \to [0,1]$, $g \circ f(x) = g(f(x))$ is bounded and continuous and therefore

$$\begin{aligned}
& \lim_{n\to\infty} \int g(y)\, df_\sharp(\mu_n)(y) \\
= \quad & \lim_{n\to\infty} \int g \circ f(x)\, d\mu_n(x) \\
= \quad & \int g \circ f(x)\, d\mu(x) \\
= \quad & \int g(y)\, df_\sharp(\mu)(y).
\end{aligned}$$

Thus, $(f_\sharp(\mu_n))_{n\in\mathbb{N}}$ converges to $f_\sharp(\mu)$. We conclude that $f_\sharp$ is continuous.

Souslin sets. Given a Hausdorff space X, a subset M of X is called *a Souslin set in* X if there is a Polish space X_M and a continuous function $f\colon X_M \to X$ such that $f(X_M) = M$ (cf. Definition 6.6.1 in [10]). The empty set is agreed to be Souslin as well. A Souslin space is a Hausdorff space that is a Souslin set.

▶ **Remark 18** (Theorem 6.6.7 in [10]). Every Borel subset of a Souslin space is a Souslin set.

▶ **Remark 19** (Theorem 6.7.3 in [10]). Let X_1 and X_2 be Souslin spaces and $f\colon X_1 \to X_2$ be a Borel function. Then, $f(M_1)$ is Souslin in X_2 for all Souslin sets $M_1 \subseteq X_1$ and $f^{-1}(M_2)$ is Souslin in X_1 for all Souslin sets $M_2 \subseteq X_2$.

▶ **Remark 20.** There exists a Borel set $M \subseteq \mathbb{R} \times \mathbb{R}$ such that its projection on the first coordinate is not Borel in $\mathbb{R}$, i.e.,

$$M_1 = \{r_1 \in \mathbb{R} \,;\ \text{there is } r_2 \in \mathbb{R} \text{ where } \langle r_1, r_2\rangle \in M\}$$

is a set, which is not Borel in $\mathbb{R}$. Observe, M_1 is Souslin set in $\mathbb{R}$ using Remark 19: since one has $f_1(M) = M_1$ where $f_1\colon X_1 \times X_2 \to X_1$, $f_1(x_1, x_2) = x_1$.

▶ **Remark 21** (Corollary 1.5.8 in [9] and Theorem 7.4.1 in [10]). Let X be Hausdorff space, $\mu \in \mathsf{Prob}(X)$, and $M \subseteq X$ be Souslin. There exist Borel sets $M_l, M_u \subseteq X$ such that $M_l \subseteq M \subseteq M_u$ and $\mu(M_l) = \mu(M_u)$.

Compact sets. Let X be a topological space. We call a set $P \subseteq \mathsf{Prob}(X)$ *tight* if for every $\varepsilon \in \mathbb{R}_{>0}$ there exists a compact subset $K \subseteq X$ such that for all $\mu \in P$ it holds $\mu(K) > 1 - \varepsilon$ (cf. Section 5 in [8]).

▶ **Remark 22** (Prokhorov's theorem, Theorems 5.1 and 5.2 in [8]). Let X be a Polish space. A set P of probability measures on X is tight iff P is relatively compact.

▶ **Remark 23.** Let X be a metric space and $(x_n)_n$ be a sequence. Then, $(x_n)_n$ converges in X iff the set $\{x_n \,;\, n \in \mathbb{N}\}$ is compact in X. The claim follows from the fact that in metric spaces the notions of compactness and sequential compactness coincide.

Couplings. Remind the definition of couplings where X_1 and X_2 are measurable spaces: Given $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$, a *coupling of* (μ_1, μ_2) is a probability measure $\mu \in \mathsf{Prob}(X_1 \times X_2)$ such that for all measurable sets $M_1 \subseteq X_1$ and $M_2 \subseteq X_2$,

$$\mu(M_1 \times X_2) = \mu_1(M_1) \quad \text{and}$$
$$\mu(X_1 \times M_2) = \mu_2(M_2).$$

▶ **Proposition 24.** *Let X_1 and X_2 be measurable spaces, $\mu_1 \in \mathsf{Prob}(X_1)$, and $x_2 \in X_2$. There exists exactly one coupling of $(\mu_1, \mathsf{Dirac}[x_2])$, namely $\mu_1 \otimes \mathsf{Dirac}[x_2]$.*

Proof. Suppose μ is a coupling of $(\mu_1, \mathsf{Dirac}[x_2])$. For all measurable sets $M_1 \subseteq X_1$ and $M_2 \subseteq X_2$, if $x_2 \notin M_2$, then $\mu(M_1 \times M_2) = 0$ since

$$\mu(M_1 \times M_2) \le \mu(X_a \times M_2) = \mathsf{Dirac}[x_2](M_2) = 0.$$

Given measurable sets $M_1 \subseteq X_1$ and $M_2 \subseteq X_2$ where $x_2 \in M_2$, we obtain

$$\mu(M_1 \times M_2) = \mu(M_1 \times X_2) - \mu(M_1 \times (X_2 \setminus M_2)) = \mu(M_1 \times X_2) = \mu_1(M_1).$$

Putting things together we have $\mu(M_1 \times M_2) = \mu_1(M_1) \cdot \mathsf{Dirac}[x_2](M_2)$ for all measurable sets $M_1 \subseteq X_1$ and $M_2 \subseteq X_2$. From this the claim follows. ◀

A.2 Relations

In this section we study relations between measurable spaces and discuss important properties for the remainder of this work.

Sections of relations. Given sets X_1 and X_2 and a relation $R \subseteq X_1 \times X_2$, for every $x_1 \in X_1$ and $x_2 \in X_2$ define

$$[x_1]_{R,-} = \{x_2' \in X_2 \,;\, x_1 \, R \, x_2'\} \quad \text{and} \quad [x_2]_{-,R} = \{x_1' \in X_1 \,;\, x_1' \, R \, x_2\}.$$

▶ **Remark 25** (Lemma 9.6.1 in [48])**.** Let X_1 and X_2 be measurable spaces and $R \subseteq X_1 \times X_2$ be measurable. For all $x_1 \in X_1$ and $x_2 \in X_2$ the sets $[x_1]_{R,-}$ and $[x_2]_{-,R}$ are measurable in X_1 and X_2, respectively.

▶ **Proposition 26.** *Let X_1 and X_2 be metric spaces and $R \subseteq X_1 \times X_2$ be closed. Then, $[x_1]_{R,-}$ and $[x_2]_{-,R}$ are closed in X_2 and X_1, respectively. Moreover, if X_1 and X_2 are sigma-compact, then $[x_1]_{R,-}$ and $[x_2]_{-,R}$ are sigma-compact in X_2 and X_1, respectively.*

Proof. For reasons of symmetry it suffices to consider the claims concerning the set $[x_1]_{R,-}$ for all $x_1 \in X_1$. Let $x_1 \in X_1$. Given the assumption that R is closed in $X_1 \times X_2$, it is easy to see that $[x_1]_{R,-}$ is closed in X_2. Let us recall the argument and suppose $(x_{2,n})_n$ is a sequence in $[x_1]_{R,-}$ that converges in X_2. Denote the limit by $x_2 \in X_2$. Then, $(\langle x_1, x_{2,n}\rangle)_n$ is a sequence in R which converges in $X_1 \times X_2$ and has limit $\langle x_1, x_2\rangle$. Since R is closed in $X_1 \times X_2$, it follows $x_1 \, R \, x_2$ and hence $x_2 \in [x_1]_{R,-}$. Thus, $[x_1]_{R,-}$ is closed in X_2.

Assume X_2 is sigma-compact and let $(K_n)_n$ be a sequence of compact sets $K_n \subseteq X_2$, $n \in \mathbb{N}$, such that $X_2 = \bigcup_{n\in\mathbb{N}} K_n$. Since $[x_1]_{R,-}$ is closed in X_2, the set $[x_1]_{R,-} \cap K_n$ is compact in X_2 for all $n \in \mathbb{N}$. Moreover,

$$[x_1]_{R,-} = \bigcup_{n\in\mathbb{N}} ([x_1]_{R,-} \cap K_n)$$

and hence, $[x_1]_{R,-}$ is sigma-compact in X_2. ◀

Stable pairs. Suppose sets X_1 and X_2 and a relation $R \subseteq X_1 \times X_2$. Given sets $A_1 \subseteq X_1$ and $A_2 \subseteq X_2$ we refer to the pair $\langle A_1, A_2 \rangle$ as R-stable if $R \cap (A_1 \times X_2) = R \cap (X_1 \times A_2)$.

Quasi-equivalence relations. In what follows we introduce quasi-equivalence relations, that generalizes the notion of equivalence relations. Let X_1 and X_2 be sets and $R \subseteq X_1 \times X_2$ be a relation. We call R *lr-total (in $X_1 \times X_2$)* if $[x_1]_{R,-}$ and $[x_2]_{-,R}$ are non-empty for all $x_1 \in X_1$ and $x_2 \in X_2$. We say that R is *z-transitive* if for all $x_1 \, R \, x_2$, $x_1 \, R \, x_2'$, and $x_1' \, R \, x_2$ it holds $x_1' \, R \, x_2'$. We call R a *quasi-equivalence (in $X_1 \times X_2$)* if R is lr-total in $X_1 \times X_2$ and z-transitive. The next proposition connects the notion of quasi-equivalences and equivalences.

▶ **Proposition 27.** *Let X be a set. Every equivalence in X is a quasi-equivalence in $X \times X$. Vice verca, a quasi-equivalence in $X \times X$, that is reflexive in X, is a equivalence in X.*

Proof. Let $R \subseteq X \times X$ be a relation. Assume R is an equivalence first. Using the reflexivity of R it follows that R is lr-total. To see that R is z-transitive assume $x_1 \, R \, x_2$, $x_1 \, R \, x_2'$, and $x_1' \, R \, x_2$. Since R is symmetric it holds $x_1' \, R \, x_2$, $x_2 \, R \, x_1$, and $x_1 \, R \, x_2'$. Applying the transitivity of R we obtain $x_1' \, R \, x_2'$ and hence R is z-transitive. This proves the first part of the proposition.

Suppose R is an quasi-equivalence that is reflexive. Let $x_1 \, R \, x_2$. Since $x_1 \, R \, x_1$ and $x_2 \, R \, x_2$, the z-transitivity of R yields $x_2 \, R \, x_1$, which shows that R is symmetric. It remains to show that R is transitive. For that assume $x_1 \, R \, x_2$ and $x_2 \, R \, x_2'$. Since $x_2 \, R \, x_2$, the z-transitivity of R again implies $x_1 \, R \, x_2'$. ◀

▶ **Proposition 28.** *Let X_1 and X_2 be sets and $R \subseteq X_1 \times X_2$ be z-transitive. Then,*

$$R = \bigcup_{x_1 \, R \, x_2} [x_2]_{-,R} \times [x_1]_{R,-}.$$

Moreover, for all $x_1 \, R \, x_2$ and $x_1' \, R \, x_2'$,

$$\begin{aligned} & [x_2]_{-,R} = [x_2']_{-,R} \\ \textit{iff} \quad & [x_2]_{-,R} \cap [x_2']_{-,R} \neq \varnothing \\ \textit{iff} \quad & [x_1]_{R,-} \cap [x_1']_{R,-} \neq \varnothing \\ \textit{iff} \quad & [x_1]_{R,-} = [x_1']_{R,-}. \end{aligned}$$

If in addition R is lr-total in $X_1 \times X_2$, then for every R-stable pair $\langle M_1, M_2 \rangle$,

$$M_1 = \bigcup_{x_2 \in M_2} [x_2]_{-,R} \quad \textit{and} \quad M_2 = \bigcup_{x_1 \in M_1} [x_1]_{R,-}.$$

Proof. The arguments are straightforward applying the definitions. ◀

We introduce a stronger variant of lr-totality in $X_1 \times X_2$ next. Suppose X_1 and X_2 are measurable spaces and $R \subseteq X_1 \times X_2$. We call R *strongly lr-total (in $X_1 \times X_2$)* if there are measurable functions $f_1 \colon X_1 \to X_2$ and $f_2 \colon X_2 \to X_1$ such that $\mathsf{graph}(f_1) \subseteq R$ and $\mathsf{graph}(f_2)^{-1} \subseteq R$ i.e., for all $x_1 \in X_1$ and $x_2 \in X_2$,

$$x_1 \, R \, f_1(x_1) \quad \text{and} \quad f_2(x_2) \, R \, x_2.$$

Moreover, R is a *strong quasi-equivalence (in $X_1 \times X_2$)* if R is strongly lr-total in $X_1 \times X_2$ and z-transitive. Obviously, every strong quasi-equivalence in $X_1 \times X_2$ is a quasi-equivalence in $X_1 \times X_2$. Reverse direction later ...

Measurable selection theorems. Let X_2 and X_1 be measurable spaces and $f\colon X_2 \to X_1$ be some surjective function. To obtain an right inverse g of f, i.e., a function $g\colon X_1 \to X_2$ where $f(g(x_1)) = x_1$ for all $x_1 \in X_1$, one can proceed as follows: For every $x_1 \in X_1$ pick some element $x_2 \in X_2$ where $f(x_2) = x_1$ and define $g(x_1) = x_2$. However, one often needs a function g which is measurable as well. The latter aspect illustrates one motivation for the research on so-called measurable selection theorems in descriptive set theory [54, 10]. To describe the (more general) setting shortly one starts with a relation $R \subseteq X_1 \times X_2$ and seeks for sufficient conditions that ensure the existence of a *measurable selection of* R, i.e., a measurable function $g\colon X_1 \to X_2$ where $\mathsf{graph}(g) \subseteq R$. For the sketched application concerning a right inverse of f we can then simply instantiate $R = \mathsf{graph}(f)^{-1}$ since here every measurable selection of R constitutes a right inverse of f.

▶ **Remark 29** (Theorem 3.5 in [54] or Theorem 6.9.6 in [10])**.** Let X_1 and X_2 be Polish spaces $R \subseteq X_1 \times X_2$ be a Borel set. Assume $[x_1]_{R,-}$ is non-empty and sigma-compact in X_2 for all $x_1 \in X_1$. Then there exists a Borel function $f\colon X_1 \to X_2$ such that $\mathsf{graph}(f) \subseteq R$.

▶ **Proposition 30.** *Let X_1 and X_2 be sigma-compact Polish spaces and $R \subseteq X_1 \times X_2$ be a closed set. Then, R is lr-total iff R is strongly lr-total. In particular, R is an quasi-equivalence iff R is a strong quasi-equivalence.*

Proof. Assume R is lr-total. Our task is to show that R is strongly lr-total. We first observe that R is Borel in $X_1 \times X_2$. Let $x_1 \in X_1$ and $x_2 \in X_2$. Since R is lr-total the sets $[x_1]_{R,-}$ and $[x_2]_{-,R}$ are non-empty. As R is closed in $X_1 \times X_2$, the sets $[x_1]_{R,-}$ and $[x_2]_{-,R}$ are closed and thus sigma-compact in X_2 and X_1, respectively (cf. Proposition 26). We are in the situation of Remark 29 that finally yields the claim. ◀

Diagonal relation. We introduce a special equivalence relation on a set X next: The *diagonal relation on* X is given by

$$\mathsf{Diag}_X = \{\langle x_1, x_2\rangle \in X \times X \; ; \; x_1 = x_2\}.$$

▶ **Remark 31.** Given a topological space X, it is well-known that X is Hausdorff iff the diagonal relation on X is closed in $X \times X$. It follows that for every standard Borel space X the diagonal relation on X is Borel in $X \times X$. The argument is straightforward: Suppose $\mathcal{O}$ is a Polish topology for X, then Diag_X is closed in $\mathcal{O} \otimes \mathcal{O}$ and hence Borel in $X \times X$.

Countably separated relations. When we discuss weight functions in Section A.3 countably separated relations will become of crucial interest. Let X_1 and X_2 be measurable spaces and $R \subseteq X_1 \times X_2$. We say that R is *countably separated* if there exists a standard Borel space $\underline{X}$ and measurable functions $\kappa_1\colon X_1 \to \underline{X}$ and $\kappa_2\colon X_2 \to \underline{X}$ such that

$$R = \{\langle x_1, x_2\rangle \in X_1 \times X_2 \; ; \; \kappa_1(x_1) = \kappa_2(x_2)\}.$$

Here, we then say that $(\underline{X}, \kappa_1, \kappa_2)$ *countably separates* R. Notice that if R is countably separated, then R is an quasi-equivalence in $X_1 \times X_2$. The idea behind the notion of countably separated relations is to distinguish the sets $[x_2]_{-,R} \times [x_1]_{R,-}$ by assigning to them an element of $\underline{X}$ where $x_1 \, R \, x_2$. Indeend, for all $x_1 \, R \, x_2$ and $\langle x_1', x_2'\rangle \in [x_2]_{-,R} \times [x_1]_{R,-}$ we have $\kappa_1(x_1') = \kappa_2(x_2')$. This idea is not new for the case where $X_1 = X_2$ (cf. Exercise 5.1.10 in [52]): Here, one defines that $R \subseteq X_1 \times X_1$ is countably separated if there are a standard Borel space $\underline{X}$ and a measurable function $\kappa\colon X_1 \to \underline{X}$ such that $R = \{\langle x_1, x_1'\rangle \in X_1 \times X_1 \; ; \; \kappa(x_1) = \kappa(x_1')\}$ (and thus R is an equivalence in X_1). Notice, in this case $X_1 = X_2$ the requirement in our definition is weaker.

We know regard the question when a given quasi-equivalence relation is countably separated. Here, we rely on results presented in [52] concerning equivalence relations. The idea is to transform an quasi-equivalence relation in an appropriate equivalence relation and then rely on results proven in [52].

▶ **Theorem 32.** *Let X_1 and X_2 be Polish spaces and $R \subseteq X_1 \times X_2$ be an strong quasi-equivalence, which is closed in $X_1 \times X_2$. Then, R is countably separated.*

Proof. Define $R' \subseteq R \times R$ by

$$R' = \{\langle\langle x_1, x_2\rangle, \langle x_1', x_2'\rangle\rangle \in R \times R \,;\, x_1 \, R \, x_2'\}.$$

We argue that R' is an equivalence in R. Reflexivity is clear. Symmetry can be seen as follows. Let $\langle x_1, x_2\rangle \, R' \, \langle x_1', x_2'\rangle$. Thus, $x_1 \, R \, x_2$, $x_1' \, R \, x_2'$, and $x_1 \, R \, x_2'$. Since R is z-transitive, we obtain $x_1' \, R \, x_2$ and thus $\langle x_1', x_2'\rangle \, R' \, \langle x_1, x_2\rangle$. It remains to show transitivity. Suppose $\langle x_1, x_2\rangle \, R' \, \langle x_1', x_2'\rangle$ and $\langle x_1', x_2'\rangle \, R' \, \langle x_1'', x_2''\rangle$. As R' is symmetric, $\langle x_1'', x_2''\rangle \, R \, \langle x_1', x_2'\rangle$. We have $x_1 \, R \, x_2'$, $x_1'' \, R \, x_2'$, and $x_1'' \, R \, x_2''$. The z-transitivity of R therefore implies $x_1 \, R \, x_2''$ and hence $\langle x_1, x_2\rangle \, R \, \langle x_1'', x_2''\rangle$.

Since R is closed in $X_1 \times X_2$, the set R equipped with the topology induced by $X_1 \times X_2$ constitutes a Polish space (cf. Remark 10). Furthermore, the set R' is closed in $R \times R$, which can be seen as follows: Define the continuous function $f \colon R \times R \to X_1 \times X_2$,

$$f(\langle x_1, x_2\rangle, \langle x_1', x_2'\rangle) = \langle x_1, x_2'\rangle.$$

Since R is closed in $X_1 \times X_2$, the set $R' = f^{-1}(R)$ is closed in $R \times R$.

We are in the situation of Proposition 5.1.11 in [52] and thus there exist a standard Borel space $\underline{X}$ and a Borel function $\kappa' \colon R' \to \underline{X}$ such that

$$R' = \{\langle x, x'\rangle \in R \times R \,;\, \kappa'(x) = \kappa'(x')\}.$$

We exploit that R is strongly lr-total in $X_1 \times X_2$ now. Let $f_1 \colon X_1 \to X_2$ and $f_2 \colon X_2 \to X_2$ be a Borel function such that $x_1 \, R \, f_1(x_1)$ for all $x_1 \in X_1$ and $f_2(x_2) \, R \, x_2$ for all $x_2 \in X_2$. Define the Borel functions $\kappa_1 \colon X_1 \to \underline{X}$ and $\kappa_2 \colon X_2 \to \underline{X}$ by

$$\kappa_1(x_1) = \kappa(x_1, f_1(x_1)) \quad \text{and}$$
$$\kappa_2(x_2) = \kappa(x_2, f_2(x_2))$$

for all $x_1 \colon X_1$ and $x_2 \in X_2$. We claim that $(\underline{X}, \kappa_1, \kappa_2)$ countably separates R.

We observe,

$$\kappa(x_1, x_2) = \kappa(x_1, x_2') \text{ for all } x_1 \, R \, x_2 \text{ and } x_1 \, R \, x_2' \quad \text{and}$$
$$\kappa(x_1, x_2) = \kappa(x_1', x_2) \text{ for all } x_1 \, R \, x_2 \text{ and } x_1' \, R \, x_2.$$

Indeed, if $x_1 \, R \, x_2$ and $x_1 \, R \, x_2'$, then $\langle x_1, x_2\rangle \, R' \, \langle x_1, x_2'\rangle$ and thus $\kappa(x_1, x_2) = \kappa(x_1, x_2')$. Analogously, given $x_1 \, R \, x_2$ and $x_1' \, R \, x_2$, then $\langle x_1, x_2\rangle \, R' \, \langle x_1', x_2\rangle$ and therefore $\kappa(x_1, x_2) = \kappa(x_1', x_2)$.

We finally show $R = \{\langle x_1, x_2\rangle \in X_1 \times X_2 \,;\, \kappa_1(x_1) = \kappa_2(x_2)\}$. Let $x_1 \in X_1$ and $x_2 \in X_2$. Assuming $x_1 \, R \, x_2$, it follows

$$\kappa_1(x_1) = \kappa(x_1, f_1(x_1)) = \kappa(x_1, x_2) = \kappa(f_2(x_2), x_2) = \kappa_2(x_2).$$

If $\kappa_1(x_1) = \kappa_2(x_2)$, then

$$\kappa(x_1, f_1(x_1)) = \kappa_1(x_1) = \kappa_2(x_2) = \kappa(f_2(x_2), x_2)$$

and with that $\langle x_1, f_1(x_1)\rangle \, R' \, \langle f_2(x_2), x_2\rangle$, which yields $x_1 \, R \, x_2$. ◀

▶ **Corollary 33.** *Let X_1 and X_2 be sigma-compact Polish spaces and $R \subseteq X_1 \times X_2$ be an quasi-equivalence, which is closed in $X_1 \times X_2$. Then, R is countably separated.*

Proof. Proposition 30 and Theorem 32 yield the claim. ◀

A.3 Weight functions

In the context of bisimulation for probabilistic systems an important question is how to lift a relation $R \subseteq X_1 \times X_2$ to probability measures, i.e., to a relation $R' \subseteq \mathsf{Prob}(X_1) \times \mathsf{Prob}(X_2)$. For a conservative notion of a lifting it is for instance desirable to have $\mathsf{Dirac}[x_1]\, R'\, \mathsf{Dirac}[x_2]$ for all $x_1\, R\, x_2$. We recall the weight lifting of relations next, discuss related literature, and present results that connect the weight lifting with ths so-called stable-pair lifting.

Let X_1 and X_2 be measurable spaces. Assuming a relation $R \subseteq X_1 \times X_2$, a *weight function for* (μ_1, R, μ_2) is a coupling W of (μ_1, μ_2) such that $x_1\, R\, x_2$ for W-almost all $\langle x_1, x_2 \rangle \in X_1 \times X_2$, i.e., there exists a measurable set R' in $X_1 \times X_2$ where $R' \subseteq R$ and $W(R') = 1$. Notice, R is not required to be measurable. The *weight lifting of* R is given by $R^w \subseteq \mathsf{Prob}(X_1) \times \mathsf{Prob}(X_2)$ such that for all $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$,

$$\mu_1\, R^w\, \mu_2 \quad \text{iff} \quad \text{there is a weight function for } (\mu_1, R, \mu_2).$$

The probability-theory community has investigated sufficient and necessary conditions for the existence of a weight function for (μ_1, R, μ_2), e.g., [53, 28, 40, 50]. The most distinguishing fact of these articles is the assumption concerning the involved spaces X_1 and X_2. For instance, [53] considers Polish spaces, [28] investigates compact Hausdorff spaces, and [50] generalizes results to Hausdorff spaces. Weight functions have been attended in the verification community as well, e.g., [49] (cf. Section 8.2) or [7] where X_1 and X_2 are supposed to be finite. In [7] one characterizes the existence of weight functions in terms of maximum flows in networks (cf. Lemma 5.1) which enables an algorithm to decide $\mu_1\, R^w\, \mu_2$ where μ_1, μ_2, and R are the inputs of the decision problem (cf. Lemma 5.2). The idea concerning maximum flows in networks is also applied in [43, 42] (cf. the proof of Proposition A.7).

▶ **Proposition 34.** *Let X_1 and X_2 be measurable spaces, $R \subseteq X_1 \times X_2$ be a relation, $x_1 \in X_1$, and $x_2 \in X_2$. Then, $\mathsf{Dirac}[x_1]\, R^w\, \mathsf{Dirac}[x_2]$ implies $x_1\, R\, x_2$. Moreover, if there is a measurable set $R' \subseteq X_1 \times X_2$ where $\langle x_1, x_2 \rangle \in R' \subseteq R$, then $\mathsf{Dirac}[x_1]\, R^w\, \mathsf{Dirac}[x_2]$. In particular, if X_1 and X_2 are standard Borel spaces, then*

$$x_1\, R\, x_2 \quad \textit{iff} \quad \mathsf{Dirac}[x_1]\, R^w\, \mathsf{Dirac}[x_2].$$

Proof. Assume W is a weight function for $(\mathsf{Dirac}[x_1], R, \mathsf{Dirac}[x_2])$ and $R' \subseteq X_1 \times X_2$ is a measurable set such that $R' \subseteq R$ and $W(R') = 1$. According to Proposition 24 we thus have $W = \mathsf{Dirac}[x_1] \otimes \mathsf{Dirac}[x_2]$ and hence $\langle x_1, x_2 \rangle \in R' \subseteq R$, that shows the first part of the proposition.

Given a measurable set $R' \subseteq X_1 \times X_2$ where $\langle x_1, x_2 \rangle \in R' \subseteq R$, then it easy to see that $\mathsf{Dirac}[x_1] \otimes \mathsf{Dirac}[x_2]$ is a weight function for $(\mathsf{Dirac}[x_1], R, \mathsf{Dirac}[x_2])$ and hence $\mathsf{Dirac}[x_1]\, R^w\, \mathsf{Dirac}[x_2]$.

If a topological space satisfies the separation axiom T_1, then the singleton subsets are closed and hence Borel. Thus, whenever X_1 and X_2 are standard Borel spaces, then $\{\langle x_1, x_2 \rangle\}$ is Borel in $X_1 \times X_2$, which yields the remaining claim. ◀

Considering the literature concerning LMPs and NLMPs [24, 25, 19, 47, 23, 55, 6] there is another approach of lifting relations to probability measures summarized next. Let X_1

and X_2 be measurable spaces and $R \subseteq X_1 \times X_2$. The *stable-pair lifting of* R is the relation $R^s \subseteq \mathsf{Prob}(X_1) \times \mathsf{Prob}(X_2)$ such that for all $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$,

$$\mu_1 \, R^s \, \mu_2 \quad \text{iff} \quad \mu_1(M_1) = \mu_2(M_2)$$
$$\text{for all measurable } M_1 \subseteq X_1 \text{ and } M_2 \subseteq X_2 \text{where } \langle M_1, M_2 \rangle \textit{ is } R\text{-stable.}$$

Remind, $\langle M_1, M_2 \rangle$ is R-stable pair if $R \cap (M_1 \times X_2) = R \cap (X_1 \times M_2)$.

▶ **Proposition 35.** *Let X_1 and X_2 be measurable spaces and $R \subseteq X_1 \times X_2$. Then, $R^w \subseteq R^s$.*

Proof. Let $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$ be such that $\mu_1 \, R^w \, \mu_2$. Assume W is a weight function for (μ_1, R, μ_2) and $R' \subseteq X_1 \times X_2$ is a measurable set where $R' \subseteq R$ and $W(R') = 1$. Suppose measurable sets $M_1 \subseteq X_1$ and $M_2 \subseteq X_2$ where $\langle M_1, M_2 \rangle$ is R-stable, i.e., $R \cap (M_1 \times X_2) = R \cap (X_1 \times M_2)$, also $R' \cap (M_1 \times X_2) = R' \cap (X_1 \times M_2)$. Since $W(R') = 1$,

$$\begin{aligned} & \mu_1(M_1) \\ = \; & W(M_1 \times X_2) \\ = \; & W(R' \cap (M_1 \times X_2)) \\ = \; & W(R' \cap (X_1 \times M_2)) \\ = \; & W(X_1 \times M_2) \\ = \; & \mu_2(M_2). \end{aligned}$$

From this we obtain $\mu_1 \, R^s \, \mu_2$, which finishes our proof. ◀

▶ **Example 36.** Let $X = \{0, 1\}$ and consider the relation $R \subseteq X \times X$ defined by $R = \{\langle 0, 0 \rangle, \langle 1, 0 \rangle, \langle 1, 1 \rangle\}$. Since $\langle 0, 1 \rangle \notin R$ we obtain $\langle \mathsf{Dirac}[0], \mathsf{Dirac}[1] \rangle \notin R^w$ by Proposition 34. However, it is easy to see that $\langle \mathsf{Dirac}[0], \mathsf{Dirac}[1] \rangle \in R^s$ since $\langle \varnothing, \varnothing \rangle$ and $\langle X, X \rangle$ are the only R-stable pairs. It follows $R^s \not\subseteq R^w$. We observe that R is no quasi-equivalence.

We now discuss condtions which ensure that the two notions of liftings coincide, i.e., $R^w = R^s$. Having Proposition 35 in mind, we aim for conditions where $R^s \subseteq R^w$. Example 36 teaches us that it is appropriate to restrict to quasi-equivalences. In the discrete setting there is the following result (cf. Lemma 8.2.2 in [49]): If R is an quasi-equivalence relation between countable sets, then $R^w = R^s$. The proof idea can be applied for quasi-equivalence relations R between arbitrary measurable spaces if R admits a countable Borel decomposition [24].

We aim to show that if R is countably separated, then $R^s = R^w$. For this purpose we present the following two lemmas first.

▶ **Lemma 37.** *Let X be a standard Borel space. For all $\mu_1, \mu_2 \in \mathsf{Prob}(X)$,*

$$\mu_1 \, \mathsf{Diag}_X^w \, \mu_2 \quad \textit{iff} \quad \mu_1 \, \mathsf{Diag}_X^s \, \mu_2 \quad \textit{iff} \quad \mu_1 = \mu_2.$$

Proof. Remind, Diag_X is Borel in $X \times X$ by Remark 31. It turns out that $\mu \, \mathsf{Diag}_X^w \, \mu$ for all $\mu \in \mathsf{Prob}(X)$. Let us see why and define the Borel function $f\colon X \to X \times X$, $f(x) = \langle x, x \rangle$. Then,

$$f_\sharp(\mu)(\mathsf{Diag}_X) = \mu(f^{-1}(\mathsf{Diag}_X)) = \mu(X) = 1$$

and for all Borel sets $M \subseteq X$ it holds $f^{-1}(M \times X) = M$ and $f^{-1}(X \times M) = M$ and thus,

$$f_\sharp(\mu)(M \times X) = \mu(f^{-1}(M \times X)) = \mu(M) \quad \text{and}$$
$$f_\sharp(\mu)(X \times M) = \mu(f^{-1}(X \times M)) = \mu(M).$$

Therefore, $f_\sharp(\mu)$ is a weight function for $(\mu, \mathsf{Diag}_X, \mu)$ and thus $\mu_1 \, \mathsf{Diag}_X^w \, \mu_2$.

Let $\mu_1, \mu_2 \in \mathsf{Prob}(X)$ be such that $\mu_1 \, \mathsf{Diag}_X^w \, \mu_2$. Proposition 35 yields $\mu_1 \, \mathsf{Diag}_X^s \, \mu_2$. For every Borel set $M \subseteq X$ the pair $\langle M, M \rangle$ is Diag_X-stable and hence $\mu_1(M) = \mu_2(M)$. It follows $\mu_1 = \mu_2$, which yields the claimed equivalence. ◀

▶ **Lemma 38.** *Let X_1, X_2, $\tilde{X}_1$, and $\tilde{X}_2$ be standard Borel spaces, $R \subseteq X_1 \times X_2$, and $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$. Moreover, assume Borel functions $f_1 \colon X_1 \to \tilde{X}_1$ and $f_2 \colon X_2 \to \tilde{X}_2$. Define*

$$\tilde{\mu}_1 = (f_1)_\sharp(\mu_1), \quad \tilde{\mu}_2 = (f_2)_\sharp(\mu_2), \quad \textit{and} \quad \tilde{R} = \{\langle f_1(x_1), f_2(x_2) \rangle \; ; \; x_1 \, R \, x_2\}.$$

The following statements hold:

1. $\mu_1 \, R^w \, \mu_2$ *implies* $\tilde{\mu}_1 \, \tilde{R}^w \, \tilde{\mu}_2$.
2. $\mu_1 \, R^s \, \mu_2$ *implies* $\tilde{\mu}_1 \, \tilde{R}^s \, \tilde{\mu}_2$.

Proof. *Ad* (1). Let $\tilde{\mathcal{O}}_1$ and $\tilde{\mathcal{O}}_2$ be Polish topologies for $\tilde{X}_1$ and $\tilde{X}_2$, respectively. According to Remark 13 there are Polish topologies $\mathcal{O}_1$ and $\mathcal{O}_2$ on X_1 and X_2, respectively, such that f_1 is $(\mathcal{O}_1, \tilde{\mathcal{O}}_1)$-continuous and f_2 is $(\mathcal{O}_2, \tilde{\mathcal{O}}_2)$-continuous. In what follows we suppose X_1 and X_2 are equipped with the topologies $\mathcal{O}_1$ and $\mathcal{O}_2$, respectively. Define the continuous function $f \colon X_1 \times X_2 \to \tilde{X}_1 \times \tilde{X}_2$,

$$f(x_1, x_2) = \langle f_1(x_1), f_2(x_2) \rangle.$$

Notice, $\tilde{R} = f(R)$. Assume $\mu_1 \, R^w \, \mu_2$ and let W be a weight function for (μ_1, R, μ_2). Define

$$\tilde{W} = f_\sharp(W).$$

We claim that $\tilde{W}$ is a weight function for $(\tilde{\mu}_1, \tilde{R}, \tilde{\mu}_2)$. For every Borel set $\tilde{M}_1 \subseteq \tilde{X}_1$ it holds $f_1^{-1}(\tilde{M}_1) \times X_2 = f^{-1}(\tilde{M}_1 \times \tilde{X}_2)$ and therefore,

$$\tilde{\mu}_1(\tilde{M}_1) = \mu_1(f_1^{-1}(\tilde{M}_1)) = W(f_1^{-1}(\tilde{M}_1) \times X_2) = W(f^{-1}(\tilde{M}_1 \times \tilde{X}_2)) = \tilde{W}(\tilde{M}_1 \times \tilde{X}_2).$$

Analogously, $\tilde{\mu}_2(\tilde{M}_2) = \tilde{W}(\tilde{X}_1 \times \tilde{M}_2)$ for all Borel sets $\tilde{M}_2 \subseteq \tilde{X}_2$. Therefore, $\tilde{W}$ is a coupling of $(\tilde{\mu}_1, \tilde{\mu}_2)$.

It remains to show that $\tilde{x}_1 \, \tilde{R} \, \tilde{x}_2$ for $\tilde{W}$-almost all $\langle \tilde{x}_1, \tilde{x}_2 \rangle \in \tilde{X}_1 \times \tilde{X}_2$. Let $R' \subseteq X_1 \times X_2$ be a Borel set such that $R' \subseteq R$ and $W(R') = 1$. Then, $f(R')$ is a Souslin set in $\tilde{X}_1 \times \tilde{X}_2$. According to Remark 21 there exist Borel sets $\tilde{R}_l, \tilde{R}_u \subseteq \tilde{X}_1 \times \tilde{X}_2$ such that $\tilde{R}_l \subseteq f(R') \subseteq \tilde{R}_u$ and $\tilde{W}(\tilde{R}_l) = \tilde{W}(\tilde{R}_u)$. Using $R' \subseteq f^{-1}(f(R')) \subseteq f^{-1}(\tilde{R}_u)$ we obtain

$$\tilde{W}(\tilde{R}_l) = \tilde{W}(\tilde{R}_u) = W(f^{-1}(\tilde{R}_u)) \geq W(R') = 1$$

Since $\tilde{R}_l \subseteq f(R') \subseteq f(R) = \tilde{R}$ it follows that $\tilde{W}$ is indeed a weight function for $(\tilde{\mu}_1, \tilde{R}, \tilde{\mu}_2)$.

Ad (2). We assume $\mu_1 \, R^s \, \mu_2$. Let $\tilde{M}_1 \subseteq \tilde{X}_1$ and $\tilde{M}_2 \subseteq \tilde{X}_2$ be measurable sets such that $\langle \tilde{M}_1, \tilde{M}_2 \rangle$ is a $\tilde{R}$-stable pair. We argue that $\langle f_1^{-1}(\tilde{M}_1), f_2^{-1}(\tilde{M}_2) \rangle$ is R-stable first. Let $x_1 \, R \, x_2$ where $x_1 \in f_1^{-1}(\tilde{M}_1)$. Then, $f_1(x_1) \, \tilde{R} \, f_2(x_2)$ and $f_1(x_1) \in \tilde{M}_1$. Since $\langle \tilde{M}_1, \tilde{M}_2 \rangle$ is $\tilde{R}$-stable, it thus follows $f_2(x_2) \in \tilde{M}_2$ and so $x_2 \in f_2^{-1}(\tilde{M}_2)$. Therefore, $R \cap (f_1^{-1}(\tilde{M}_1) \times X_2) \subseteq R \cap (X_1 \times f_2^{-1}(\tilde{M}_2))$. One analogously shows reverse inclusion $R \cap (f_1^{-1}(\tilde{M}_1) \times X_2) \supseteq R \cap (X_1 \times f_2^{-1}(\tilde{M}_2))$. We obtain that $\langle f_1^{-1}(\tilde{M}_1), f_2^{-1}(\tilde{M}_2) \rangle$ is R-stable. Using $\mu_1 \, R^s \, \mu_2$,

$$\tilde{\mu}_1(\tilde{M}_1) = \mu_1(f_1^{-1}(\tilde{M}_1)) = \mu_2(f_2^{-1}(\tilde{M}_2)) = \tilde{\mu}_2(\tilde{M}_2).$$

This finally implies $\tilde{\mu}_1 \, \tilde{R}^s \, \tilde{\mu}_2$. ◀

▶ **Theorem 39.** *Let X_1 and X_2 be standard Borel spaces and $R \subseteq X_1 \times X_2$. Suppose $(\underline{X}, \kappa_1, \kappa_2)$ countably separates R. Then, $R^w = R^s$. Moreover, for all $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$,*

$$\mu_1 \, R^w \, \mu_2 \quad \textit{iff} \quad \mu_1 \, R^s \, \mu_2 \quad \textit{iff} \quad (\kappa_1)_\sharp(\mu_1) = (\kappa_2)_\sharp(\mu_2).$$

Proof. Let $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$. In Proposition 35 we have already seen the first implication, i.e., $\mu_1 \, R^w \, \mu_2$ implies $\mu_1 \, R^s \, \mu_2$. Assume $\mu_1 \, R^s \, \mu_2$ now. Since $(\underline{X}, \kappa_1, \kappa_2)$ countably separates R,

$$\mathsf{Diag}_{\underline{X}} = \{\langle \kappa_1(x_1), \kappa_2(x_2)\rangle \; ; \; x_1 \, R \, x_2\}.$$

Therefore, $(\kappa_1)_\sharp(\mu_1) \, \mathsf{Diag}^s_{\underline{X}} \, (\kappa_2)_\sharp(\mu_2)$ by Lemma 38 (2). and hence $(\kappa_1)_\sharp(\mu_1) = (\kappa_2)_\sharp(\mu_2)$ applying Lemma 37. For the remaining implication, i.e., there is a weight function for (μ_1, R^w, μ_2) provided $(\kappa_1)_\sharp(\mu_1) = (\kappa_2)_\sharp(\mu_2)$, see Proposition A.7 in [43, 42]. ◀

We use the notations from Theorem 39 and consider the following implication,

$$(\kappa_1)_\sharp(\mu_1) = (\kappa_2)_\sharp(\mu_2) \quad \text{implies} \quad \mu_1 \, R^w \, \mu_2.$$

This implication has been also attended in [24, 25, 26, 27] where one discusses this problem in a categorial setting in the context of the existence of pullbacks. For a quick overview and more details we refer to [47] (cf. the first part of Chapter 13).

Let us state two important corollaries of Theorem 39.

▶ **Corollary 40.** *Let X_1 and X_2 be Polish spaces and $R \subseteq X_1 \times X_2$ be an strong quasi-equivalence, which is closed in $X_1 \times X_1$. Then, $R^w = R^s$.*

Proof. Theorems 32 and 39 yield the claim. ◀

▶ **Corollary 41.** *Let X_1 and X_2 be sigma-compact Polish spaces and $R \subseteq X_1 \times X_2$ be an quasi-equivalence, that is closed in $X_1 \times X_2$. Then, $R^w = R^s$.*

Proof. The claim follows from Corollary 33 and Theorem 39. ◀

B Spans

B.1 Basic notions

▶ **Definition 42.** A *span* is a tuple $(X, X_1, X_2, \iota_1, \iota_2)$ consisting of Polish spaces X, X_1, and X_2 and continuous functions $\iota_1 \colon X \to X_1$ and $\iota_2 \colon X \to X_2$. We call a span $(X, X_1, X_2, \iota_1, \iota_2)$ *proper* if $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is compact in X for all compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$.

Let $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ be a span. We refer to ι_1 and ι_2 as the *$\mathcal{X}$-projections.* If there is no room for confusion, then we write $x_{|1}$ and $x_{|2}$ instead of $\iota_1(x)$ and $\iota_2(x)$, respectively, for all $x \in X$. For a short notation we sometimes drop the $\mathcal{X}$-projections from the notation and refer to (X, X_1, X_2) as a span. Given $x_1 \in X_1$ and $x_2 \in X_2$, then we write $x_1 \, \mathcal{X} \, x_2$, if there exists $x \in X$ such that $x_{|1} = x_1$ and $x_{|2} = x_2$. We define the relation $\mathsf{Rel}(\mathcal{X}) \subseteq X_1 \times X_2$ by

$$\mathsf{Rel}(\mathcal{X}) = \{\langle x_1, x_2\rangle \in X_1 \times X_2 \; ; \; x_1 \, \mathcal{X} \, x_2\}.$$

▶ **Example 43.** We consider important instances of proper spans where $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$:

1. $\mathcal{X}$ is a *Cartesian span* if $X = X_1 \times X_2$ and ι_1 and ι_2 are the natural projections.

2. $\mathcal{X}$ is a *variable span* if $X_1 = \mathsf{Ev}(\mathsf{Var}_1)$, $X_2 = \mathsf{Ev}(\mathsf{Var}_2)$, and $X = \mathsf{Ev}(\mathsf{Var}_1 \cup \mathsf{Var}_2)$ for some sets of variables Var_1 and Var_2, and ι_1 and ι_2 are the natural projections.
3. $\mathcal{X}$ is an *identity span* if $X = X_1 = X_2$ and $\iota_1(x) = x$ and $\iota_2(x) = x$ for all $x \in X$.

It is not difficult to give a span, that is not proper: Consider for instance the span $\mathcal{X} = (\mathbb{R}^3, \mathbb{R}, \mathbb{R}, \iota_1, \iota_2)$ where $\iota_1 \colon \mathbb{R}^3 \to \mathbb{R}$, $\iota_1(r_1, r_2, r_3) = r_1$ and $\iota_2 \colon \mathbb{R}^3 \to \mathbb{R}$, $\iota_2(r_1, r_2, r_3) = r_2$. Here, for every compact sets $K_1 \subseteq \mathbb{R}$ and $K_2 \subseteq \mathbb{R}$ it holds $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2) = K_1 \times K_2 \times \mathbb{R}$ and hence, $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is not compact in $\mathbb{R}^3$.

▶ **Proposition 44.** *If $\mathcal{X} = (X, X_1, X_2)$ is a span where X is compact, then $\mathcal{X}$ is proper. In particular, every span (X, X_1, X_2) with finite set X is proper.*

Proof. Let $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ be a span and assume X is compact. Let $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$ be compact sets. It is well-known that compact sets in Hausdorff spaces are closed and thus, K_1 and K_2 are closed in X_1 and X_2, respectively. Using the continuity of ι_1 and ι_2 the set $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is closed in X. Since X is compact, $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is therefore compact in X. This yields the claim. ◀

▶ **Proposition 45.** *Let $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ be a span. For all compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$ the set $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is compact in X iff it is relatively compact in X.*

Proof. The argument is as in Proposition 44: Given compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$, then $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is closed in X. A closed set is compact iff it is relatively compact and hence we are done. ◀

▶ **Proposition 46.** *Let $\mathcal{X} = (X, X_1, X_2)$ be a span. The following statements are equivalent:*

1. *$\mathcal{X}$ is proper.*
2. *The function $\iota \colon X \to X_1 \times X_2$, $\iota(x) = \langle x_{|1}, x_{|2} \rangle$ is proper, i.e., $\iota^{-1}(K_{12})$ is compact in X for all compact sets $K_{12} \subseteq X_1 \times X_2$*
3. *For all converging sequences $(x_{1,n})_n$ in X_1 and $(x_{2,n})_n$ in X_2 the statement below holds: If $(x_n)_n$ is a sequence in X such that $x_{n|1} = x_{1,n}$ and $x_{n|2} = x_{2,n}$ for all $n \in \mathbb{N}$, then $(x_n)_n$ has a subsequence, that converges in X.*

Proof. Throughout the proof denote the $\mathcal{X}$-projections by ι_1 and ι_2.

(1)→(3). Suppose $\mathcal{X}$ is proper. Let $(x_{1,n})_n$ and $(x_{2,n})_n$ be converging sequences in X_1 and X_2, repsectively. Assume $(x_n)_n$ is a sequence in X such that $x_{n|1} = x_{1,n}$ and $x_{n|2} = x_{2,n}$ for all $n \in \mathbb{N}$. Define $K_1 = \{x_{1,n} \,;\, n \in \mathbb{N}\}$ and $K_2 = \{x_{2,n} \,;\, n \in \mathbb{N}\}$. Having Remark 23 in mind, K_1 and K_2 are compact in X_1 and X_2, respectively. Since $\mathcal{X}$ is proper, $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$ is compact in X. As $(x_n)_n$ is a sequene in $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$, there thus exists a subsequence of $(x_n)_n$, that converges in X.

(3)→(2). Assume statement (3) holds. Let K_{12} be a compact set in $X_1 \times X_2$ and define $K = \iota^{-1}(K_{12})$. As in Propositions 44 and 45, K is closed in X. Thus, it order to show that K is compact, it suffices to justify that every sequence in K has a subsequence, that converges in X. Let $(x_n)_n$ be a sequence in K. Then, $(x_{n|1})_n$ and $(x_{n|2})_n$ are sequences in K_1 and K_2, respectively. Since K_1 is compact in X_1, there exists a strictly increasing function $\sigma_1 \colon \mathbb{N} \to \mathbb{N}$ such that $(x_{\sigma_1(n)|1})_n$ converges in X_1. Since $(x_{\sigma_1(n)|2})_n$ is a sequence in K_2 as well, there is a strictly increasing function $\sigma_2 \colon \mathbb{N} \to \mathbb{N}$ where $(x_{\sigma_2(\sigma_1(n))|2})_n$ converges in X_2. Define $\sigma \colon \mathbb{N} \to \mathbb{N}$ by $\sigma = \sigma_2 \diamond \sigma_1$. Subsequences of converging sequences converge as well and thus, $(x_{\sigma(n)|1})_n$ and $(x_{\sigma(n)|2})_n$ converge in X_1 and X_2, respectively. Applying (3), $(x_{\sigma(n)})_n$ has a subsequence that converges in X. Of course, this subsequence is also a subsequence of $(x_n)_n$, that finally yields the claim.

(2)→(1). Given compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$, the Cartesian product $K_1 \times K_2$ is compact in $X_1 \times X_2$. Since $\iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2) = \iota^{-1}(K_1 \times K_2)$, the claim follows. ◀

▶ **Proposition 47.** *Let $\mathcal{X} = (X, X_1, X_2)$ be a span. The following statements hold:*

1. $\mathsf{Rel}(\mathcal{X})$ *is a Souslin set in* $X_1 \times X_2$.
2. *If $\mathcal{X}$ is proper, then* $\mathsf{Rel}(\mathcal{X})$ *is closed in* $X_1 \times X_2$.

Proof. Define the continuous function $\iota\colon X \to X_1 \times X_2$, $\iota(x) = \langle x_{|1}, x_{|2}\rangle$. Since $\mathsf{Rel}(\mathcal{X}) = \iota(X)$, claim (1) follows by the definition of a Souslin set.

We consider statement (2) and assume that $\mathcal{X}$ is proper. We justify that the limit of every convergent sequences in $\mathsf{Rel}(\mathcal{X})$ is contained $\mathsf{Rel}(\mathcal{X})$. From this the closeness of $\mathsf{Rel}(\mathcal{X})$ immediately follows. Let $(\langle x_{1,n}, x_{2,n}\rangle)_n$ be a convergent sequence in $\mathsf{Rel}(\mathcal{X})$ and $\langle x_1, x_2\rangle \in X_1 \times X_2$ its limit. Notice, $(x_{1,n})_n$ converges to x_1 and $(x_{1,n})_n$ converges to x_2. By the definition of $\mathsf{Rel}(\mathcal{X})$ there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $x_{n|1} = x_{1,n}$ and $x_{n|2} = x_{2,n}$ for all $n \in \mathbb{N}$. According to Proposition 46 (3) there is a strictly increasing function $\sigma\colon \mathbb{N} \to \mathbb{N}$ such that $(x_{\sigma(n)})_n$ converges in X. Denote the corresponding limit by x. Since ι_1 is continuous, $(x_{\sigma(n)|1})_n$ converges to $x_{|1}$ in X_1. Therefore,

$$x_1 = \lim_{n\to\infty} x_{1,n} = \lim_{n\to\infty} x_{1,\sigma(n)} = \lim_{n\to\infty} x_{\sigma(n)|1} = x_{|1}.$$

One similiar shows $x_2 = x_{|2}$ and hence $x_1 \,\mathcal{X}\, x_2$. This finishes the proof. ◀

▶ **Example 48.** There exists a span $\mathcal{X} = (X, X_1, X_2)$ such that $\mathsf{Rel}(\mathcal{X})$ is not Borel in $X_1 \times X_2$. Let us give the details. Having Remark 20 in mind there exists a Borel set $M \subseteq \mathbb{R} \times \mathbb{R}$ such that $M_1 = M_1 = \{r_1 \in \mathbb{R}\,;\ \text{there is } r_2 \in \mathbb{R} \text{ where } \langle r_1, r_2\rangle \in M\}$ is not Borel in $\mathbb{R}$. Denote the natural topology on $\mathbb{R} \times \mathbb{R}$ by $\mathcal{O}_N$. Since M is Borel in $\mathbb{R} \times \mathbb{R}$ there is a Polish topology $\mathcal{O}$ on $\mathbb{R} \times \mathbb{R}$ such that $M \in \mathcal{O}$ and $\mathcal{O}_N \subseteq \mathcal{O}$. (cf. Remark 11). Thus, the set M equipped with the induced topology of $(\mathbb{R} \times \mathbb{R}, \mathcal{O})$ constitutes a Polish space (cf. Remark 10). Moreover, using $\mathcal{O}_N \subseteq \mathcal{O}$, the function $\iota_1\colon M \to \mathbb{R}$, $\iota_1(r_1, r_2) = r_1$ is continuous. Of course, the constant function $\iota_2\colon M \to \{0\}$, $\iota_2(r_1, r_2) = 0$ is continuous as well. Therefore, $\mathcal{X} = (M, \mathbb{R}, \{0\}, \iota_1, \iota_2)$ is a span. However, $\mathsf{Rel}(\mathcal{X})$ is not Borel in $\mathbb{R} \times \{0\}$ since $\mathsf{Rel}(\mathcal{X}) = M_1 \times \{0\}$. As a side remark it follows that $\mathcal{X}$ is not proper by Proposition 47 (2).

B.2 Span couplings

Given a span $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ and $\mu \in \mathsf{Prob}(X)$, then we write $\mu_{|1}$ and $\mu_{|2}$ as shorthand notations for $(\iota_1)_\sharp(\mu) = \mu_1$ and $(\iota_2)_\sharp(\mu) = \mu_2$, respectively.

▶ **Definition 49.** Let $\mathcal{X} = (X, X_1, X_2)$ be a span and $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$, A probability measure $\mu \in \mathsf{Prob}(X)$ is called a *$\mathcal{X}$-coupling of* (μ_1, μ_2) if

$$\mu_{|1} = \mu_1 \quad \text{and} \quad \mu_{|2} = \mu_2.$$

If there exists a $\mathcal{X}$-coupling of (μ_1, μ_2), then we denote $\mu_1 \,\mathcal{X}^c\, \mu_2$.

▶ **Proposition 50.** *If $\mathcal{X} = (X, X_1, X_2)$ is a span, then for all $x_1 \in X_1$ and $x_2 \in X_2$,*

$$x_1 \,\mathcal{X}\, x_2 \quad \textit{iff} \quad \mathsf{Dirac}[x_1] \,\mathcal{X}^c\, \mathsf{Dirac}[x_2].$$

Proof. The claim follows directly from Proposition 34. ◀

▶ **Proposition 51.** *If $\mathcal{X} = (X, X_1, X_2)$ is a Cartesian span, then $\mu_1 \,\mathcal{X}^c\, \mu_2$ for all $\mu_1 \in$* $\mathsf{Prob}(X_1)$ *and $\mu_2 \in$* $\mathsf{Prob}(X_2)$.

Proof. Given a Cartesian span $\mathcal{X} = (X, X_1, X_2)$, i.e., $X = X_1 \times X_2$, then $\mu_1 \otimes \mu_2$ is a $\mathcal{X}$-coupling of (μ_1, μ_2) for all $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$. ◀

▶ **Proposition 52.** *Given a span* $\mathcal{X} = (X, X_1, X_2)$*, for all* $\mu_1 \in \mathsf{Prob}(X_1)$ *and* $\mu_2 \in \mathsf{Prob}(X_2)$*,*

$$\mu_1\, \mathcal{X}^c\, \mu_2 \quad \textit{implies} \quad \mu_1\, \mathsf{Rel}(\mathcal{X})^w\, \mu_2 \quad \textit{implies} \quad \mu_1\, \mathsf{Rel}(\mathcal{X})^s\, \mu_2.$$

Proof. Denote the $\mathcal{X}$-projections by $\iota_1 \colon X \to X_1$ and $\iota_2 \colon X \to X_2$. Let $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$ be such that $\mu_1\, \mathcal{X}^c\, \mu_2$. Suppose μ is a $\mathcal{X}$-coupling of (μ_1, μ_2). Define the Borel function $\iota \colon X \to X_1 \times X_2$, $\iota(x) = \langle x_{|1}, x_{|2}\rangle$ and let $W = \iota_\sharp(\mu)$. We claim that W is a weight function for $(\mu_1, \mathsf{Rel}(\mathcal{X}), \mu_2)$. It is not hard to see that W is a coupling of (μ_1, μ_2). Indeed, for every Borel set $M_1 \subseteq X_1$,

$$W(M_1 \times X_2) = \mu(\iota^{-1}(M_1 \times X_2)) = \mu(\iota_1^{-1}(M_1)) = \mu_{|1}(M_1) = \mu_1(M_1).$$

Analogously, $W(X_1 \times M_2) = \mu_2(M_2)$ for all Borel sets $M_2 \subseteq X_2$.

We show that $x_1\, \mathsf{Rel}(\mathcal{X})\, x_2$ for W-almost all $\langle x_1, x_2\rangle \in X_1 \times X_2$. According to Proposition 47 (1), $\mathsf{Rel}(\mathcal{X})$ is a Souslin set in $X_1 \times X_2$. By Remark 21 there exist Borel sets $R_l, R_u \subseteq X_1 \times X_2$ such that $R_l \subseteq \mathsf{Rel}(\mathcal{X}) \subseteq R_u$ and $W(R_l) = W(R_u)$. Notice, $X = \iota^{-1}(\mathsf{Rel}(\mathcal{X})) \subseteq \iota^{-1}(R_u) \subseteq X$ and thus $X = \iota^{-1}(R_u)$. It follows

$$W(R_l) = W(R_u) = \mu(\iota^{-1}(R_u)) = \mu(X) = 1.$$

Therefore, W is a weight function for $(\mu_1, \mathsf{Rel}(\mathcal{X}), \mu_2)$ and hence $\mu_1\, \mathsf{Rel}(\mathcal{X})^w\, \mu_2$.

By the help of Proposition 35, $\mu_1\, \mathsf{Rel}(\mathcal{X})^s\, \mu_2$ follows from $\mu_1\, \mathsf{Rel}(\mathcal{X})^w\, \mu_2$ and hence our proof is complete. ◀

For proper spans we can strengthen the first part of Proposition 52: We will see that for all $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$ one has $\mu_1\, \mathcal{X}^c\, \mu_2$ iff $\mu_1\, \mathsf{Rel}(\mathcal{X})^w\, \mu_2$ provided $\mathcal{X}$ is proper (cf. Corollary 60 below). Remind that we have already discussed the connections between the weight lifting $\mathsf{Rel}(\mathcal{X})^w$ and the stable-pair lifting $\mathsf{Rel}(\mathcal{X})^s$ in Section A.3.

B.3 Operations for spans

There are various operations for spans, that yield complex spans out of some given basic spans. The question whether the operation preserves properness is important for practical purposes.

Probabilistic version.

▶ **Definition 53.** Given a span $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$, its *probilistic version* is defined to be the tuple

$$\mathsf{Prob}(\mathcal{X}) = (\mathsf{Prob}(X), \mathsf{Prob}(X_1), \mathsf{Prob}(X_2), (\iota_1)_\sharp, (\iota_2)_\sharp).$$

▶ **Proposition 54.** *The probabilistic version of a span is a span. Moreover, the probabilistic version of a proper span is proper as well.*

Proof. Let $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ be a span. According to Remark 17 the pushforwards of ι_1 and ι_2 are continuous. Since $\mathsf{Prob}(X_1)$, $\mathsf{Prob}(X_2)$, and $\mathsf{Prob}(X)$ are Polish spaces when equipped with the topology of weak-convergence (cf. Remark 10), respectively, the tuple $\mathsf{Prob}(\mathcal{X})$ constitutes indeed a span.

Assume $\mathcal{X}$ is proper now. Our task is to show that $\mathsf{Prob}(\mathcal{X})$ is proper. To that end let $P_1 \subseteq \mathsf{Prob}(X_1)$ and $P_2 \subseteq \mathsf{Prob}(X_2)$ be compact sets and define $P = (\iota_1)_\sharp^{-1}(P_1) \cap (\iota_2)_\sharp^{-1}(P_2)$, i.e.,

$$P = \{\mu \in \mathsf{Prob}(X) \; ; \; \mu_{|1} \in P_1 \text{ and } \mu_{|_2} \in P_2\}.$$

It is enought to show that P is tight in $\mathsf{Prob}(X)$ since it then follows that P is relatively compact in X (cf. Remark 22) and thus compact in X (cf. Proposition 45). Let $\varepsilon \in \mathbb{R}_{>0}$. According to Remark 22 the sets P_1 and P_2 are tight in $\mathsf{Prob}(X_1)$ and $\mathsf{Prob}(X_2)$, respectively. There hence exist compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$ such that $\mu_1(K_1) > 1 - \varepsilon/2$ for all $\mu_1 \in P_1$ and $\mu_2(K_2) > 1 - \varepsilon/2$ for all $\mu_2 \in P_2$. Define

$$K = \iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2).$$

Using that $\mathcal{X}$ is proper we conclude that K is compact in X. Moreover, for all $\mu \in P$,

$$\begin{aligned} & \mu(X \setminus K) \\ = \; & \mu(\iota_1^{-1}(X_1 \setminus K_1) \cup \iota_2^{-1}(X_2 \setminus K_2)) \\ \leq \; & \mu(\iota_1^{-1}(X_1 \setminus K_1)) + \mu_2(\iota_2^{-1}(X_2 \setminus K_2) \\ = \; & \mu_{|1}(X_1 \setminus K_1) + \mu_{|2}(X_2 \setminus K_2) \\ < \; & \varepsilon/2 + \varepsilon/2 \\ = \; & \varepsilon, \end{aligned}$$

and thus $\mu(K) > 1 - \varepsilon$. It follows that P is tight in $\mathsf{Prob}(X)$. As discussed before this yields that the probabilistic version of $\mathcal{X}$ is proper. ◀

Cartesian product.

▶ **Definition 55.** Let $N \subseteq \mathbb{N}$ be a non-empty set and for every $n \in N$ suppose a span $\mathcal{X}_n = (X_n, X_{1,n}, X_{2,n}, \iota_{1,n}, \iota_{2,n})$. The *Cartesian product of* $(\mathcal{X}_n)_{n \in N}$ is defined by the tuple $\prod_{n \in N} \mathcal{X}_n = (X, X_1, X_2, \iota_1, \iota_2)$, where

$$X = \prod_{n \in N} X_n, \quad X_1 = \prod_{n \in N} X_{1,n}, \quad X_2 = \prod_{n \in N} X_{2,n},$$

and $\iota_1 \colon X \to X_1$ and $\iota_2 \colon X \to X_2$ are given by

$$\iota_1(x)[n] = \iota_{1,n}(x[n]) \quad \text{and} \quad \iota_2(x)[n] = \iota_{2,n}(x[n]),$$

respectively, for all $x \in X$ and $n \in N$.

▶ **Proposition 56.** *Let $N \subseteq \mathbb{N}$ be a non-empty set and for every $n \in N$ let $\mathcal{X}_n$ be a span. Then, $\prod_{n \in N} \mathcal{X}_n$ is a span. Moreover, if $\mathcal{X}_n$ is proper for all $n \in N$, then $\prod_{n \in N} \mathcal{X}_n$ is proper.*

Proof. For every $n \in N$ assume the span $\mathcal{X}_n$ is given by $\mathcal{X}_n = (X_n, X_{1,n}, X_{2,n}, \iota_{1,n}, \iota_{2,n})$. Let X, X_1, X_2, ι_1 and ι_2 be given as in Definition 55 and abbreviate $\mathcal{X} = \prod_{n \in N} \mathcal{X}_n$. Reminding Remark 10 the sets X, X_1, and X_2 constitute Polish spaces, respectively. It is easy to see that ι_1 and ι_2 are continuous and hence, $\mathcal{X}$ is indeed a span.

Suppose $\mathcal{X}_n$ is proper for all $n \in N$ and let K_1 and K_2 be compact sets in X_1 and X_2, respectively. In order to show that $\mathcal{X}$ is proper, we argue that K is compact in X, where

$K = \iota_1^{-1}(K_1) \cap \iota_2^{-1}(K_2)$. For every $n \in N$ define the continuous functions $\iota_{1,[n]} \colon X_1 \to X_{1,n}$ and $\iota_{2,[n]} \colon X_2 \to X_{2,n}$ by

$$\iota_{1,[n]}(x_1) = x_1[n] \quad \text{and} \quad \iota_{2,[n]}(x_2) = x_2[n]$$

for all $x_1 \in X_1$ and $x_2 \in X_2$, respectively. For every $n \in N$ let

$$K_{1,[n]} = \iota_{1,[n]}(K_1) \quad \text{and} \quad K_{2,[n]} = \iota_{2,[n]}(K_2).$$

Define $K' \subseteq X$ by

$$K' = \iota_1^{-1}\Big(\prod_{n\in N} K_{1,[n]}\Big) \cap \iota_2^{-1}\Big(\prod_{n\in N} K_{2,[n]}\Big).$$

It is well-known that the continuous image of a compact set is compact and thus $K_{1,[n]}$ and $K_{2,[n]}$ are compact in $X_{1,n}$ and $X_{2,n}$, respectively, for all $n \in N$. Since

$$K' = \prod_{n\in N} \iota_{1,n}^{-1}(K_{1,[n]}) \cap \iota_{2,n}^{-1}(K_{2,[n]})$$

and $\mathcal{X}_n$ is proper for all $n \in N$, the set K' is hence compact in X applying Tychonoff's Theorem. We conclude that K is compact in X observing $K \subseteq K'$ and the fact that K is closed in X (cf. our argument for Proposition 44). This finishes our argument. ◀

B.4 Span inverse

▶ **Definition 57.** Let $\mathcal{X} = (X, X_1, X_2)$ be a span. A Borel function $f \colon X_1 \times X_2 \to X$ is called an *$\mathcal{X}$-inverse*, if for all $x_1 \in X_1$ and $x_2 \in X_2$,

$$x_1 \,\mathcal{X}\, x_2 \quad \text{implies} \quad f(x_1, x_2)_{|1} = x_1 \text{ and } f(x_1, x_2)_{|2} = x_2.$$

▶ **Example 58.** We present a span $\mathcal{X}$, which has no $\mathcal{X}$-inverse. According to Corollary 6.9.10 in [10] (for definitions see also for page 33 in [10]) there are Polish spaces X and X_1 as well as a continuous and surjective function $\iota_1 \colon X \to X_1$ satisfying the property below: ι_1 has no Borel right inverse, i.e., there is no Borel function $g \colon X_1 \to X$ such that for all $x_1 \in X_1$,

$$\iota_1(g(x_1)) = x_1.$$

We consider the span $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ where $X_2 = \{0\}$ and $\iota_2 \colon X \to X_2$, $\iota_2(x) = 0$. Notice, for every $x_1 \in X_1$ and $x_2 \in X_2$ it exists $x \in X$ where $x_{|1} = x_1$ and $x_{|2} = x_2$. Towards a contradiction assume $f \colon X_1 \times X_2 \to X$ is an $\mathcal{X}$-inverse. Then, f is Borel and hence $h \colon X_1 \to X$, $h(x_1) = f(x_1, 0)$ is Borel as well. Moreover, for all $x_1 \in X_1$,

$$\iota_1(h(x_1)) = \iota_1(f(x_1, 0)) = x_1.$$

This contradicts the fact that there is no Borel right inverse of ι_1. It hence follows that there exists no $\mathcal{X}$-inverse.

▶ **Theorem 59.** *Every proper span $\mathcal{X}$ has an $\mathcal{X}$-inverse.*

Proof. Our argument relies on a measurable selection theorem. Let $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ be a proper span and $\hat{x} \in X$ and abbreviate $X_{12} = X_1 \times X_2$. Define the Borel function $\iota \colon X \to X_{12}$, $\iota(x) = \langle x_{|1}, x_{|2} \rangle$. Moreover, let $R \subseteq X_{12} \times X$ be given by

$$R = \mathsf{graph}(\iota)^{-1} \cup ((X_{12} \setminus \mathsf{Rel}(\mathcal{X})) \times \{\hat{x}\}).$$

Observe that R is a Borel set in $X_{12}\times X$ since $\mathsf{graph}(\iota)$ is a Borel set in $X_{12}\times X$ (cf. Remark 16) and $\mathsf{Rel}(\mathcal{X})$ is a Borel set in X_{12} (cf. Proposition 47 (2)). For all $\langle x_1,x_2\rangle \in X_{12}\cap \mathsf{Rel}(\mathcal{X})$,

$$[\langle x_1,x_2\rangle]_{R,-} = \iota_1^{-1}(\{x_1\})\cap \iota_2^{-1}(\{x_2\})$$

is compact in X using the assumption that $\mathcal{X}$ is proper. For all $\langle x_1,x_2\rangle \in X_{12}\setminus \mathsf{Rel}(\mathcal{X})$,

$$[\langle x_1,x_2\rangle]_{R,-} = \{\hat{x}\}$$

is compact in X as well. Putting things together $[\langle x_1,x_2\rangle]_{R,-}$ is compact in X for all $\langle x_1,x_2\rangle \in X_{12}$. We are in the situation of the measurable selection theorem given in Remark 29 and hence there exists a Borel function $f\colon X_{12}\to X$ such that $\mathsf{graph}(f)\subseteq R$, i.e., for all $x_1\in X_1$ and $x_2\in X_2$,

$$\langle x_1,x_2\rangle \, R\, f(x_1,x_2).$$

It is easy to see that f is an $\mathcal{X}$-inverse, that finishes our argumentation. ◀

▶ **Corollary 60.** *Let $\mathcal{X} = (X,X_1,X_2)$ be a proper span. For every $\mu_1\in \mathsf{Prob}(X_1)$ and $\mu_2\in\mathsf{Prob}(X_2)$ the following equivalence holds,*

$$\mu_1\,\mathcal{X}^c\,\mu_2 \quad \textit{iff} \quad \mu_1\,\mathsf{Rel}(\mathcal{X})^w\,\mu_2.$$

Proof. Denote the $\mathcal{X}$-projections by $\iota_1\colon X\to X_1$ and $\iota_2\colon X\to X_2$. Let $\mu_1\in\mathsf{Prob}(X_1)$ and $\mu_2\in\mathsf{Prob}(X_2)$. Proposition 52 already shows one implication of the claimed equivalence. We consider the remaining implication and assume $\mu_1\,\mathsf{Rel}(\mathcal{X})^w\,\mu_2$. Let W be a weight function for $(\mu_1,\mathsf{Rel}(\mathcal{X}),\mu_2)$. Remind, $\mathsf{Rel}(\mathcal{X})$ is a Borel set in $X_1\times X_2$ by Proposition 47 (2). According to Theorem 59 there exists an $\mathcal{X}$-inverse f. Set $\mu = f_\sharp(W)$. It turns out that μ is a $\mathcal{X}$-coupling of (μ_1,μ_2). Indeed, for all Borel sets $M_1\subseteq X_1$,

$$f^{-1}(\iota_1^{-1}(M_1))\cap\mathsf{Rel}(\mathcal{X}) = (M_1\times X_2)\cap\mathsf{Rel}(\mathcal{X})$$

and therefore

$$\begin{aligned}
& \mu_{|1}(M_1)\\
=\; & \mu(\iota_1^{-1}(M_1))\\
=\; & W(f^{-1}(\iota_1^{-1}(M_1)))\\
=\; & W(f^{-1}(\iota_1^{-1}(M_1))\cap\mathsf{Rel}(\mathcal{X}))\\
=\; & W((M_1\times X_2)\cap\mathsf{Rel}(\mathcal{X}))\\
=\; & W(M_1\times X_2)\\
=\; & \mu_1(M_1).
\end{aligned}$$

Thus, $\mu_{|1}=\mu_1$. One similiar shows $\mu_{|2}=\mu_2$ and hence we are done. ◀

▶ **Corollary 61.** *Let $\mathcal{X}_a = (X_a,X_{a1},X_{a2})$ be a span and $\mathcal{X}_b=(X_b,X_{b1},X_{b2})$ be a proper span. Suppose Markov kernels $k_1\colon X_{a1}\to\mathsf{Prob}(X_{b1})$ and $k_2\colon X_{a2}\to\mathsf{Prob}(X_{b2})$. Then there is a Markov kernel $k\colon X_a\to\mathsf{Prob}(X_b)$ where for all $x_a\in X_a$, if $k_1(x_{a|1})\,\mathcal{X}_b^c\,k_2(x_{a|2})$, then*

$$k(x_a) \textit{ is a } \mathcal{X}_b\textit{-coupling of } (k_1(x_{a|1}),k_2(x_{a|2})).$$

Proof. Let $f\colon\mathsf{Prob}(X_{b1})\times\mathsf{Prob}(X_{b2})\to\mathsf{Prob}(X_b)$ be a $\mathsf{Prob}(\mathcal{X}_b)$-inverse (cf. Proposition 54 and Theorem 59). Define $k\colon X_a\to\mathsf{Prob}(X_b)$,

$$k(x_a) = f(k_1(x_{a|1}),k_2(x_{a|2})).$$

Notice, k is a Markov kernel, since the composition of Borel functions yields a Borel function. Using the fact that f is an $\mathsf{Prob}(\mathcal{X}_b)$-inverse (cf. Definition 57), the claim follows. ◀

B.5 Countably separated spans

▶ **Definition 62.** We call a span $\mathcal{X} = (X, X_1, X_2)$ *countably separated* if $\mathsf{Rel}(\mathcal{X})$ is countably separated, i.e., there exists a Polish space $\underline{X}$ and Borel functions $\kappa_1 \colon X_1 \to \underline{X}$ and $\kappa_2 \colon X_2 \to \underline{X}$ such that

$$\mathsf{Rel}(\mathcal{X}) = \{\langle x_1, x_2\rangle \in X_1 \times X_2 \ ; \ \kappa_1(x_1) = \kappa_2(x_2)\},$$

i.e., for all $x_1 \in X_1$ and $x_2 \in X_2$,

$$x_1 \, \mathcal{X} \, x_2 \quad \text{iff} \quad \kappa_1(x_1) = \kappa_2(x_2).$$

Here, we then say that $(\underline{X}, \kappa_1, \kappa_2)$ *countably separates* $\mathcal{X}$.

▶ **Example 63.** Assume $\mathcal{X} = (X, X_1, X_2)$ is a variable span, i.e., $X_1 = \mathsf{Ev}(\mathsf{Var}_1)$, $X_2 = \mathsf{Ev}(\mathsf{Var}_2)$, and $X = \mathsf{Ev}(\mathsf{Var}_1 \cup \mathsf{Var}_2)$ for sets of variables Var_1 and Var_2. Denote the set of shared variables by $\mathsf{SVar} = \mathsf{Var}_1 \cap \mathsf{Var}_2$. Then, $(\mathsf{Ev}(\mathsf{SVar}), \kappa_1, \kappa_2)$ countably separates $\mathcal{X}$, where $\kappa_1(e_1) = e_{1|\mathsf{SVar}}$ for all $e_1 \in \mathsf{Ev}(\mathsf{Var}_1)$ and $\kappa_2(e_2) = e_{2|\mathsf{SVar}}$ for all $e_2 \in \mathsf{Ev}(\mathsf{Var}_2)$.

▶ **Proposition 64.** $\mathsf{Rel}(\mathcal{X})$ *is an quasi-equivalence in* $X_1 \times X_2$ *for every countably separated span* $\mathcal{X} = (X, X_1, X_2)$.

Proof. The claim follows directly from the definitions. ◀

▶ **Proposition 65.** *Let* $\mathcal{X} = (X, X_1, X_2)$ *be a proper span and assume* $(\underline{X}, \kappa_1, \kappa_2)$ *countably separates* $\mathcal{X}$*. For all* $\mu_1 \in \mathsf{Prob}(X_1)$ *and* $\mu_2 \in \mathsf{Prob}(X_2)$*,*

$$\mu_1 \, \mathcal{X}^c \, \mu_2 \quad \textit{iff} \quad \mu_1 \, \mathsf{Rel}(\mathcal{X})^w \, \mu_2 \quad \textit{iff} \quad \mu_1 \, \mathsf{Rel}(\mathcal{X})^s \, \mu_2 \quad \textit{iff} \quad (\kappa_1)_\sharp(\mu_1) = (\kappa_2)_\sharp(\mu_2).$$

Proof. The claim is a consequence of Theorem 39. and Corollary 60. ◀

▶ **Proposition 66.** *If* $(\underline{X}, \kappa_1, \kappa_2)$ *countably separates the proper span* $\mathcal{X}$*, then* $\mathsf{Prob}(\mathcal{X})$ *is countably separated and* $(\mathsf{Prob}(\underline{X}), (\kappa_1)_\sharp, (\kappa_2)_\sharp)$ *countably separates* $\mathsf{Prob}(\mathcal{X})$*.*

Proof. Assume $(\underline{X}, \kappa_1, \kappa_2)$ countably separates the span $\mathcal{X} = (X, X_1, X_2)$. Of course, $\mathsf{Prob}(\underline{X})$ is a Polish space when equipped with the topology of weak-convergence of probability measures (cf. Remark 10) and $(\kappa_1)_\sharp$ and $(\kappa_2)_\sharp$ are Borel functions (cf. Remark 17). Let $\mu_1 \in \mathsf{Prob}(X_1)$ and $\mu_2 \in \mathsf{Prob}(X_2)$. To see that $(\mathsf{Prob}(\underline{X}), (\kappa_1)_\sharp, (\kappa_2)_\sharp)$ countably separates $\mathsf{Prob}(\mathcal{X})$ we use Proposition 65 and observe

$$\mu_1 \, \mathsf{Rel}(\mathsf{Prob}(\mathcal{X})) \, \mu_2 \quad \text{iff} \quad \mu_1 \, \mathcal{X}^c \, \mu_2 \quad \text{iff} \quad (\kappa_1)_\sharp(\mu_1) = (\kappa_2)_\sharp(\mu_2).$$

This completes the argument. ◀

B.6 Span connections

▶ **Definition 67.** A *span connection* is a tuple $(\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ consisting of spans $\mathcal{X}_a = (X_a, X_{a1}, X_{a2})$ and $\mathcal{X}_b = (X_b, X_{b1}, X_{b2})$ and relations $R_1 \subseteq X_{a1} \times X_{b1}$ and $R_2 \subseteq X_{a2} \times X_{b2}$. Let $\mathcal{C} = (\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ be a span connection. Set

$$R_1 \wedge_{\mathcal{C}} R_2 = \{\langle x_a, x_b\rangle \in X_a \times X_b \ ; \ x_{a|1} \, R_1 \, x_{b|1} \text{ and } x_{a|2} \, R_2 \, x_{b|2}\}.$$

We call $\mathcal{C}$ *l-adequate* if for all $\mu_{a1} \, R_1^w \, \mu_{b1}$ and $\mu_{a2} \, R_2^w \, \mu_{b2}$,

$$\mu_{a1} \, \mathcal{X}_a^c \, \mu_{a2} \quad \text{implies} \quad \mu_{b1} \, \mathcal{X}_b^c \, \mu_{b2}.$$

Vice verca, $\mathcal{C}$ is called *r-adequate*, if for all $\mu_{a1}\, R_1^w\, \mu_{b1}$ and $\mu_{a2}\, R_2^w\, \mu_{b2}$,

$$\mu_{b1}\, \mathcal{X}_b^c\, \mu_{b2} \quad \text{implies} \quad \mu_{a1}\, \mathcal{X}_a c\, \mu_{a2}.$$

$\mathcal{C}$ is defined to be *adequate* iff $\mathcal{C}$ is both, l-adequate and r-adequate.

▶ **Proposition 68.** *Let $\mathcal{C} = (\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ be a span connection. If $\mathcal{X}_b$ is a Cartesian span, then $\mathcal{C}$ is l-adequate. Similiar, if $\mathcal{X}_a$ is a Cartesian span, then $\mathcal{C}$ is r-adequate.*

Proof. Trivial (cf. Proposition 51). ◀

▶ **Proposition 69.** *Let $\mathcal{C} = (\mathcal{X}, \mathcal{X}, R_1, R_2)$ be a span connection with $\mathcal{X} = (X, X_1, X_2, \iota_1, \iota_2)$ being a proper span. Suppose $(\underline{X}, \kappa_1, \kappa_2)$ countably separates $\mathcal{X}$. Then, $\mathcal{C}$ is adequate if*

$$R_1 \subseteq \{\langle x_{a1}, x_{b1}\rangle \in X_1 \times X_1 \;;\, \kappa_1(x_{a1}) = \kappa_1(x_{b1})\} \quad \textit{and}$$
$$R_2 \subseteq \{\langle x_{a2}, x_{b2}\rangle \in X_2 \times X_2 \;;\, \kappa_2(x_{a2}) = \kappa_2(x_{b2})\}.$$

Proof. For every $\underline{M} \subseteq \underline{X}$ we justify the following statements first,

$$\langle \kappa_1^{-1}(\underline{M}), \kappa_1^{-1}(\underline{M})\rangle \text{ is } R_1\text{-stable,}$$
$$\langle \kappa_2^{-1}(\underline{M}), \kappa_2^{-1}(\underline{M})\rangle \text{ is } R_2\text{-stable,} \quad \text{and}$$
$$\langle \kappa_1^{-1}(\underline{M}), \kappa_2^{-1}(\underline{M})\rangle \text{ is } \mathsf{Rel}(\mathcal{X})\text{-stable.}$$

We argue that $\langle \kappa_1^{-1}(\underline{M}), \kappa_1^{-1}(\underline{M})\rangle$ is R_1-stable. If $\langle x_{a1}, x_{b1}\rangle \in R_1 \cap (\kappa_1^{-1}(\underline{M}) \times X_1)$, then $\kappa_1(x_{b1}) = \kappa_1(x_{a1}) \in \underline{M}$ and hence, $\langle x_{a1}, x_{b1}\rangle \in R_1 \cap (X_1 \times \kappa_1^{-1}(\underline{M}))$. One similar shows the reverse inclusion, i.e., $R_1 \cap (X_1 \times \kappa_1^{-1}(\underline{M})) \subseteq R_1 \cap (\kappa_1^{-1}(\underline{M}) \times X_1)$. It analogously follows that $\langle \kappa_2^{-1}(\underline{M}), \kappa_2^{-1}(\underline{M})\rangle$ is R_2-stable.

We show that $\langle \kappa_1^{-1}(\underline{M}), \kappa_2^{-1}(\underline{M})\rangle$ is $\mathsf{Rel}(\mathcal{X})$-stable. To do so let $\langle x_1, x_2\rangle \in \mathsf{Rel}(\mathcal{X}) \cap (\kappa_1^{-1}(\underline{M}) \times X_2)$. Since $(\underline{X}, \kappa_1, \kappa_2)$ countably separates $\mathsf{Rel}(\mathcal{X})$, we have $\kappa_2(x_2) = \kappa_1(x_1) \in \underline{M}$. Thus, $\langle x_1, x_2\rangle \in \mathsf{Rel}(\mathcal{X}) \cap (X_1 \times \kappa_2^{-1}(\underline{M}))$. The reverse inclusion $\mathsf{Rel}(\mathcal{X}) \cap (X_1 \times \kappa_2^{-1}(\underline{M})) \subseteq \mathsf{Rel}(\mathcal{X}) \cap (\kappa_1^{-1}(\underline{M}) \times X_2)$ follows similarly.

In what follows we conclude the claim of the proposition. Let $\mu_{a1}\, R_1^w\, \mu_{b1}$ and $\mu_{a2}\, R_2^w\, \mu_{b2}$ be such that $\mu_{a1}\, \mathcal{X}^c\, \mu_{a2}$. In order to show that $\mathcal{C}$ is l-adequate our task is to argue $\mu_{b1}\, \mathcal{X}^c\, \mu_{b2}$. According to Propositions 35 and 52, for every Borel set $\underline{M} \subseteq X$,

$$\mu_{b1}(\kappa_1^{-1}(\underline{M})) = \mu_{a1}(\kappa_1^{-1}(\underline{M})) = \mu_{a2}(\kappa_2^{-1}(\underline{M})) = \mu_{b2}(\kappa_2^{-1}(\underline{M})).$$

Thus, $(\kappa_1)_\sharp(\mu_{b1}) = (\kappa_2)_\sharp(\mu_{b2})$ and hence $\mu_{b1}\, \mathcal{X}^c\, \mu_{b2}$ by Proposition 65. It follows that $\mathcal{C}$ is l-adequate. R-adequacy of $\mathcal{C}$ is shown accordingly. ◀

▶ **Proposition 70.** *Let $\mathcal{C} = (\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ be a span connection. For all $\mu_a\, (R_1 \wedge_{\mathcal{C}} R_1)^w\, \mu_b$,*

$$\mu_{a|1}\, R_1^w\, \mu_{b|1} \quad \textit{and} \quad \mu_{a|2}\, R_2^w\, \mu_{b|2}.$$

Proof. Abbreviate $R = R_1 \wedge_{\mathcal{C}} R_1$. Assume the spans are given by $\mathcal{X}_a = (X_a, X_{a1}, X_{a2})$ and $\mathcal{X}_b = (X_b, X_{b1}, X_{b2})$. Define the Borel functions $\iota_1 \colon X_a \times X_b \to X_{a1} \times X_{b1}$ and $\iota_2 \colon X_a \times X_b \to X_{a2} \times X_{b2}$ for all $x_a \in X_a$ and $x_b \in X_b$ by

$$\iota_1(x_a, x_b) = \langle x_{a|1}, x_{b|1}\rangle \quad \text{and} \quad \iota_2(x_a, x_b) = \langle x_{a|2}, x_{b|2}\rangle.$$

Let $\mu_a \in \mathsf{Prob}(X_a)$, $\mu_b \in \mathsf{Prob}(X_b)$, and W be a weight function for (μ_a, R, μ_b). Then, $(\iota_1)_\sharp(W)$ is a weight function for $(\mu_{a|1}, R_1, \mu_{b|1})$ and $(\iota_2)_\sharp(W)$ is a weight function for $(\mu_{a|2}, R_2, \mu_{b|2})$. For reasons of symmetry we concentrate on $(\iota_1)_\sharp(W)$ only in the following.

Let $R' \subseteq X_a \times X_b$ be a Borel set such that $R' \subseteq R$ and $W(R') = 1$. Abbreviate $W_1 = (\iota_1)_\sharp(W)$ and $R'_1 = \iota_1(R')$. The set R'_1 is a Souslin set in $X_{a1} \times X_{b1}$ (cf. Remarks 18 and 19). Hence, there are Borel sets $R'_{1l}, R'_{1u} \subseteq X_{a1} \times X_{b1}$ such that $R'_{1l} \subseteq R'_1 \subseteq R'_{1u}$ and $W_1(R'_{1l}) = W_1(R'_{1u})$ (cf. Remark 21). Since $R' \subseteq \iota_1^{-1}(R'_{1u})$,

$$W_1(R'_{1l}) = W_1(R'_{1u}) = W(\iota_1^{-1}(R'_{1u})) \geq W(R') = 1.$$

It remains to show that W_1 is a coupling of $(\mu_{a|1}, \mu_{b|1})$. Denote the $\mathcal{X}_b$-projections by ι_{a1} and ι_{b1}. For all Borel sets $M_{a1} \subseteq X_{a1}$, it holds $\iota_1^{-1}(M_{a1} \times X_{b1}) = \iota_{a1}^{-1}(M_{a1}) \times X_b$ and therefore,

$$\begin{aligned} & W_1(M_{a1} \times X_{b1}) \\ = \; & W(\iota_1^{-1}(M_{a1} \times X_{b1})) \\ = \; & W(\iota_{a1}^{-1}(M_{a1}) \times X_b) \\ = \; & \mu_a(\iota_{a1}^{-1}(M_{a1})) \\ = \; & \mu_{a|1}(M_{a1}). \end{aligned}$$

One similar shows that $W_1(X_{a1} \times M_{b1}) = \mu_{b|1}(M_{b1})$ for all Borel sets $M_{b1} \subseteq X_{b1}$. ◀

▶ **Lemma 71.** *Let X_a and X_b be standard Borel spaces and $R \subseteq X_a \times X_b$. Let $\mu_a \in \mathsf{Prob}(X_a)$ and $k\colon X_a \to \mathsf{Prob}(X_b)$ be a Markov kernel such that $x_a \, R \, x_b$ for $(\mu_a \rtimes k)$-almost all $\langle x_a, x_b\rangle \in X_a \times X_b$. Then, for μ_a-almost all $x_a \in X_a$,*

$$\mathsf{Dirac}[x_a] \, R^w \, k(x_a).$$

Proof. Abbreviate $W = \mu_a \rtimes k$. Let $R' \subseteq X_a \times X_b$ be a Borel set such that $R' \subseteq R$ and $W(R') = 1$. For all $x_a \in X_a$ the set $[x_a]_{R',-}$ is Borel in X_b (cf. Remark 25). For every $x_a \in X_a$ define

$$W_{x_a} = \mathsf{Dirac}[x_a] \otimes k(x_a).$$

Notice that W_{x_a} is the only candidate for a weight function for $(\mathsf{Dirac}[x_a], R, k(x_a))$ for all $x_a \in X_a$ (cf. Proposition 24). For every $M_a \subseteq X_a$ and $x_a \in X_a$ observe

$$\begin{aligned} & x_a \in M_a \quad \text{implies} \quad [x_a]_{R',-} = [x_a]_{R' \cap (M_a \times X_b),-} \quad \text{and} \\ & x_a \notin M_a \quad \text{implies} \quad [x_a]_{R' \cap (M_a \times X_b),-} = \varnothing. \end{aligned}$$

According to Remark 15 and using $W(R') = 1$ we therefore obtain for all Borel sets $M_a \subseteq X_a$,

$$\begin{aligned} & \int_{M_a} k(x_a)([x_a]_{R',-}) \, d\mu_a(x_a) \\ = \; & \int_{M_a} k(x_a)([x_a]_{R' \cap (M_a \times X_b),-}) \, d\mu_a(x_a) \\ = \; & \int k(x_a)([x_a]_{R' \cap (M_a \times X_b),-}) \, d\mu_a(x_a) \\ = \; & W(R' \cap (M_a \times X_b)) \\ = \; & W(M_a \times X_b) \\ = \; & \mu_a(M_a) \\ = \; & \int_{M_a} f(x_a) \, d\mu_a(x_a), \end{aligned}$$

where $f\colon X_a \to \mathbb{R}$, $f(x_a) = 1$. It follows $k(x_a)([x_a]_{R',-}) = f(x_a) = 1$ for μ_a-almost all $x_a \in X_a$ (cf. Folgerung 9.2.5 in [48]). Therefore,

$$W_{x_a}(\{x_a\} \times [x_a]_{R',-}) = \mathsf{Dirac}[x_a](\{x_a\}) \cdot k(x_a)([x_a]_{R',-}) = 1$$

for μ_a-almost all $x_a \in X_a$. Since $\{x_a\} \times [x_a]_{R',-} \subseteq R' \subseteq R$ for all $x_a \in X_a$ our argumentation for the lemma is complete. ◀

▶ **Theorem 72.** *Let $\mathcal{C} = (\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ be a span connection with $\mathcal{X}_a = (X_a, X_{a1}, X_{a2})$ and $\mathcal{X}_b = (X_b, X_{b1}, X_{b2})$. Assuming $\mathcal{C}$ is l-adequate, for all $\mu_a \in \mathsf{Prob}(X_a)$, $\mu_{b1} \in \mathsf{Prob}(X_{b1})$, and $\mu_{b2} \in \mathsf{Prob}(X_{b2})$, if $\mu_{a|1}\, R_1^w\, \mu_{b1}$ and $\mu_{a|2}\, R_2^w\, \mu_{b2}$, then*

there exists a $\mathcal{X}_b$-coupling of (μ_{b1}, μ_{b2}) where $\mu_a\, (R_1 \wedge_{\mathcal{C}} R_2)^w\, \mu_b$.

Analogously, assuming $\mathcal{C}$ is r-adequate, for all $\mu_b \in \mathsf{Prob}(X_b)$, $\mu_{a1} \in \mathsf{Prob}(X_{a1})$, and $\mu_{a2} \in \mathsf{Prob}(X_{a2})$, if $\mu_{a1}\, R_1^w\, \mu_{b|1}$ and $\mu_{a2}\, R_2^w\, \mu_{b|2}$, then

there exists a $\mathcal{X}_a$-coupling of (μ_{a1}, μ_{a2}) where $\mu_a\, (R_1 \wedge_{\mathcal{C}} R_2)^w\, \mu_b$.

Proof. Denote $\mathcal{X}_a$-projections by ι_{a1} and ι_{a2} and the $\mathcal{X}_b$-projections by ι_{b1} and ι_{b2}. Abbreviate $R = R_1 \wedge_{\mathcal{C}} R_2$. Let $\mu_a \in \mathsf{Prob}(X_a)$, $\mu_{b1} \in \mathsf{Prob}(X_{b1})$, and $\mu_{b2} \in \mathsf{Prob}(X_{b2})$ be such that $\mu_{a|1}\, R_1^w\, \mu_{b1}$ and $\mu_{a|2}\, R_2^w\, \mu_{b2}$. Suppose W_1 and W_2 are weight functions for $(\mu_{a|1}, R_1, \mu_{b1})$ and $(\mu_{a|2}, R_2, \mu_{b2})$, respectively. Using disintegration of measures (cf. Remark 15) there are Markov kernels $k_1\colon X_{a1} \to \mathsf{Prob}(X_{b1})$ and $k_2\colon X_{a2} \to \mathsf{Prob}(X_{b2})$ such that

$$W_1 = \mu_{a|1} \rtimes k_1 \quad \text{and} \quad W_2 = \mu_{a|2} \rtimes k_2.$$

According to Lemma 71, for $\mu_{a|1}$-almost all $x_{a1} \in X_{a1}$ it holds $\mathsf{Dirac}[x_{a1}]\, R_1^w\, k_1(x_{a1})$ and similiar, for $\mu_{a|2}$-almost all $x_{a2} \in X_{a2}$, $\mathsf{Dirac}[x_{a2}]\, R_2^w\, k_2(x_{a2})$. From this we conclude that for μ_a-almost all $x_a \in X_a$,

$$\mathsf{Dirac}[x_{a|1}]\, R_1^w\, k_1(x_{a|1}) \quad \text{and} \quad \mathsf{Dirac}[x_{a|2}]\, R_2^w\, k_2(x_{a|2}).$$

Remind, $\mathsf{Dirac}[x_{a|1}]\, \mathcal{X}_a^c\, \mathsf{Dirac}[x_{a|2}]$ for all $x_a \in X_a$ (cf. Proposition 50). Using the assumption that $\mathcal{C}$ is l-adequate, for μ_a-almost all $x_a \in X_a$ it holds $k_1(x_{a|1})\, \mathcal{X}_b^c\, k_2(x_{a|2})$, i.e., there exists a $\mathcal{X}_b$-coupling of $(k_1(x_{a|1}), k_2(x_{a|2}))$. We are in the situation of Corollary 61 and hence there is a Markov kernel $k\colon X_a \to \mathsf{Prob}(X_b)$ such that for μ_a-almost all $x_a \in X_a$, the probability measure $k(x_a)$ is a $\mathcal{X}_b$-coupling of $(k_1(x_{a|1}), k_2(x_{a|2}))$. Define

$$W = \mu_a \rtimes k$$

and let $\mu_b \in \mathsf{Prob}(X_b)$ be given by

$$\mu_b(M_b) = W(X_a \times M_b)$$

for all Borel sets $M_b \subseteq X_b$.

In what follows we justify that μ_b is a $\mathcal{X}_b$-coupling of (μ_{b1}, μ_{b2}) such that $\mu_a\, R^w\, \mu_b$. For that purpose we observe that W is a $(\mathcal{X}_a \times \mathcal{X}_b)$-coupling of (W_1, W_2) first. Define the Borel functions $\iota_1\colon X_a \times X_b \to X_{a1} \times X_{b1}$ and $\iota_2\colon X_a \times X_b \to X_{a2} \times X_{b2}$ by

$$\iota_1(x_a, x_b) = \langle x_{a|1}, x_{b|1} \rangle \quad \text{and} \quad \iota_2(x_a, x_b) = \langle x_{a|2}, x_{b|2} \rangle$$

for every $x_a \in X_a$ and $x_b \in X_b$, respectively. For all Borel sets $M_{a1} \subseteq X_{a1}$ and $M_{b1} \subseteq X_{b1}$ it holds $\iota_1^{-1}(M_{a1} \times M_{b1}) = \iota_{a1}^{-1}(M_{a1}) \times \iota_{b1}^{-1}(M_{b1})$ and therefore, using integration by substitution,

$$\begin{aligned} & (\iota_1)_\sharp(W)(M_{a1} \times M_{b1}) \\ = & \int_{\iota_{a1}^{-1}(M_{a1})} k(x_a)(\iota_{b1}^{-1}(M_{b1}))\, d\mu_a(x_a) \\ = & \int_{\iota_{a1}^{-1}(M_{a1})} k_1(\iota_{a1}(x_a))(M_{b1})\, d\mu_a(x_a) \\ = & \int_{M_{a1}} k_1(x_{a1})(M_{b1})\, d\mu_{a|1}(x_{a1}) \\ = & W_1(M_{a1} \times M_{b1}). \end{aligned}$$

Carathéodory's measure extension theorem yields $(\iota_1)_\sharp(W) = W_1$. One similiar shows $(\iota_2)_\sharp(W) = W_2$ and thus, W is indeed a $(\mathcal{X}_a \times \mathcal{X}_b)$-coupling of (W_1, W_2).

Now it is not hard to see that μ_b is a $\mathcal{X}_b$-coupling of (μ_{b1}, μ_{b2}). Indeed, for all Borel sets $M_{b1} \subseteq X_{b1}$ we have $X_a \times \iota_{b1}^{-1}(M_{b1}) = \iota_1^{-1}(X_{a1} \times M_{b1})$ and hence,

$$\mu_{b|1}(M_{b1}) = \mu_b(\iota_{b1}^{-1}(M_{b1})) = W(X_a \times \iota_{b1}^{-1}(M_{b1})) = W_1(X_{a1} \times M_{b1}) = \mu_{b1}(M_{b1}).$$

One analogously justifies $\mu_{b|2} = \mu_{b2}$.

It remains to show $\mu_a \, R^w \, \mu_b$. To this end we justify that W is a weight function for (μ_a, R, μ_b). Obviously, W is a coupling of (μ_a, μ_b). Let $R'_1 \subseteq X_{a1} \times X_{b1}$ and $R'_2 \subseteq X_{a2} \times X_{b2}$ be Borel sets where $W_1(R'_1) = 1$, $W_2(R'_2) = 1$, $R'_1 \subseteq R_1$, and $R'_2 \subseteq R_2$. Define

$$R' = \iota_1^{-1}(R'_1) \cap \iota_2^{-1}(R'_2).$$

Then, R' is a Borel set in $X_1 \times X_2$ and $R' \subseteq R$. Moreover, $W(\iota_1^{-1}(R'_1)) = W_1(R'_1) = 1$ and $W(\iota_2^{-1}(R'_2)) = W_2(R'_2) = 1$ and therefore

$$W(R') = W(\iota_1^{-1}(R'_1) \cap \iota_2^{-1}(R'_2)) = 1.$$

We conclude that W is a weight function for (μ_a, R, μ_b) and hence $\mu_a \, R^w \, \mu_b$. The remaining claim of the theorem is proven accordingly. ◀

Weakly adequate span connections.

▶ **Definition 73.** We call a span connection $(\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ *weakly adequate* if for all $x_{a1} \, R_1 \, x_{b1}$ and $x_{a2} \, R_2 \, x_{b2}$,

$$x_{a1} \, \mathcal{X}_a \, x_{a2} \quad \text{iff} \quad x_{b1} \, \mathcal{X}_b \, x_{b2}.$$

▶ **Proposition 74.** *Every adequate span connection is weakly adequate.*

Proof. The claim follows from Propositions 34 and 50. ◀

▶ **Theorem 75.** *Let $\mathcal{C} = (\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ be a span connection where $\mathcal{X}_a = (X_a, X_{a1}, X_{a2})$ and $\mathcal{X}_b = (X_b, X_{b1}, X_{b2})$. Assume that $\mathsf{Rel}(\mathcal{X}_a)$ and $\mathsf{Rel}(\mathcal{X}_b)$ are stable-iff-weight and R_1 and R_2 are strongly lr-total in $X_{a1} \times X_{b1}$ and $X_{a2} \times X_{b2}$, respectively. If $\mathcal{C}$ is weakly adequate, then $\mathcal{C}$ is adequate.*

Proof. Assume that $\mathcal{C}$ weakly adequate. For reasons of symmetry we argue that $\mathcal{C}$ is l-adequate only. R-adequacy is proven analogously. Let $\mu_{a1}\, R_1^w\, \mu_{b1}$ and $\mu_{a2}\, R_2^w\, \mu_{b2}$ be such that $\mu_{a1}\, \mathcal{X}_a^c\, \mu_{a2}$. Our task is to justify $\mu_{b1}\, \mathcal{X}_b^c\, \mu_{b2}$. Let $\langle M_{b1}, M_{b2}\rangle$ be a $\mathsf{Rel}(\mathcal{X}_b)$-stable pair, where M_{b1} and M_{b2} are Borel sets in X_{b1} and X_{b2}, respectively. The argument proceeds as follows: We construct a $\mathsf{Rel}(\mathcal{X}_a)$-stable pair $\langle M_{a1}, M_{a2}\rangle$, where M_{a1} and M_{a2} are Borel in X_{a1} and X_{a2}, respectively, such that

$$\langle M_{a1}, M_{b1}\rangle \text{ is } R_1\text{-stable} \quad \text{and}$$
$$\langle M_{a2}, M_{b2}\rangle \text{ is } R_2\text{-stable.}$$

Notice, if we prove this claim, then

$$\mu_{b1}(M_{b1}) = \mu_{a1}(M_{a1}) = \mu_{a2}(M_{a2}) = \mu_{b2}(M_{b2}),$$

that yiels $\mu_{b1}\, \mathsf{Rel}(\mathcal{X}_b)^s\, \mu_{b2}$ and also $\mu_{b1}\, \mathsf{Rel}(\mathcal{X}_b)^w\, \mu_{b2}$ using that $\mathsf{Rel}(\mathcal{X}_b)$ is stable-iff-weight. Corollary 60 thus yields $\mu_{b1}\, \mathcal{X}_b^c\, \mu_{b2}$.

We regard the remaining claim. Let $f_a\colon X_{a1} \to X_{b1}$, $f_b\colon X_{b1} \to X_{a1}$, $g_a\colon X_{a2} \to X_{b2}$ and $g_b\colon X_{b2} \to X_{a2}$ be Borel functions where for all $x_{a1} \in X_{a1}$, $x_{b1} \in X_{b1}$, $x_{a2} \in X_{a2}$, and $x_{b2} \in X_{b2}$,

$$x_{a1}\, R_1\, f_a(x_{a1}) \quad \text{and} \quad f_b(x_{b1})\, R_1\, x_{b1} \quad \text{and}$$
$$x_{a2}\, R_2\, g_a(x_{a2}) \quad \text{and} \quad g_b(x_{b2})\, R_2\, x_{b2}.$$

Notice, such functions indeed exist due to assumption that R_1 and R_2 are strongly rl-total in $X_{a1} \times X_{b1}$ and $X_{a2} \times X_{b2}$, respectively. The relation $\mathsf{Rel}(\mathcal{X}_b)$ is an quasi-equivalence in $R_{b1} \times R_{b2}$ (cf. Proposition 64). Therefore and as $\langle M_{b1}, M_{b2}\rangle$ is $\mathsf{Rel}(\mathcal{X}_b)$-stable, we obtain

$$M_{b1} = \bigcup_{x_{b2} \in M_{b2}} [x_{b2}]_{-,\mathsf{Rel}(\mathcal{X}_b)} \quad \text{and} \quad M_{b2} = \bigcup_{x_{b1} \in M_{b1}} [x_{b1}]_{\mathsf{Rel}(\mathcal{X}_b),-}.$$

applying Proposition 28. Define

$$M_{a1} = \bigcup_{x_{b2} \in M_{b2}} [g_b(x_{b2})]_{-,\mathsf{Rel}(\mathcal{X}_a)} \quad \text{and} \quad M_{a2} = \bigcup_{x_{b1} \in M_{b1}} [f_b(x_{b1})]_{\mathsf{Rel}(\mathcal{X}_a),-}.$$

We show that $\langle M_{a1}, M_{b1}\rangle$ is R_1-stable first. Let $\langle x_{a1}, x_{b1}\rangle \in R_1 \cap (M_{a1} \times X_{b1})$. According to the definition of M_{a1}, there is $x_{b2} \in M_{b2}$ such that $x_{a1}\, \mathsf{Rel}(\mathcal{X}_a)\, g_b(x_{b2})$. As $x_{a1}\, R_1\, x_{b1}$ and $g_b(x_{b2})\, R_2\, x_{b2}$, we have $x_{b1}\, \mathsf{Rel}(\mathcal{X}_b)\, x_{b2}$ using the fact that $\mathcal{C}$ is weakly adequate. Since $\langle M_{b1}, M_{b2}\rangle$ is $\mathsf{Rel}(\mathcal{X}_b)$-stable and $x_{b2} \in M_{b2}$, we hence obtain $x_{b1} \in M_{b1}$. This shows $R_1 \cap (M_{a1} \times X_{b1}) \subseteq R_1 \cap (X_{a1} \times M_{b1})$. We justify the reverse inclusion and assume $\langle x_{a1}, x_{b1}\rangle \in R_1 \cap (X_{a1} \times M_{b1})$. Considering the properties of M_{b1}, there is $x_{b2} \in M_{b2}$ such that $x_{b1}\, \mathsf{Rel}(\mathcal{X}_b)\, x_{b2}$. Using $x_{a1}\, R_1\, x_{b1}$ and $g_b(x_{b2})\, R_2\, x_{b2}$, weakly-adequacy of $\mathcal{C}$ yields $x_{a1}\, \mathsf{Rel}(\mathcal{X}_a)\, g_b(x_{b2})$. Hence, $x_{a1} \in M_{a1}$ by the definition of M_{a1}, which yields $R_1 \cap (X_{a1} \times M_{b1}) \subseteq R_1 \cap (M_{a1} \times X_{b1})$. It follows that $\langle M_{a1}, M_{b1}\rangle$ is R_1-stable. One analogously shows that $\langle M_{a2}, M_{b2}\rangle$ is R_2-stable.

We argue that $\langle M_{a1}, M_{a2}\rangle$ is $\mathsf{Rel}(\mathcal{X}_a)$-stable now. Let $\langle x_{a1}, x_{a2}\rangle \in \mathsf{Rel}(\mathcal{X}_a) \cap (M_{a1} \times X_{a2})$. The definition of M_{a1} justifies the existence of some $x_{b2} \in M_{b2}$ where $x_{a1}\, \mathsf{Rel}(\mathcal{X}_a)\, g_b(x_{b2})$. Since $g_b(x_{b2})\, R_2\, x_{b2}$ and as $\langle M_{a2}, M_{b2}\rangle$ is R_2-stable, we obtain $g_b(x_{b2}) \in M_{a2}$. By the definition of M_{a2} there exists $x_{b1} \in M_{b1}$ such that $f_b(x_{b1})\, \mathsf{Rel}(\mathcal{X}_a)\, g_b(x_{b2})$. As $x_{a1}\, \mathsf{Rel}(\mathcal{X}_a)\, g_b(x_{b2})$ and $x_{a1}\, \mathsf{Rel}(\mathcal{X}_a)\, x_{a2}$, the z-transitivity of $\mathsf{Rel}(\mathcal{X}_a)$ yields $f_b(x_{b1})\, \mathsf{Rel}(\mathcal{X}_a)\, x_{a2}$. It follows $x_{a2} \in M_{a2}$ and thus $\mathsf{Rel}(\mathcal{X}_a) \cap (M_{a1} \times X_{a2}) \subseteq \mathsf{Rel}(\mathcal{X}_a) \cap (X_{a1} \times M_{a2})$. The reverse inclusion is shown accordingly and thus, $\langle M_{a1}, M_{a2}\rangle$ is $\mathsf{Rel}(\mathcal{X}_a)$-stable.

To complete our argumentation it remains to show M_{a1} and M_{a2} are Borel in X_{a1} and X_{a2}, respectively. We concentrate on M_{a1} since M_{a2} can treated similarly. Define the function $h_1 \colon X_{a1} \to R_1$,

$$h_1(x_{a1}) = \langle x_{a1}, f_a(x_{a1}) \rangle$$

and equip the set R_1 with the induced sigma-algebra from $X_{a1} \times X_{b1}$. Then, h_1 is measurable: The function $\tilde{h}_1 \colon X_{a1} \to X_{a1} \times X_{b1}$, $\tilde{h}_1(x_{a1}) = \langle x_{a1}, f_a(x_{a1}) \rangle$ is Borel and for all $M_{12} \subseteq X_{a1} \times X_{b1}$ it holds $h_1^{-1}(R \cap M_{12}) = \tilde{h}_1^{-1}(M_{12})$. Using that $R_1 \cap (X_{a1} \times M_{b1})$ is measurable in R_1 and $\langle M_{a1}, M_{b1} \rangle$ is R_1-stable,

$$M_{a1} = h_1^{-1}(R_1 \cap (M_{a1} \times X_{b1})) = h_1^{-1}(R_1 \cap (X_{a1} \times M_{b1}))$$

is measurable in X_{a1}. This finally finishes our proof as already discussed before. ◀

▶ **Corollary 76.** *Let $\mathcal{C} = (\mathcal{X}_a, \mathcal{X}_b, R_1, R_2)$ be a span connection where $\mathcal{X}_a = (X_a, X_{a1}, X_{a2})$ and $\mathcal{X}_b = (X_b, X_{b1}, X_{b2})$ are countably separated and proper spans. Assume that R_1 and R_2 are strongly lr-total in $X_{a1} \times X_{b1}$ and $X_{a2} \times X_{b2}$, respectively. If $\mathcal{C}$ is weakly adequate, then $\mathcal{C}$ is adequate.*

Proof. Theorem 75 together with Proposition 65 yield the claim. ◀

Span connections for variable spans. Throughout this paragraph let $\mathcal{C} = (\mathcal{X}, \mathcal{X}, R_1, R_2)$ be a span connection where $\mathcal{X} = (X, X_1, X_2)$ is a variable span, i.e., $X_1 = \mathsf{Ev}(\mathsf{Var}_1)$, $X_2 = \mathsf{Ev}(\mathsf{Var}_2)$, and $X = \mathsf{Ev}(\mathsf{Var}_1 \cup \mathsf{Var}_2)$ for sets of variables Var_1 and Var_2. Abbreviate

$$\mathsf{LVar}_1 = \mathsf{Var}_1 \setminus \mathsf{Var}_2, \quad \mathsf{LVar}_2 = \mathsf{Var}_2 \setminus \mathsf{Var}_1, \quad \text{and} \quad \mathsf{SVar} = \mathsf{Var}_1 \cap \mathsf{Var}_2.$$

Remind, $\mathcal{X}$ is a proper span (cf. Example 43 (2)) and the tuple $(\mathsf{Ev}(\mathsf{SVar}), \kappa_1, \kappa_2)$ countably separates $\mathcal{X}$. Here, $\kappa_1(e_1) = e_{1|\mathsf{SVar}}$ and $\kappa_2(e_2) = e_{2|\mathsf{SVar}}$ for all $e_1 \in \mathsf{Ev}(\mathsf{Var}_1)$ and $e_2 \in \mathsf{Ev}(\mathsf{Var}_2)$ (cf. Example 63).

▶ **Definition 77.** We say that the span connection $\mathcal{C}$ *does not involve shared variables* if the following two statements hold:

1. There is $\tilde{R}_1 \subseteq \mathsf{Ev}(\mathsf{LVar}_1) \times \mathsf{Ev}(\mathsf{LVar}_1)$ such that

$$R_1 = \{\langle e_{a1}^L \uplus e^S, e_{b1}^L \uplus e^S \rangle \; ; \; e_{a1}^L \, \tilde{R}_1 \, e_{b1}^L \text{ and } e^S \in \mathsf{Ev}(\mathsf{SVar})\}.$$

2. There is $\tilde{R}_2 \subseteq \mathsf{Ev}(\mathsf{LVar}_2) \times \mathsf{Ev}(\mathsf{LVar}_2)$ such that

$$R_2 = \{\langle e_{a2}^L \uplus e^S, e_{b2}^L \uplus e^S \rangle \; ; \; e_{a2}^L \, \tilde{R}_2 \, e_{b2}^L \text{ and } e^S \in \mathsf{Ev}(\mathsf{SVar})\}.$$

▶ **Proposition 78.** *If $\mathcal{C}$ does not involve shared variables, then $\mathcal{C}$ is adequate and*

$$R_1 \subseteq \{\langle e_{a1}, e_{b1} \rangle \in \mathsf{Ev}(\mathsf{Var}_1) \times \mathsf{Ev}(\mathsf{Var}_1) \; ; \; e_{a1|\mathsf{SVar}} = e_{b1|\mathsf{SVar}}\} \quad \textit{and}$$
$$R_2 \subseteq \{\langle e_{a2}, e_{b2} \rangle \in \mathsf{Ev}(\mathsf{Var}_2) \times \mathsf{Ev}(\mathsf{Var}_2) \; ; \; e_{a2|\mathsf{SVar}} = e_{b2|\mathsf{SVar}}\}$$

as well as

$$R_1^w \subseteq \{\langle \eta_{a1}, \eta_{b1} \rangle \in \mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}_1)) \times \mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}_1)) \; ; \; \eta_{a1|\mathsf{SVar}} = \eta_{b1|\mathsf{SVar}}\} \quad \textit{and}$$
$$R_2^w \subseteq \{\langle \eta_{a2}, \eta_{b2} \rangle \in \mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}_2)) \times \mathsf{Prob}(\mathsf{Ev}(\mathsf{Var}_2)) \; ; \; \eta_{a2|\mathsf{SVar}} = \eta_{b2|\mathsf{SVar}}\}.$$

Proof. Suppose $\mathcal{C}$ does not involve shared variables. The claims concerning R_1 and R_2 is direct consequence of Definition 77. Thus, $\mathcal{C}$ is adequate due to Proposition 69. In what follows we regard the relations R_1^w and R_2^w. Notice,

$$\mathsf{Diag}_{\mathsf{Ev(SVar)}} = \{\langle e_{a1|\mathsf{SVar}}, e_{b1|\mathsf{SVar}}\rangle \ ; \ e_{a1}\, R_1\, e_{b1}\} \quad \text{and}$$
$$\mathsf{Diag}_{\mathsf{Ev(SVar)}} = \{\langle e_{a2|\mathsf{SVar}}, e_{b2|\mathsf{SVar}}\rangle \ ; \ e_{a2}\, R_2\, e_{b2}\}.$$

Let $\eta_{a1}\, R_1^w\, \eta_{b1}$. Lemma 38 (1) yields $\eta_{a1|\mathsf{SVar}}\, \mathsf{Diag}^w_{\mathsf{Ev(SVar)}}\, \eta_{b1|\mathsf{SVar}}$. Lemma 37 thus implies $\eta_{a1|\mathsf{SVar}} = \eta_{b1|\mathsf{SVar}}$. The statement concerning R_2^w can be treated analogously. ◀

▶ **Proposition 79.** *Assume $\mathcal{C}$ does not involve shared variables. For all $\eta_a, \eta_b \in \mathsf{Prob(Ev(Var))}$ where $\eta_{a|\mathsf{SVar}} = \eta_{b|\mathsf{SVar}}$ the following statements hold:*

1. *For all $e_{a1}\, R_1\, e_{b1}$ it holds $\eta_{a|\mathsf{Var}_1}\, R_1^w\, \eta_{b|\mathsf{Var}_1}$ if*

$$\eta_{a|\mathsf{LVar}_1} = \mathsf{Dirac}[e_{a1|\mathsf{LVar}_1}] \quad \textit{and} \quad \eta_{b|\mathsf{LVar}_1} = \mathsf{Dirac}[e_{b1|\mathsf{LVar}_1}].$$

2. *For all $e_{a2}\, R_2\, e_{b2}$ it holds $\eta_{a|\mathsf{Var}_2}\, R_2^w\, \eta_{b|\mathsf{Var}_2}$ if*

$$\eta_{a|\mathsf{LVar}_2} = \mathsf{Dirac}[e_{a2|\mathsf{LVar}_2}] \quad \textit{and} \quad \eta_{b|\mathsf{LVar}_2} = \mathsf{Dirac}[e_{b2|\mathsf{LVar}_2}].$$

Proof. For reasons of symmetry we only show (1) since (2) can be treated analogously. Let $\eta_a, \eta_b \in \mathsf{Prob(Ev(Var))}$ and $\eta^S \in \mathsf{Prob(Ev(SVar))}$ be such that $\eta_{a|\mathsf{SVar}} = \eta^S = \eta_{b|\mathsf{SVar}}$. Moreover, let $e_{a1}\, R_1\, e_{b1}$. Abbreviate $e_{a1|\mathsf{LVar}_1} = e^L_{a1}$ and $e_{b1|\mathsf{LVar}_1} = e^L_{b1}$. Assume $\eta_{a|\mathsf{LVar}_1} = \mathsf{Dirac}[e^L_{a1}]$ and $\eta_{b|\mathsf{LVar}_1} = \mathsf{Dirac}[e^L_{b1}]$. Our task is to show $\eta_{a|\mathsf{Var}_1}\, R_1^w\, \eta_{b|\mathsf{Var}_1}$.

Define $f\colon \mathsf{Ev(SVar)} \to \mathsf{Ev(Var_1)} \times \mathsf{Ev(Var_1)}$,

$$f(e^S) = \langle e^L_{a1} \uplus e^S, e^L_{b1} \uplus e^S\rangle.$$

Then, f is a Borel function and thus we can safely define $W_1 \in \mathsf{Prob(Ev(Var_1) \times Ev(Var_1))}$ by

$$W_1 = f_\sharp(\eta^S).$$

We argue that W_1 is a weight function for $(\eta_{a|\mathsf{Var}_1}, R_1, \eta_{b|\mathsf{Var}_1})$. For that purpose introduce

$$R_1' = \{\langle e^L_{a1} \uplus e^S, e^L_{b1} \uplus e^S\rangle \ ; \ e^S \in \mathsf{Ev(SVar)}\}.$$

It is easy to see that R_1' is a Borel set in $\mathsf{Ev(Var_1) \times Ev(Var_1)}$. Moreover, $R_1' \subseteq R_1$ using $e_{a1}\, R_1\, e_{b1}$ and Proposition 78. Since $f^{-1}(R_1') = \mathsf{Ev(SVar)}$,

$$W_1(R_1') = \eta^S(f^{-1}(R_1')) = \eta^S(\mathsf{Ev(SVar)}) = 1.$$

It remains to show that W is a coupling of $(\eta_{a|\mathsf{Var}_1}, \eta_{b|\mathsf{Var}_1})$. For every $V \subseteq \mathsf{Var}_1$ define the Borel function $g_V\colon \mathsf{Ev(Var_1)} \to \mathsf{Ev}(V)$, $g_V(e_1) = e_{1|V}$. Let $M^L \subseteq \mathsf{Ev(LVar_1)}$ and $M^S \subseteq \mathsf{Ev(SVar_1)}$ be Borel sets and define the Borel set $M \subseteq \mathsf{Ev(Var_1)}$ by $M' = g^{-1}_{\mathsf{LVar}_1}(M^L) \cap g^{-1}_{\mathsf{SVar}}(M^S)$. Since $\eta_{a|\mathsf{Var}_1}$ is a coupling of $(\eta^S, \mathsf{Dirac}[e^L_{a1}])$, Proposition 24 yields $\eta_{a|\mathsf{Var}_1}(M') = \eta^S(M^S) \cdot \mathsf{Dirac}[e^L_{a1}](M^L)$ and thus

$$\begin{aligned}
& W_1(M' \times \mathsf{Ev(Var_1)}) \\
= \ & \eta^S(f^{-1}(M' \times \mathsf{Ev(Var_1)})) \\
= \ & \eta^S(M^S) \cdot \mathsf{Dirac}[e^L_{a1}](M^L) \\
= \ & \eta_{a|\mathsf{Var}_1}(M').
\end{aligned}$$

Observe, the sigma-algebra on $\mathsf{Ev(Var)}$ is generated by sets of the form $g^{-1}_{\mathsf{LVar}_1}(M^L) \cap g^{-1}_{\mathsf{SVar}}(M^S)$ where $M^L \subseteq \mathsf{Ev(LVar_1)}$ and $M^S \subseteq \mathsf{Ev(SVar_1)}$ are Borel sets. Carathéodory's measure extension theorem thus yields $W_1(M \times \mathsf{Ev(Var_1)}) = \eta_{a|\mathsf{Var}_1}(M)$ for all Borel sets $M \subseteq \mathsf{Ev(Var)}$. One similar shows $W_1(\mathsf{Ev(Var_1)} \times M) = \eta_{b|\mathsf{Var}_1}(M)$ for every Borel set $M \subseteq \mathsf{Ev(Var)}$, that completes our proof. ◀

C Stochastic transition systems

C.1 Composition

Local constraints.

▶ **Definition 80.** Let $\mathcal{S} = (S, S_1, S_2)$ be a span. A $\mathcal{S}_2$*-local constraint* is a relation $LC_2 \subseteq S \times \mathsf{Prob}(S)$, that enjoys the conditions below:

1. For all $s \in S$ and $\mu \in \mathsf{Prob}(S)$, if $\mu_{|2} = \mathsf{Dirac}[s_{|2}]$, then $s\,LC_2\,\mu$.
2. For all $s\,LC_2\,\mu$ and $\mu' \in \mathsf{Prob}(S)$, if $\mu_{|1} = \mu'_{|1}$ and $\mu_{|2} = \mu'_{|2}$, then $s\,LC_2\,\mu'$.
3. For all $s\,LC_2\,\mu$, if $\mu_{|1}\,\mathcal{S}^c\,\mathsf{Dirac}[s_{|2}]$, then μ is a $\mathcal{S}$-coupling of $(\mu_{|1}, \mathsf{Dirac}[s_{|2}])$.

Similar, a $\mathcal{S}_1$*-local constraint* is a relation $LC_1 \subseteq S \times \mathsf{Prob}(S)$, such that:

4. For all $s \in S$ and $\mu \in \mathsf{Prob}(S)$, if $\mu_{|1} = \mathsf{Dirac}[s_{|1}]$, then $s\,LC_1\,\mu$.
5. For all $s\,LC_1\,\mu$ and $\mu' \in \mathsf{Prob}(S)$, if $\mu_{|1} = \mu'_{|1}$ and $\mu_{|2} = \mu'_{|2}$, then $s\,LC_1\,\mu'$.
6. For all $s\,LC_1\,\mu$, if $\mathsf{Dirac}[s_{|1}]\,\mathcal{S}^c\,\mu_{|2}$, then μ is a $\mathcal{S}$-coupling of $(\mathsf{Dirac}[s_{|1}], \mu_{|2})$.

A pair (LC_1, LC_2) consisting of a LC_1-local constraint LC_1 and a LC_2-local constraint LC_2 is called $\mathcal{S}$*-agreement*.

▶ **Example 81.** Let $\mathcal{S} = (S, S_1, S_2)$ be a Cartesian span, i.e., $S = S_1 \times S_2$. Then,

$$LC_2 = \{\langle\langle s_1, s_2\rangle, \mu_1 \otimes \mathsf{Dirac}[s_2]\rangle \ ; \ s_1 \in S_1,\ s_2 \in S_2, \text{ and } \mu_1 \in \mathsf{Prob}(S_1)\}$$

constitute a $\mathcal{S}_2$-local constraint Moreover, there exists no $\mathcal{S}_2$-local contraint different from LC_2. Let us have a closer look and justify these claims. Requirement (1) in Definition 80 yield $\langle\langle s_1, s_2\rangle, \mu_1 \otimes \mathsf{Dirac}[s_2]\rangle \in LC_2$ for all $s_1 \in S_1$, $s_2 \in S_2$, and $\mu_1 \in \mathsf{Prob}(S_1)$. Conversely, let $\langle\langle s_1, s_2\rangle, \mu\rangle \in LC_2$. Since $\mathcal{S}$ is supposed to be a Cartesian span we have $\mathsf{Rel}(\mathcal{S}) = S_1 \times S_2$ and thus, $\mu_{|1}\,\mathsf{Rel}(\mathcal{S})^w\,\mathsf{Dirac}[s_{|2}]$. Hence, $\mu_{|1}\,\mathcal{S}^c\,\mathsf{Dirac}[s_{|2}]$ (cf. Corollary 60). Therefore, Definition 80 (3) yields that μ is a $\mathcal{S}$-coupling of $(\mu_{|1}, \mathsf{Dirac}[s_2])$. We obtain $\mu = \mu_{|1} \otimes \mathsf{Dirac}[s_2]$ by Proposition 24, i.e., $\mu = \mu_1 \otimes \mathsf{Dirac}[s_2]$ for some $\mu_1 \in \mathsf{Prob}(S_1)$. From this all the claims concerning LC_2 follow.

▶ **Example 82.** Let $\mathcal{S} = (S, S_1, S_2)$ be an identity span, i.e., $S = S_1 = S_2$. Then,

$$LC_2 = S \times \mathsf{Prob}(S)$$

is a $\mathcal{S}_2$-local contraint. Obviously, LC_2 satisfies the requirements (1) and (2) in Definition 80. To justify Definition 80 (3) suppose $\langle s, \mu\rangle \in LC_2$ where $\mu_{|1}\,\mathcal{S}^c\,\mathsf{Dirac}[s_{|2}]$. Since $\mathcal{S}$ is an identiy span, $\mu_{|1} = \mu$ and $\mathsf{Dirac}[s_{|2}] = \mathsf{Dirac}[s]$. Hence, $\mu\,\mathcal{S}^c\,\mathsf{Dirac}[s]$ and also $\mu\,\mathsf{Rel}(\mathcal{S})^w\,\mathsf{Dirac}[s]$ (cf. Proposition 52). We obtain $\mu = \mathsf{Dirac}[s]$ using $\mathsf{Rel}(\mathcal{S}) = \mathsf{Diag}_S$ and Lemma 37. Of course, μ is thus a $\mathcal{S}$-coupling of $(\mu_{|1}, \mathsf{Dirac}[s_{|2}])$. This finally shows that LC_2 is indeed a $\mathcal{S}_2$-local constraint.

▶ **Example 83.** Let $\mathcal{S} = (S, S_1, S_2)$ be a variable span, i.e., $S_1 = \mathsf{Ev}(\mathsf{Var}_1)$, $S_2 = \mathsf{Ev}(\mathsf{Var}_2)$, and $S = \mathsf{Ev}(\mathsf{Var}_1 \cup \mathsf{Var}_2)$ for sets of variables Var_1 and Var_2. Abbreviate $\mathsf{LVar}_1 = \mathsf{Var}_1 \setminus \mathsf{Var}_2$, $\mathsf{LVar}_2 = \mathsf{Var}_2 \setminus \mathsf{Var}_1$, and $\mathsf{SVar} = \mathsf{Var}_1 \cap \mathsf{Var}_2$. Then,

$$\begin{aligned}
LC'_2 &= \{\langle e, \eta\rangle \in S \times \mathsf{Prob}(S) \ ; \ \eta_{|\mathsf{SVar}} = \mathsf{Dirac}[e_{|\mathsf{SVar}}] \text{ implies } \eta_{|\mathsf{Var}_2} = \mathsf{Dirac}[e_{|\mathsf{Var}_2}]\},\\
LC''_2 &= \{\langle e, \eta\rangle \in S \times \mathsf{Prob}(S) \ ; \ \eta_{|\mathsf{LVar}_2} = \mathsf{Dirac}[e_{|\mathsf{LVar}_2}]\}, \quad \text{and}\\
LC'''_2 &= \{\langle e, \eta\rangle \in S \times \mathsf{Prob}(S) \ ; \ \eta_{|\mathsf{Var}_2} = \mathsf{Dirac}[e_{|\mathsf{Var}_2}]\}
\end{aligned}$$

are $\mathcal{S}_2$-local constraints such that $LC_2' \supseteq LC_2'' \supseteq LC_2'''$. Moreover, for every $\mathcal{S}_2$-local constraint LC_2 we have $LC_2' \supseteq LC_2$. We consider the last claim first and suppose a $\mathcal{S}_2$-local constraint LC_2. Let $\langle e, \eta \rangle \in LC_2$ where $\eta_{|\mathsf{SVar}} = \mathsf{Dirac}[e_{|\mathsf{SVar}}]$. In order to show $\langle e, \eta \rangle \in LC_2'$ we have to argue $\eta_{|\mathsf{Var}_2} = \mathsf{Dirac}[e_{|\mathsf{Var}_2}]$. Since $\eta_{|\mathsf{SVar}} = \mathsf{Dirac}[e_{|\mathsf{SVar}}]$ we obtain $\eta_{|\mathsf{Var}_1}\, \mathcal{S}^c\, \mathsf{Dirac}[e_{|\mathsf{Var}_2}]$ (cf. Proposition 65). Definition 80 (3) thus ensures that η is a $\mathcal{S}$-coupling of $(\eta_{|\mathsf{Var}_1}, \mathsf{Dirac}[e_{|\mathsf{Var}_2}])$. It follows $\eta_{|\mathsf{Var}_2} = \mathsf{Dirac}[e_{|\mathsf{Var}_2}]$, which justifies $LC_2' \supseteq LC$.

Span composition. Throughout this paragraph let $\mathcal{T}_1 = (S_1, \Gamma_1, \rightarrow_1)$ and $\mathcal{T}_1 = (S_2, \Gamma_2, \rightarrow_2)$ be STSs where S_1 and S_2 are supposed to be Polish spaces. Let $\mathcal{S} = (S, S_1, S_2)$ be a span, $\mathsf{Sync} \subseteq \Gamma_1 \cap \Gamma_2$ be a set of synchronization labels, and $\mathcal{G} = (LC_1, LC_2)$ an $\mathcal{S}$-agreement.

▶ **Definition 84.** We define the STS

$$\mathcal{T}_1 \parallel_{\mathcal{S},\mathcal{G},\mathsf{Sync}} \mathcal{T}_2 = (S, \Gamma_1 \cup \Gamma_2, \rightarrow),$$

where for all $s \in S$, $\gamma \in \Gamma$, and $\mu \in \mathsf{Prob}(S)$ it holds $s \rightarrow^{\gamma} \mu$ iff one of the conditions below hold:

1. $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$ and $s_{|1} \rightarrow_1^{\gamma} \mu_{|1}$ and $s\, LC_2\, \mu$.
2. $\gamma \in \Gamma_2 \setminus \mathsf{Sync}$ and $s\, LC_1\, \mu$ and $s_{|2} \rightarrow_2^{\gamma} \mu_{|2}$.
3. $\gamma \in \mathsf{Sync}$ and $s_{|1} \rightarrow_1^{\gamma} \mu_{|1}$ and $s_{|2} \rightarrow_2^{\gamma} \mu_{|2}$.

In case $\mathcal{S}$ is a Cartesian span, i.e., $S = S_1 \times S_2$, the $(\mathcal{T}_1, \mathcal{T}_2)$-agreement is uniquely determined (cf. Example 81) and hence we simply write $\mathcal{T}_1 \parallel_{\times,\mathsf{Sync}} \mathcal{T}_2$ instead of $\mathcal{T}_1 \parallel_{\mathcal{S},\mathcal{G},\mathsf{Sync}} \mathcal{T}_2$.

▶ **Proposition 85.** *Assume $\mathcal{S}$ is a Cartesian span, i.e., $S = S_1 \times S_2$. For all $s_1 \in S_1$, $s_2 \in S_2$, $\gamma \in \Gamma_1 \cup \Gamma_2$, and $\mu \in \mathsf{Prob}(S)$, we have $\langle s_1, s_1 \rangle \hookrightarrow^{\gamma} \mu$ in $\mathcal{T}_1 \parallel_{\times,H} \mathcal{T}_2$ iff one of the following statements hold:*

1. $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$ *and* $s_1 \rightarrow_1^{\gamma} \mu_{|1}$ *and* $\mu_2 = \mathsf{Dirac}[s_2]$.
2. $\gamma \in \Gamma_2 \setminus \mathsf{Sync}$ *and* $s_2 \rightarrow_1^{\gamma} \mu_{|2}$ *and* $\mu_1 = \mathsf{Dirac}[s_1]$.
3. $\gamma \in \mathsf{Sync}$ *and* $s_1 \rightarrow_1^{\gamma} \mu_{|1}$ *and* $s_2 \rightarrow_2^{\gamma} \mu_{|2}$.

Proof. The claim is a direct consequence of Definition 84 and Example 81. ◀

▶ **Remark 86.** We do not use any properties of Polish spaces within the definition of our composition operator (cf. Definition 84). At the beginning of the paragraph we require that S_1 and S_2 are Polish spaces since spans involve Polish spaces. However, one could introduce a relaxed notion of a span (e.g., a tuple $(X, X_1, X_2, \iota_1, \iota_2)$ consisting of measurable spaces X, X_1, and X_2 and measurable functions $\iota_1 \colon X \rightarrow X_1$ and $\iota_2 \colon X \rightarrow X_2$) and define the composition operator with respect to such a span.

Span versus independent composition. Recall the standard composition for STSs [18]: Given STSs $\mathcal{T}_1 = (S_1, \Gamma_1, \rightarrow_1)$ and $\mathcal{T}_2 = (S_2, \Gamma_2, \rightarrow_2)$ and a set of synchronization labels $\mathsf{Sync} \subseteq \Gamma_1 \cap \Gamma_2$, the STS

$$\mathcal{T}_1 \parallel_{\mathsf{Sync}}^{\otimes} \mathcal{T}_2 = (S_1 \times S_2, \Gamma_1 \cup \Gamma_2, \rightarrow)$$

with $\langle s_1, s_2 \rangle \rightarrow^{\gamma} \mu_1 \otimes \mu_2$ iff the following holds:

- If $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$, then $s_1 \rightarrow^{\gamma} \mu_1$ and $\mu_2 = \mathsf{Dirac}[s_2]$.
- If $\gamma \in \Gamma_2 \setminus \mathsf{Sync}$, then $\mu_1 = \mathsf{Dirac}[s_1]$ and $s_2 \rightarrow^{\gamma} \mu_2$.
- If $\gamma \in \mathsf{Sync}$, then $s_1 \rightarrow^{\gamma} \mu_1$ and $s_2 \rightarrow^{\gamma} \mu_2$.

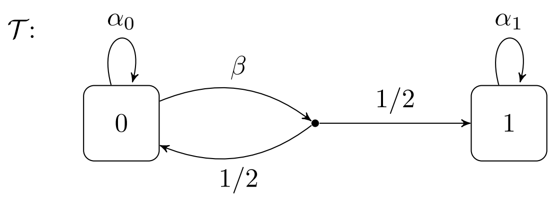


Figure 1 STS $\mathcal{T} = (S, \Gamma, \rightarrow)$ where $S = \{0, 1\}$ and $\Gamma = \{\alpha_0, \alpha_1, \beta\}$.

Consider the STS $\mathcal{T} = (S, \Gamma, \rightarrow)$ illustrated in Figure 1 where $S = \{0, 1\}$ and $\Gamma = \{\alpha_0, \alpha_1, \beta\}$. Assume $\mu \in \mathsf{Prob}(S)$ is given by $\mu(\{0\}) = \mu(\{1\}) = 1/2$. Notice, $0 \rightarrow^{\beta} \mu$. Abbreviate

$$\mathcal{T}_a = \mathcal{T} \parallel_{\times,\Gamma} \mathcal{T} \quad \text{and} \quad \mathcal{T}_b = \mathcal{T} \parallel_{\Gamma}^{\otimes} \mathcal{T} \quad \text{and} \quad \hat{S} = S \times S.$$

We aim to show $\mathcal{T}_a$ and $\mathcal{T}_b$ are not bisimilar. Towards a contradiction assume that $\mathcal{T}_a$ and $\mathcal{T}_b$ are bisimilar. Suppoe $R \subseteq \hat{S} \times \hat{S}$ is the coarsest bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$, i.e., $R' \subseteq R$ for every bisimulation R' for $(\mathcal{T}_a, \mathcal{T}_b)$. Remind, the union of arbitrary many bisimulations for $(\mathcal{T}_a, \mathcal{T}_b)$ yields again a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$ and thus R equals the union of all bisimulations for $(\mathcal{T}_a, \mathcal{T}_b)$. As $\mathcal{T}_a$ and $\mathcal{T}_b$ are supposed to be bisimilar, there is at least one bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$. Given $s_a \in S_a$ and $s_b \in S_b$, we use $s_a s_b$ as a shorthand notation for $\langle s_a, s_b \rangle$ in the following.

Since 0 is the only state in $\mathcal{T}$ with an outgoing α_0-transition and α_0 is a synchronization label, it necessarily holds $00\,R\,00$ using the fact that R is lr-total in $\hat{S} \times \hat{S}$. For similar reasons it holds $11\,R\,11$ and moreover,

$$\{\langle 00, 01\rangle, \langle 00, 10\rangle, \langle 00, 11\rangle, \langle 01, 00\rangle, \langle 10, 00\rangle, \langle 11, 00\rangle\} \cap R = \varnothing \text{ and}$$
$$\{\langle 11, 00\rangle, \langle 11, 01\rangle, \langle 11, 10\rangle, \langle 00, 11\rangle, \langle 01, 11\rangle, \langle 11, 10\rangle\} \cap R = \varnothing.$$

As R is the coarsest bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$, it follows

$$01\,R\,01, \quad 01\,R\,10, \quad 10\,R\,01, \quad \text{and} \quad 10\,R\,10.$$

Putting things together,

$$R = \{\langle 00, 00\rangle, \langle 11, 11\rangle, \langle 01, 01\rangle, \langle 01, 10\rangle, \langle 10, 01\rangle, \langle 10, 10\rangle\}.$$

Consider the transition $00 \hookrightarrow^{\beta} \hat{\mu}$ in $\mathcal{T}_a$ where $\hat{\mu} \in \mathsf{Prob}(\hat{S})$ is given by $\hat{\mu}(\{00\}) = \hat{\mu}(\{11\}) = 1/2$ (notice $\hat{\mu}$ is indeed a coupling of (μ, μ)). Since R is a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$ where $00\,R\,00$ and $00 \hookrightarrow^{\beta} \mu \otimes \mu$ is the only β-transition in $\mathcal{T}_b$ outgoing from 00, it holds $\hat{\mu}\,R^w\,(\mu \otimes \mu)$. Thus, $\hat{\mu}\,R^s\,(\mu \otimes \mu)$ by Proposition 35. The pair $\langle \{00\}, \{00\} \rangle$ is R-stable since

$$R \cap (\{00\} \times \hat{S}) = \{\langle 00, 00 \rangle\} = R \cap (\hat{S} \times \{00\}).$$

But,

$$\hat{\mu}(\{00\}) = 1/2 \neq 1/4 = 1/2 \cdot 1/2 = \mu(\{0\}) \cdot \mu(\{0\}) = \mu \otimes \mu(\{00\}),$$

that yields a contradiction. Therefore, $\mathcal{T}_a$ and $\mathcal{T}_b$ are not bisimilar.

Parallel composition of NLMPs. Labeled Markov processes and non-deterministic labeled Markov processes (NLMPs) [25, 23, 55] are elegant formalisms that allow among others for an rich theory on bisimulation and its logical characterization. An NLMP can be seen as an STS that enjoys requirements concerning measurability issues. Unfortunately, NLMPs are a priori not appropriate for our purposes as the class of NLMPs is not closed under composition. Let us give some details and assume an STS $\mathcal{T} = (S, \Gamma, \rightarrow)$. For every $s \in S$ and $\gamma \in \Gamma$ define

$$\mathsf{Enabled}_{\mathcal{T}}(s,\gamma) = \{\mu \in \mathsf{Prob}(S) \; ; \; s \rightarrow^{\gamma} \mu\}.$$

The STS $\mathcal{T}$ is called *NLMP-like*, if for all $\gamma \in \Gamma$ the following two conditions are fulfilled:

- $\{\mu \in \mathsf{Prob}(S) \; ; \; s \rightarrow^{\gamma} \mu\}$ is Borel in $\mathsf{Prob}(S)$ for all $s \in S$.
- $\{s \in S \; ; \; \mathsf{Enabled}_{\mathcal{T}}(s,\gamma) \cap P \neq \varnothing\}$ is Borel in S for all Borel sets $P \subseteq \mathsf{Prob}(S)$.

We now given two NLMP-like STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ such that their composition is not NLMP-like when synchronizing on all common labels. Consider the STSs $\mathcal{T}_1 = (\mathbb{R}, \Gamma, \rightarrow_1)$ and $\mathcal{T}_2 = (\mathbb{R}, \Gamma, \rightarrow_2)$ with $\Gamma = \{\gamma\}$ and

$$\rightarrow_1 = \{\langle r_1, \gamma, \mathsf{Dirac}[r_1]\rangle \; ; \; r_1 \in \mathbb{R}\} \quad \text{and}$$
$$\rightarrow_2 = \{\langle r_2, \gamma, \mathsf{Dirac}[r_2']\rangle \; ; \; r_2, r_2' \in \mathbb{R}\}.$$

Abbreviate $\mathcal{T} = \mathcal{T}_1 \parallel_{\times,\Gamma} \mathcal{T}_2$. It is not hard to see that $\mathcal{T}_1$ and $\mathcal{T}_2$ are NLMP-like (cf. Remark 14). However, $\mathcal{T}$ is not NLMP-like. In what follows we give a formal argument. Let $M \subseteq \mathbb{R} \times \mathbb{R}$ be a Borel set such that M_1 is not Borel in $\mathbb{R}$, where $M_1 = \{r_1 \in \mathbb{R} \; ; \;$ there is $r_2 \in \mathbb{R}$ where $\langle r_1, r_2\rangle \in M\}$ (cf. Remark 20). Define $\mathsf{Dirac}_M = \{\mathsf{Dirac}[r] \; ; \; r \in M\}$. Then, Dirac_M is Borel in $\mathsf{Prob}(\mathbb{R} \times \mathbb{R})$ (cf. Remark 14), but

$$\{\langle r_1, r_2\rangle \in \mathbb{R} \times \mathbb{R} \; ; \; \mathsf{Enabled}_{\mathcal{T}}(\langle r_1, r_2\rangle, \gamma) \cap \mathsf{Dirac}_M \neq \varnothing\} = M_1 \times \mathbb{R}$$

is not Borel $\mathbb{R} \times \mathbb{R}$ (the function $f \colon \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R}$, $f(r) = \langle r, r\rangle$ is Borel and if M_1 is Borel in $\mathbb{R}$, then $f^{-1}(M_1 \times \mathbb{R}) = M_1$ is Borel in $\mathbb{R}$, which is not case). The STS $\mathcal{T}$ is therefore not NLMP-like. Notice, this has nothing to do with our notion of composition since $\mathcal{T} = \mathcal{T}_1 \parallel_{\Gamma}^{\otimes} \mathcal{T}_2$ (cf. Proposition 24). Moreover, the present STS $\mathcal{T}_1$ and $\mathcal{T}_2$ do not even incorporate non-trival probability measures and can be seen as non-probabilistic models.

C.2 Congruence

Let $\mathcal{T}_{a1} = (S_{a1}, \Gamma_1, \rightarrow_{a1})$, $\mathcal{T}_{a2} = (S_{a2}, \Gamma_2, \rightarrow_{a2})$, $\mathcal{T}_{b1} = (S_{b1}, \Gamma_1, \rightarrow_{b1})$, and $\mathcal{T}_{b2} = (S_{b2}, \Gamma_2, \rightarrow_{b2})$ be STSs such that $\mathcal{T}_{a1} \sim \mathcal{T}_{b1}$ and $\mathcal{T}_{a2} \sim \mathcal{T}_{b2}$. Notice, $\mathcal{T}_{a1}$ and $\mathcal{T}_{b1}$ have the same sets of labels Γ_1 and similar, $\mathcal{T}_{a2}$ and $\mathcal{T}_{b2}$ have the same sets of labels Γ_2. Define

$$\mathcal{T}_a = \mathcal{T}_{a1} \parallel_{\mathcal{S}_a, \mathcal{G}_a, \mathsf{Sync}} \mathcal{T}_{a2} \quad \text{and} \quad \mathcal{T}_b = \mathcal{T}_{b1} \parallel_{\mathcal{S}_b, \mathcal{G}_b, \mathsf{Sync}} \mathcal{T}_{b2},$$

where $\mathsf{Sync} \subseteq \Gamma_1 \cap \Gamma_2$, $\mathcal{S}_a = (S_a, S_{a1}, S_{a2})$ and $\mathcal{S}_b = (S_b, S_{b1}, S_{b2})$ are proper spans, $\mathcal{G}_a = (LC_{a1}, LC_{a2})$ is a $(\mathcal{T}_{a1}, \mathcal{T}_{a2})$-agreement, and $\mathcal{G}_b = (LC_{b1}, LC_{b2})$ is a $(\mathcal{T}_{b1}, \mathcal{T}_{b2})$-agreement. Assume relations $R_1 \subseteq S_{a1} \times S_{b1}$ and $R_2 \subseteq S_{a2} \times S_{b2}$. We consider the span connection $\mathcal{C} = (\mathcal{S}_a, \mathcal{S}_b, R_1, R_2)$. Reminding Definition 67,

$$R_1 \wedge_{\mathcal{C}} R_2 = \{\langle s_a, s_b\rangle \in S_a \times S_b \; ; \; s_{a|1} \, R_1 \, s_{b|1} \text{ and } s_{a|2} \, R_2 \, s_{b|2}\}.$$

We aim to show that $R_1 \wedge_{\mathcal{C}} R_2$ is a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$ if R_1 and R_2 are bisimulations for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$, respectively. For that purpose we need a notion to compare the agreements $\mathcal{G}_a$ and $\mathcal{G}_b$.

▶ **Definition 87.** LC_{a2} and LC_{b2} are *$\mathcal{C}$-bisimilar* whenever for all $s_a\,(R_1 \wedge_{\mathcal{C}} R_2)\,s_b$ the following statements hold:

1. For all $\mu_a \in \mathsf{Prob}(S_a)$ and $\mu_{b1} \in \mathsf{Prob}(S_{b1})$, if $s_a\,LC_{a2}\,\mu_a$ and $\mu_{a|1}\,R_1^w\,\mu_{b1}$, then there is $\mu_b \in \mathsf{Prob}(S_b)$ where
$$s_b\,LC_{b2}\,\mu_b, \quad \mu_{b|1} = \mu_{b1}, \quad \text{and} \quad \mu_{a|2}\,R_2^w\,\mu_{b|2}.$$
2. For all $\mu_b \in \mathsf{Prob}(S_b)$ and $\mu_{a1} \in \mathsf{Prob}(S_{a1})$, if $s_b\,LC_{b2}\,\mu_b$ and $\mu_{a1}\,R_1^w\,\mu_{b|1}$, then there is $\mu_a \in \mathsf{Prob}(S_a)$ where
$$s_a\,LC_{a2}\,\mu_a, \quad \mu_{a|1} = \mu_{a1}, \quad \text{and} \quad \mu_{a|2}\,R_2^w\,\mu_{b|2}.$$

Similar, LC_{a1} and LC_{b1} are *$\mathcal{C}$-bisimilar* whenever for all $s_a\,(R_1 \wedge_{\mathcal{C}} R_2)\,s_b$ the following statements hold:

3. For all $\mu_a \in \mathsf{Prob}(S_a)$ and $\mu_{b2} \in \mathsf{Prob}(S_{b2})$, if $s_a\,LC_{a1}\,\mu_a$ and $\mu_{a|2}\,R_2^w\,\mu_{b2}$, then there is $\mu_b \in \mathsf{Prob}(S_b)$ where
$$s_b\,LC_{b1}\,\mu_b, \quad \mu_{b|2} = \mu_{b2}, \quad \text{and} \quad \mu_{a|1}\,R_1^w\,\mu_{b|1}.$$
4. For all $\mu_b \in \mathsf{Prob}(S_b)$ and $\mu_{a2} \in \mathsf{Prob}(S_{a2})$, if $s_b\,LC_{b1}\,\mu_b$ and $\mu_{a2}\,R_2^w\,\mu_{b|2}$, then there is $\mu_a \in \mathsf{Prob}(S_a)$ where
$$s_a\,LC_{a1}\,\mu_a, \quad \mu_{a|2} = \mu_{a2}, \quad \text{and} \quad \mu_{a|1}\,R_1^w\,\mu_{b|1}.$$

$\mathcal{G}_a$ and $\mathcal{G}_b$ are *$\mathcal{C}$-bisimilar* if LC_{a2} and LC_{b2} as well as LC_{a1} and LC_{b1} are $\mathcal{C}$-bisimilar.

▶ **Proposition 88.** *If $\mathcal{S}_a$ and $\mathcal{S}_b$ are Cartesian spans, i.e., $S_a = S_{a1} \times S_{a2}$ and $S_b = S_{b1} \times S_{b2}$, then $\mathcal{G}_a$ and $\mathcal{G}_b$ are $\mathcal{C}$-bisimilar.*

Proof. Abbreviate $R = R_1 \wedge_{\mathcal{C}} R_2$. Assume $\mathcal{S}_a$ and $\mathcal{S}_b$ are Cartesian spans and $\mathcal{C}$ is adequate. For reasons of symmetry we only show that LC_{a2} and LC_{b2} are $\mathcal{C}$-bisimilar. According to Example 81,

$$LC_{a2} = \{\langle\langle s_{a1}, s_{a2}\rangle, \mu_{a1} \otimes \mathsf{Dirac}[s_{a2}]\rangle \;;\; s_{a1} \in S_{a1},\ s_{a2} \in S_{a2}, \text{ and } \mu_{a1} \in \mathsf{Prob}(S_{a1})\},$$
$$LC_{b2} = \{\langle\langle s_{b1}, s_{b2}\rangle, \mu_{b1} \otimes \mathsf{Dirac}[s_{b2}]\rangle \;;\; s_{b1} \in S_{b1},\ s_{b2} \in S_{b2}, \text{ and } \mu_{b1} \in \mathsf{Prob}(S_{b1})\}.$$

Suppose $s_a\,R\,s_b$ with $s_a = \langle s_{a1}, s_{a2}\rangle$ and $s_b = \langle s_{b1}, s_{b2}\rangle$. Let $\mu_a \in \mathsf{Prob}(S_a)$ and $\mu_{b2} \in \mathsf{Prob}(S_{b2})$ be such that $s_a\,LC_{a2}\,\mu_a$ and $\mu_{a|1}\,R_1^w\,\mu_{b1}$. Define $\mu_b \in \mathsf{Prob}(S_b)$ by

$$\mu_b = \mu_{b1} \otimes \mathsf{Dirac}[s_{b2}].$$

Obviously, $s_b\,LC_{b2}\,\mu_b$ and $\mu_{b|1} = \mu_{b1}$. It remains to show $\mu_{a|2}\,R_2^w\,\mu_{b|2}$. As $s_a\,R\,s_b$ we have $s_{a2}\,R_2\,s_{b2}$ and thus $\mathsf{Dirac}[s_{a2}]\,R_2^w\,\mathsf{Dirac}[s_{b2}]$ by Proposition 34. Since $s_a\,LC_{a2}\,\mu_a$ it follows $\mu_a = \mu_{a1} \otimes \mathsf{Dirac}[s_{a2}]$ and therefore $\mu_{a|2} = \mathsf{Dirac}[s_{a2}]$. As $\mu_{b|2} = \mathsf{Dirac}[s_{b2}]$ we hence obtain $\mu_{a|2}\,R_2^w\,\mu_{b|2}$. This finally yields requirement (1) in Definition 87. One similar shows requirement (2) in Definition 87. ◀

▶ **Proposition 89.** *Suppose $\mathcal{S} = \mathcal{S}_a = \mathcal{S}_b$ and $\mathcal{G} = \mathcal{G}_a = \mathcal{G}_b$, where $\mathcal{S}$ is a variable span and $\mathcal{G}$ is the standard $\mathcal{S}$-agreement. Moreover, assume $\mathcal{C}$ does not involve shared variables. Then, $\mathcal{G}_a$ and $\mathcal{G}_b$ are $\mathcal{C}$-bisimilar.*

Proof. For reasons of symmetry we only show that LC_{a2} and LC_{b2} are $\mathcal{C}$-bisimilar. Assume $\mathcal{S}$ is given by $\mathcal{S} = (S, S_1, S_2)$ where $S_1 = \mathsf{Ev}(\mathsf{Var}_1)$, $S_2 = \mathsf{Ev}(\mathsf{Var}_2)$, and $S = \mathsf{Ev}(\mathsf{Var}_1 \cup \mathsf{Var}_2)$ for sets of variables Var_1 and Var_2. Reminding Example 83,

$$LC_{a2} = \{\langle e_a, \eta_a\rangle \in S \times \mathsf{Prob}(S) \; ; \; \eta_{a|\mathsf{Var}_2\setminus\mathsf{Var}_1} = \mathsf{Dirac}[e_{a|\mathsf{Var}_2\setminus\mathsf{Var}_1}]\},$$
$$LC_{b2} = \{\langle e_b, \eta_b\rangle \in S \times \mathsf{Prob}(S) \; ; \; \eta_{b|\mathsf{Var}_2\setminus\mathsf{Var}_1} = \mathsf{Dirac}[e_{b|\mathsf{Var}_2\setminus\mathsf{Var}_1}]\}.$$

Abbreviate $R = R_1 \wedge_{\mathcal{C}} R_2$. Let $e_a \, R \, e_b$, $\eta_a \in \mathsf{Prob}(S)$, and $\eta_{b2} \in \mathsf{Prob}(S_2)$ be such that $e_a \, LC_{a2} \, \eta_a$ and $\eta_{a1} \, R_1^w \, \eta_{b1}$. Here, we abbreviate $\eta_{a1} = \eta_{a|\mathsf{Var}_1}$. Let $\mu_b \in \mathsf{Prob}(S)$ be the probability measure uniquely determined by

$$\eta_{b|\mathsf{Var}_1} = \eta_{b1} \quad \text{and} \quad \eta_{b|\mathsf{Var}_2\setminus\mathsf{Var}_1} = \mathsf{Dirac}[e_{b|\mathsf{Var}_2\setminus\mathsf{Var}_1}].$$

Obviously, $e_b \, LC_{b2} \, \eta_b$. It remains to justify $\eta_{a|\mathsf{Var}_2} \, R_2^w \, \eta_{b|\mathsf{Var}_2}$ in order to justify requirement (1) in Definition 87. Since $\eta_{a1} \, R_1^w \, \eta_{b1}$ and as $\mathcal{C}$ does not involve shared variables, Proposition 78 yields $\eta_{a|\mathsf{SVar}} = \eta_{a1|\mathsf{SVar}} = \eta_{b1|\mathsf{SVar}} = \eta_{b|\mathsf{SVar}}$. Since $e_a \, LC_{a2} \, \eta_a$, we have $\eta_{a|\mathsf{Var}_2\setminus\mathsf{Var}_1} = \mathsf{Dirac}[e_{a|\mathsf{Var}_2\setminus\mathsf{Var}_1}]$. Moreover, $e_{a|\mathsf{Var}_2} \, R_2 \, e_{b|\mathsf{Var}_2}$ and hence we are in the situation of Proposition 79 (2), that finally yields $\eta_{a|\mathsf{Var}_2} \, R_2^w \, \eta_{b|\mathsf{Var}_2}$.

Requirement (2) in Definition 87 is shown analogously and therefore, LC_{a2} and LC_{b2} are $\mathcal{C}$-bisimilar. ◀

▶ **Theorem 90.** *Assume R_1 and R_2 are bisimulations for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$, respectively, $\mathcal{C}$ is adequate, and $\mathcal{G}_a$ and $\mathcal{G}_b$ are $\mathcal{C}$-bisimilar. Then,*

$R_1 \wedge_{\mathcal{C}} R_2$ is a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$.

Proof. Abbreviate $R = R_1 \wedge_{\mathcal{C}} R_2$. For reasons of symmetry we only justify that every transition in $\mathcal{T}_a$ can be mimicked by a corresponding transition in $\mathcal{T}_b$. Let $s_a \, R \, s_b$, $\gamma \in \Gamma_1 \cup \Gamma_2$, and $\mu_a \in \mathsf{Prob}(S_a)$ and assume $s_a \to_a^\gamma \mu_a$. Notice, $s_{a|1} \, R_1 \, s_{b|1}$ and $s_{a|2} \, R_2 \, s_{b|2}$. We justify that there is a $\mu_b \in \mathsf{Prob}(S_b)$ such that $s_b \to_b^\gamma \mu_b$ and $\mu_a \, R \, \mu_b$.

We consider the case where $\gamma \in \mathsf{Sync}$ first. According to the definition of our composition operator (cf. Definition 84), it holds $s_{a|1} \to_{a1}^\gamma \mu_{a|1}$ and $s_{a|2} \to_{a2}^\gamma \mu_{a|2}$. Since R_1 and R_2 are bisimulations for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$, respectively, there are $\mu_{b1} \in \mathsf{Prob}(S_{b1})$ and $\mu_{b2} \in \mathsf{Prob}(S_{b2})$ such that

$$s_{b|1} \to_{b1}^\gamma \mu_{b1} \quad \text{and} \quad \mu_{a|1} \, R_1^w \, \mu_{b1} \quad \text{and}$$
$$s_{b|2} \to_{b2}^\gamma \mu_{b2} \quad \text{and} \quad \mu_{a|2} \, R_2^w \, \mu_{b2}.$$

Since $\mathcal{C}$ is adequate we can apply Theorem 72 and thus there exists a $\mathcal{S}_b$-coupling μ_b of (μ_{b1}, μ_{b2}) such that $\mu_a \, R^w \, \mu_b$. As it also holds $s_b \to_b^\gamma \mu_b$ by Definition 84, the first part of our proof is complete.

We attend the case where $\gamma \in \Gamma_1 \setminus \mathsf{Sync}$. Then, $s_{a|1} \to_{a1}^\gamma \mu_{a|1}$ and $s_a \, LC_{a2} \, \mu_a$. Using that R_1 is a bisimulation for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$, there exists $\mu_{b1} \in \mathsf{Prob}(S_{b1})$ such that

$$s_{b|1} \to_{b1}^\gamma \mu_{b1} \quad \text{and} \quad \mu_{a|1} \, R_1^w \, \mu_{b1}.$$

Since LC_{a2} and LC_{b2} are $\mathcal{C}$-bisimilar, there is $\mu_b' \in \mathsf{Prob}(S_b)$ such that

$$s_b \, LC_{b2} \, \mu_b', \quad \mu_{b|1}' = \mu_{b1}, \quad \text{and} \quad \mu_{a|2} \, R_2^w \, \mu_{b|2}'$$

(cf. Definition 87 (1)). Again, as $\mathcal{C}$ is adequate Theorem 72 yields the existence of a $\mathcal{S}_b$-coupling μ_b of $(\mu_{b1}, \mu_{b|2}')$ such that $\mu_a \, R^w \, \mu_b$. Since $\mu_{b|1}' = \mu_{b1} = \mu_{b|1}$ and $\mu_{b|2}' = \mu_{b|2}$, we have $s_b \, LC_{b2} \, \mu_b$ (cf. Definition 80 (2)). It follows $s_b \rightarrow_b^\gamma \mu_b$ and hence the next part of our proof that R is a bisimulation is complete.

The case where $\gamma \in \Gamma_2 \setminus \mathsf{Sync}$ can be treated similarly. However let us recall the argument. We have $s_{a|2} \rightarrow_{a2}^\gamma \mu_{a|2}$ and $s_a \, LC_{a1} \, \mu_a$. Since R_2 is a bisimulation for $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$, there is $\mu_{b2} \in \mathsf{Prob}(S_{b2})$ such that

$$s_{b|2} \rightarrow_{b2}^\gamma \mu_{b2} \quad \text{and} \quad \mu_{a|2} \, R_2^w \, \mu_{b2}.$$

Since LC_{a1} and LC_{b1} are $\mathcal{C}$-bisimilar, there is $\mu_b' \in \mathsf{Prob}(S_b)$ such that

$$s_b \, LC_{b1} \, \mu_b', \quad \mu_{b|2}' = \mu_{b2}, \quad \text{and} \quad \mu_{a|1} \, R_1^w \, \mu_{b|1}'$$

(cf. Definition 87 (3)). Using that $\mathcal{C}$ is adequate, Theorem 72 yields a $\mathcal{S}_b$-coupling μ_b of $(\mu_{b|1}', \mu_{b2})$ such that $\mu_a \, R^w \, \mu_b$. Since $\mu_{b|1}' = \mu_{b|1}$ and $\mu_{b|2}' = \mu_{b2} = \mu_{b|2}$, we have $s_b \, LC_{b1} \, \mu_b$ (cf. Definition 80 (5)). It follows $s_b \rightarrow_b^\gamma \mu_b$ and hence the next part of our proof that R is a bisimulation is complete. ◀

▶ **Corollary 91.** *Assume R_1 and R_2 are bisimulations for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$, respectively. If $\mathcal{S}_a$ and $\mathcal{S}_b$ are Cartesian spans, then*

$R_1 \wedge_{\mathcal{C}} R_2$ is a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$.

Proof. Propositions 68 and 88 together with Theorem 90 justify the claim. ◀

▶ **Corollary 92.** *Suppose $\mathcal{S} = \mathcal{S}_a = \mathcal{S}_b$ and $\mathcal{G} = \mathcal{G}_a = \mathcal{G}_b$, where $\mathcal{S}$ is a variable span and $\mathcal{G}$ is the standard $\mathcal{S}$-agreement. Moreover, assume $\mathcal{C}$ does not involve shared variables. Assume R_1 and R_2 are bisimulations for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$, respectively. Then,*

$R_1 \wedge_{\mathcal{C}} R_2$ is a bisimulation for $(\mathcal{T}_a, \mathcal{T}_b)$.

Proof. Notice, $\mathcal{C}$ is adequate due to Proposition 78 and $\mathcal{G}_a$ and $\mathcal{G}_b$ are $\mathcal{C}$-bisimilar according to Proposition 89. Hence, we are in the situation of Theorem 90, that yields the claim. ◀