# Stochastic transition systems: bisimulation, logic, and composition

## Dissertation

zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)

vorgelegt an der

Technischen Universität Dresden
Fakultät Informatik

eingereicht von

Daniel Gburek, M. Sc.
geboren am 24. Oktober 1990 in Gera

verteidigt am

27. März 2018

begutachtet von

Prof. Dr. Christel Baier
Technische Universität Dresden
Prof. Dr. Prakash Panangaden
McGill University Montréal

Dresden, im September 2018

**Dinge, die ich beim Schreiben der vorliegenden Dissertation gelernt habe:**

(1) Ausreden sind ein Zeichen von Faulheit.

(2) Kompromisslosigkeit ist eine Komponente von Erfolg.

(3) Kompromissbereitschaft ist eine Komponente von Erfolg.

(4) Es gibt keine Low-Hanging-Fruits oder Quick-Wins.

(5) Manchmal braucht es sehr viel Zeit, ein schönes Problem zu finden.

(6) Manchmal braucht es sehr viel Zeit, ein schönes Problem zu lösen.

(7) Kreative Lösungen benötigen Reifezeit.

(8) Details haben bei der Ideenfindung nichts zu suchen.

(9) Baue Argumentationen nur auf bereits komplett durchdachten Dingen auf.

(10) Leite E-Mails stets mit einem freundlichen Satz ein.

**Danke an alle, die mich während dieses Lernprozesses unterstützt haben!**

# Abstract

Cyber-physical systems and the Internet of Things raise various challenges concerning the modelling and analysis of large modular systems. Models for such systems typically require uncountable state and action spaces, samplings from continuous distributions, and non-deterministic choices over uncountable many alternatives. In this thesis we focus on a general modelling formalism for stochastic systems called stochastic transition system. We introduce a novel composition operator for stochastic transition systems that is based on couplings of probability measures. Couplings yield a declarative modelling paradigm appropriate for the formalisation of stochastic dependencies that are caused by the interaction of components. Congruence results for our operator with respect to standard notions for simulation and bisimulation are presented for which the challenge is to prove the existence of appropriate couplings. In this context a theory for stochastic transition systems concerning simulation, bisimulation, and trace-distribution relations is developed. We show that under generic Souslin conditions, the simulation preorder is a subset of trace-distribution inclusion and accordingly, bisimulation equivalence is finer than trace-distribution equivalence. We moreover establish characterisations of the simulation preorder and the bisimulation equivalence for a broad subclass of stochastic transition systems in terms of expressive action-based probabilistic logics and show that these characterisations are still maintained by small fragments of these logics, respectively. To treat associated measurability aspects, we rely on methods from descriptive set theory, properties of Souslin sets, as well as prominent measurable-selection principles.

# Contents

# 1 Introduction

Modelling, control, analysis, and verification of cyber-physical systems and the Internet of Things represent key challenges of the 21st century in the areas of computer science, mathematics, and engineering [EG16]. Mathematical challenges arise when the formal models include uncountable state and action spaces, discrete and continuous dynamics, samplings from continuous distributions, and non-deterministic choices over uncountable many alternatives. The article [HH15] provides an excellent overview of prominent classes of models. For instance, dynamics obtained by interaction of digital sensors with the continuous environment lead to hybrid systems [Hen96, BBM98, Pla08]. Communication and interaction of complex and heterogenous systems motivate distributed and compositional approaches of hybrid systems [AH97, AH99] as well as stochastic extension thereof [WSS97, LSV03, HHHK12, Pla12].

This thesis investigates a general model for modelling complex systems called stochastic transition system (Part I). We develop a theory on simulation, bisimulation, and trace-distribution relations (Part II), provide action-based probabilistic logics characterising the simulation preorder and the bisimulation equivalence (Part III), and present a composition operator not imposing that stochastic transition systems to be composed behave stochastically independent (Part IV). Discussions and concepts referring to stochastic systems with uncountable state spaces are inspired by the pioneering works on labelled Markov processes [BDEP97, Des99, DEP02, DGJP03, Pan09], the contributions on non-deterministic labelled Markov processes [DTW12, Wol12, DLM16], as well as the investigations for simple stochastic transition systems in [CSKN05, Cat05]. Considering the world of countable-state-space systems, this thesis is in the spirit of [SL94, Seg95, LSV07] studying probabilistic automata and of [D'A99, BD04, DK05] investigating stochastic automata.

**Part I: stochastic transition systems.** We focus on a general stochastic model, called stochastic transition system (STS), which consists of three basic ingredients: a state space, an action space, and a transition relation. Both the state and the action space of an STS may be uncountable, more precisely, are required to form Polish spaces. A state intuitively represents the configuration of the modelled system at a certain moment of its execution, e.g., the recent performance mode of a hardware component or the current value of variables referring to physical quantities such as temperature, velocity, or pressure. An action stands for an observable process activity, e.g., the switch of a performance mode affecting the energy consumption or the adjustment of system parameters influencing controlled physical quantities. Besides this, the sojourn time in a performance mode can be also reported by means of specific actions. The transition relation finally specifies how the execution of an action changes the system state and includes arbitrary (continuous) distributions over action-state pairs.

There are two views on operational systems concerning time: the linear and path-based [Pnu77] as well as the branching notion of time [EC82, Mil82, HJ94], see also [vG90, vG93, GSS95]. Simulation and bisimulation notions [vB76, HM80] relate states on the basis of their branching-time structure and provide prominent formalisms that have been adapted in many flavours. Whereas bisimulation equivalence assures the equivalence of the branching-time structure, simulation preorders aim to identify states where a state arises from the other by means of an abstraction. In the context of this thesis, we consider probabilistic simulation and bisimulation [LS91, Den15] where we focus on a purely action-based setting, i.e., the set of all actions represents the relevant set of basic atomic observables.

Intuitively, if the state $s_b$ of an STS simulates the state $s_a$, then $s_b$ can mimic the behaviour of $s_a$. Given a third state $s_c$ that in turn simulates $s_b$, one naturally expects that $s_c$ also simulates $s_a$. Thus, it is desired that simulation and also bisimulation yield transitive relations on the corresponding state space of the STS under consideration. The following theorem summarises the main contribution of Part I of this thesis:

**Theorem A.** *For every STS the following two statements hold:*

(1) *The simulation relation $\preceq$ is a preorder on the state space.*

(2) *The bisimulation relation $\simeq$ is an equivalence on the state space.*

*In particular, both relations $\preceq$ and $\simeq$ are transitive.* ⌟

The challenging part of the proof for this theorem is to justify the transitivity of the induced relations of simulation and bisimulation. The complete argument can be found in Section 3.4. Theorem A covers already established transitivity results for instances of STSs such as labelled Markov processes [FKP17] (see also [Des99, DEP02, DGJP03]) and their conservative extension of non-deterministic labelled Markov processes including internal non-determinism [DTW12, Wol12].

An STS may involve non-deterministic choices over a possible uncountable number of successor distributions in a state. Thus, reasoning about probabilities of sets of (infinite) paths requires the resolution of this non-determinism in terms of schedulers, also referred to as policies, adversaries, or strategies. Here, we follow standard concepts for discrete systems, e.g., [Seg95, BK08], and continuous systems, e.g., [CSKN05, Cat05, WJ06, Pan09, Wol12]. A trace is obtained by projecting a path in an STS on the sequence of the basic atomic observables, i.e., the actions. Hence, every scheduler also induces a distribution over traces. In the linear-time setting for stochastic models, two systems are considered to be equivalent if they induce the same trace distributions over sequences of observables. While the coinductive nature of bisimulations often facilitates elegant arguments about states and their transitions, reasoning about trace distributions is typically more involved as we also see in Part II of this thesis.

**Part II: simulations and trace distributions for Souslin systems.** In the non-stochastic setting, it is well known that bisimulation equivalence is finer than trace equivalence (see, e.g., [BK08]). A corresponding statement for stochastic models involving trace distributions has been shown, e.g., for models with discrete state spaces such as probabilistic automata [Seg95], continuous-time Markov decision processes [NK07], labelled concurrent Markov chains [DGJP10], and for simple STSs in the case of a global notion of bisimulation [Cat05]. The main objective of Part II is to investigate the mentioned connection between the bisimulation equivalence and the trace-distribution equivalence for STSs where our main contribution can be summarised as follows:

**Theorem B.** *For every Souslin STS the Souslin-simulation preorder is a subset of the trace-distribution preorder and accordingly, the Souslin-bisimulation equivalence is finer that the trace-distribution equivalence, i.e., for every states $s_a$ and $s_b$ the following two implications hold:*

(1) $s_a \preceq^{\text{sou}} s_b$ *implies* $s_a \leq^{\text{tr}} s_b$.

(2) $s_a \simeq^{\text{sou}} s_b$ *implies* $s_a =^{\text{tr}} s_b$.

⌟

A proof of this theorem is finally established in Section 4.6. The sketched result provides important sufficient criteria for proving trace-distribution preorder and equivalence of STSs. The latter are often intricate as one has to consider infinite paths or more precisely, probability measures on the set of all infinite paths. We show, roughly speaking, that local reasoning about states and their outgoing transitions suffices to determine whether there exists a simulation or a bisimulation relating given states of an STS. Theorem B is associated with the following scheduler synthesis problem: given two states $s_a$ and $s_b$ such that $s_b$ simulates $s_a$ and an $s_a$-scheduler, the task is to generate an $s_b$-scheduler inducing the same trace distribution. The challenging part in the proof of this theorem originates from the fact that schedulers are required to form Borel functions.

Theorem B includes two requirements referring to Souslin sets. First of all, the STS under consideration is supposed to be Souslin, meaning that the corresponding transition relation forms a Souslin set. It even turns out that the theorem does not hold any longer if one drops this Souslin assumption on the STS. Besides this, an existing (bi)simulation relating the states under consideration is required to constitute a Souslin set. These Souslin requirements enable the applicability of measurable-selection principles from the literature [Wag80, HPV81, AB06, Bog07] that yield a key ingredient for our proof of Theorem B for the construction of specific schedulers. The potential of measurable-selection principles has been already discovered in the context of stochastic control problems [DIY79, BS96], optimisation problems [Rie78] and stochastic relations [Dob07].

Measurability considerations constitute an essential element in the study of systems that involve uncountable state spaces or uncountable action spaces. For instance, known techniques for proving the logical characterisation of (bi)simulation for labelled Markov processes [Des99, DEP02, DGJP03, DDLP06, Pan09, FKP17] and non-deterministic labelled Markov processes [DTW12, Wol12] heavily exploit the measurability requirements on the underlying stochastic model and rely on argumentation principles from measure theory. The proof in [NK07] showing that bisimulation preserves continuous stochastic logic for continuous-time Markov decision processes is also substantially based on measure theory. Furthermore, the articles [CSKN05, Cat05, WJ06] study measurability questions associated to schedulers and executions of STSs. In general, the challenge is to find generic measurability assumptions that are satisfied by a large and important class of stochastic systems while ensuring strong structural properties for theoretical arguments.

Despite the Souslin assumptions, Theorem B has various important consequences. Indeed, the logical characterisations of the simulation preorder and the bisimulation equivalences to investigated in Part III of this thesis yield a powerful subclass of Souslin STSs where the Souslin-simulation preorder and the simulation preorder are the same and accordingly, where the Souslin-bisimulation equivalence and the bisimulation equivalence collapse. This class includes probabilistic automata [Seg95], Markov decision processes [Put94], labelled Markov processes [Des99, DGJP03, Pan09], image-finite non-deterministic labelled Markov processes [DTW12, Wol12], and continuous controlled Markov processes [DIY79, BS96, ZEM$^+$14, TMKA16].

Independently of the latter mentioned results, many prominent modelling formalisms known from the literature turn out to admit a semantics in terms of Souslin STSs, e.g., stochastic automata [D'A99, BD04, DK05], probabilistic rectangular hybrid automata [Spr01, KNSS02, Spr11, ZSR$^+$12, Spr15], and o-minimal hybrid automata for common o-minimal theories [LPS00, BM05, BBC06]. There are also tight connections to the semantical model for the prominent modelling language HMODEST [BDHK06, HHHK12]. Besides HMODEST, instances of Souslin STS can be also analysed with the software tools HyTech [HHWt97], PHAVER [Fre08], FAUST2 [SGA15], and KeYmaera [Pla08, Pla10, Pla15]. There are other prominent approaches for the analysis of stochastic systems with uncountable state spaces, e.g., based on (stochastic) satisfiability modulo theories [FTE10, EGF15], statistical model-checking techniques [ZBC12, EGF15], or counterexample-guided abstraction refinement approaches [NDN$^+$16].

**Theorem C.** *For every deterministic purely stochastic Souslin STS the Souslin-bisimulation equivalence and the trace-distribution equivalence are the same, i.e., for every states $s_a$ and $s_b$ it holds:*

$$s_a \simeq s_b \quad \textit{iff} \quad s_a \simeq^{\mathrm{sou}} s_b \quad \textit{iff} \quad s_a =^{\mathrm{tr}} s_b.$$

*In particular, the relation $=^{\mathrm{tr}}$ is a bisimulation.* ⌟

For arbitrary STSs two states whose associated sets of induced trace distributions coincide are not bisimilar to each other in general. However, by Theorem C, the corresponding result holds for a subclass of STSs. This theorem is presented in Section 4.8 of this thesis (see also Section 7.2). Although our proof showing that the trace-distribution equivalence $=^{\mathrm{tr}}$ forms a bisimulation is technical at some places, the argument only uses basic concepts from measure theory. The transition relation of a purely stochastic STS is completely determined by a control law that assigns to every state a uniquely determined successor distribution over action-state pairs. From this it directly follows that simulation preorder and bisimulation equivalence coincide for this model. To this end, Theorem C is presented from the bisimulation perspective only. Intuitively, a purely stochastic STS is deterministic provided in every state $s$ the execution of an action *act* almost surely leads to a uniquely determined successor state. It turns out that (deterministic) purely stochastic Souslin STSs yield an interesting subclass of STSs as they already cover powerful modelling formalisms from the literature such as stochastic timed automata [BBB$^+$14, BBCMar], semi Markov processes [Whi80, LHK01, BHHK03, GJP06], and discrete-time stochastic hybrid automata [AKLP10, AKM11, SA13].

For further related work concerning Part II of this thesis we also refer to [HJS07, KK12] and [JS09, SS11, JSS15] where a coalgebraic framework on a trace semantics for stochastic systems is developed. Moreover, the article [FKS16] investigates algorithmic questions for discrete Markovian models referring to trace distributions. For instance, it is shown that the following general trace-refinement problem is undecidable: given two discrete Markov decision processes and a scheduler for one of them, the general trace-refinement problem asks whether there exists a scheduler for the other system such that the corresponding trace distributions coincide.

**Part III: action-based probabilistic temporal logics.** The main objective of Part III in this thesis is to provide a characterisation of simulation preorder and bisimulation equivalence for a large subclass of (Souslin) STSs in terms of respective temporal action-based probabilistic logics. For bisimulation we consider an expressive logic called APCTL$^*$ to formalise constraints for probabilities for possibly complex path properties with conditions on the accumulated reward. The second logic for bisimulation called APCTL$_\circ$ is in the spirit of probabilistic Hennessy-Milner logic [HM85] where the Boolean fragment is restricted to conjunctions. Indeed, APCTL$^*$ includes the until modality and hence, is convenient for specifying conditions on infinite behaviours such as liveness properties (e.g., repeated reachability) and safety properties (e.g., infinite-horizon invariants). In contrast to that, properties formulated by APCTL$_\circ$ only concern conditions on direct successors in a state. For the characterisation of the simulation preorder we consider the two temporal logics $\exists$APCTL$^*$ and APCTL$_\bullet$. Again, while $\exists$APCTL$^*$ constitutes a comparable expressive logic capturing, e.g., the until modality, the second logic APCTL$_\bullet$ yields an inexpressive sublogic of $\exists$APCTL$^*$ similar to APCTL$_\circ$.

The temporal logic APCTL$^*$ is in the spirit of PCTL [HJ94, Bd95], PCTL$^*$ [ASB$^+$95], and CSL [ASSB00, BHHK03] that extend the classical branching-time logic CTL [EC82] with probabilities and discrete respective continuous time. Besides this, APCTL$^*$ also includes an accumulation modality in the spirit of the non-probabilistic logics in [BKKW14, BCHK14] and of the extension of CSL given by CSRL [HCH$^+$02, Clo06]. In view of the action-based setting, we are influenced by the action-based variant of CTL in [DV90]. Intuitively, $\exists$APCTL$^*$ is given by the existential fragment of APCTL$^*$ where in addition every form of negation is absent. The logics APCTL$_\circ$ and APCTL$_\bullet$ yield very restricted fragments of APCTL$^*$ and are in particular inspired by the modal logics in [BDEP97, Des99, DEP02, Pan09] for labelled Markov processes and [DTW12, Wol12] for non-deterministic labelled Markov processes.

**Theorem D.** *We consider an STS augmented with an action event family and a reward function. Assume the STS is non-blocking, image-finite, and Borel concerning the hit sigma algebra. Then the simulation preorder, the Souslin-simulation preorder, as well as the preorders induced by the* *temporal logics APCTL$_\bullet$ and $\exists$APCTL$^*$ are the same. Accordingly, the bisimulation equivalence, the Souslin-bisimulation equivalence, as well as the equivalences induced by APCTL$_\circ$ and APCTL$^*$ coincide. Thus, for every states $s_a$ and $s_b$ one has:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\mathrm{sou}} s_b$ *iff* $s_a \preceq_\bullet s_b$ *iff* $s_a \preceq^\exists s_b$.

(2) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$ *iff* $s_a \simeq_\circ s_b$ *iff* $s_a \simeq^* s_b$.

┘

As a consequence of the logical characterisation of bisimilarity in terms of the full logic APCTL$^*$ we get that bisimilar states are indeed equivalent for a wide range of properties with possibly complex temporal and reward-bounded constraints. By the logical characterisation in terms of the small sublogic of APCTL$_\circ$, non-bisimilar systems can be distinguished by comparably simple formulas. The same discussion applies for the simulation preorder and the corresponding logics $\exists$APCTL$^*$ and APCTL$_\bullet$. Besides this, Theorem D identifies a large class of STSs where the simulation preorder and the Souslin-simulation preorder as well as the bisimulation equivalence and the Souslin-bisimulation equivalence are the same, respectively. As every STS in the presented theorem is Souslin, we can thus apply Theorem B to obtain the following corollary referring to the trace-distribution relations:

**Corollary E.** *We consider an STS that is non-blocking, image-finite, and Borel concerning the hit sigma algebra. For every states $s_a$ and $s_b$ the following two implications hold:*

(1) $s_a \preceq s_b$ *implies* $s_a \leq^{\mathrm{tr}} s_b$.

(2) $s_a \simeq s_b$ *implies* $s_a =^{\mathrm{tr}} s_b$.

┘

To show that the properties formulated by APCTL$^*$ and $\exists$APCTL$^*$ are preserved by simulation and bisimulation, respectively, we can basically rely on slight extensions of the techniques developed in Part II of this thesis. To show that non-bisimilar states can be distinguished by an APCTL$_\circ$ state formula, we follow the argumentation scheme in [FKP17] that relies on Dynkin's $\pi$-$\lambda$ theorem and the unique structure theorem. Replacing these two main ingredients by corresponding new ones [FKP17], the same technique also applies for the characterisation of the simulation relation.

The approach in [FKP17] heavily exploits the countability of the logic. In our setting with uncountable action spaces, however, we cannot rely on this argumentation with a naive definition of the logics APCTL$_\bullet$ and APCTL$_\circ$, i.e., where all the individual actions serve as the basic atomic building blocks. Our main idea to treat this problem is as follows: instead of focusing on individual actions, the logics $\mathrm{APCTL}_\bullet$ and $\mathrm{APCTL}_\circ$ use a countable number of certain Borel subsets of the action space as the basic atomic observables. These subsets are required to satisfy some natural conditions that are formalised by the notion of an action event space. As a consequence, Theorem D also covers STSs whose action spaces may be uncountable.

We also provide a variant of Theorem D showing that the provided logical charactersations in particular hold for image-finite non-deterministic labeled Markov processes with countable action spaces [DTW12, Wol12]. Consequently, the contributions of Part III of this thesis cover labelled Markov processes with countable action spaces [Des99, DEP02, Pan09], stochastic automata [D'A99, BD04, DK05], probabilistic guarded-command language [MM04], and a subclass of stochastic hybrid systems [FHH$^+$11, Hah13].

**Part IV: parallel composition based on spans and couplings.** A major objective in defining compositional frameworks is to separate concerns into components (specifying the operational behaviour) and composition operators (addressing the communication and interaction of the components). Compositionality has its roots in the theory of process calculi, i.e., CCS [Mil82], CSP [Hoa85], and ACP [Bae05]. There are various investigations in the stochastic setting concerning, e.g., variants for CCS [HJ90, Yi91, YL92, GJS94, Tof94, Yi94] and CSP [Low91, DHK99] as well as process algebras for performance modelling and evaluation [Hil96], stochastic automata [D'A99, DK05], and interactive generalised semi-Markov processes [Gly89, BG02, Bra04, GJP04]. A process-algebraic approach to model and analyse software applications can be found in [ABC10]. There are also parallel operators for probabilistic automata [Seg95], simple STSs [CSKN05, Cat05], stochastic hybrid systems [Str05, BDHK06, HHHK12], and weighted Markov decision processes [DH13a]. A unifying approach is proposed, e.g., in [BNL13]. There are many variants and approaches for describing the interaction and communication of components within models for distributed systems, e.g., input/ output automata [LT87, LS89] where the set of actions is partitioned into input, output, and internal actions. The latter model has been extended to timed and hybrid systems [WSS97, LSV03] and probabilistic variants thereof [ML07, DLM16]. A related approach is given by interface automata, e.g., [dAH01]. Interface automata yield an automata-based formalism convenient for the specification of temporal aspects referring to the communication of components.

Within the above approaches for stochastic systems, the composition operator relates probability distributions of the individual components to be composed in terms of the product measure. Therefore, these operators are based on the assumption that components interact stochastically independent. However, there are situations where multiple components of a system are governed by a common context that yields stochastic dependencies, e.g., common cause failures. Such failures originate from the same cause and affect multiple components at the same time. A common source might be given by, e.g., a manufacturing defect inherent in components of the same type or common external influences the components are exposed to. In these situations it is not adequate to rely on the assumption that components behave stochastically independent.

In Part IV of this thesis we focus on a compositional framework convenient for the declarative modelling of stochastic dependencies in, e.g., common cause situations and other compositional scenarios including stochastic dependencies. Relying on specific couplings of probability measures, a new parallel operator for stochastic systems is introduced where stochastic dependent transitions are combined. In this way we can incorporate stochastic dependencies referring to the effect of actions in certain situations. The operator is moreover indexed by a span. Intuitively, spans allow for arbitrary sets and associated projections functions to specify the global state space of the STSs to be composed.

A coupling relates two a priori unrelated probability measures in the same space. As a consequence, couplings yield a flexible and declarative modelling formalism convenient for expressing stochastic dependencies between components. The potential of couplings has been already recognised in many different areas. For instance, couplings enable elegant argumentation techniques in probability and optimal-transport theory [Lin92, Lin99, LPW09, Vil09]. These proof schemes have been applied by the formal-verification community for proving differential privacy [BEG$^+$15, BGG$^+$16a, BGG$^+$16b, BGHS17]. Simulations and bisimulations using weight functions also rely on couplings. An overview on different applications of couplings in computer science can be also found in [DD09].

**Theorem F.** *Under some generic side constraints, for every simple STSs $\mathcal{T}_{a1}$, $\mathcal{T}_{a2}$, $\mathcal{T}_{b1}$, and $\mathcal{T}_{b2}$ simulation and bisimulation are congruences with respect to the span-coupling composition operator denoted by $\|$, i.e., the following two statements hold:*

(1) $\mathcal{T}_{a1} \preceq \mathcal{T}_{b1}$ *and* $\mathcal{T}_{a2} \preceq \mathcal{T}_{b2}$ *implies* $\mathcal{T}_{a1} \parallel \mathcal{T}_{a2} \preceq \mathcal{T}_{b1} \parallel \mathcal{T}_{b2}$.

(2) $\mathcal{T}_{a1} \simeq \mathcal{T}_{b1}$ *and* $\mathcal{T}_{a2} \simeq \mathcal{T}_{b2}$ *implies* $\mathcal{T}_{a1} \parallel \mathcal{T}_{a2} \simeq \mathcal{T}_{b1} \parallel \mathcal{T}_{b2}$.

⌟

In the context of process calculi, an important aspect of (bi)simulation is the compatibility with syntactic operators in the process calculus. Theorem F provides a generic congruence result for the simulation preorder and the bisimulation equivalence with respect to our newly introduced composition operator. The precise formulation of this result is presented in Section 6.6. While Theorem F is usually trivial for standard composition operators, the challenge towards a proof is the construction of suitable couplings representing given stochastic dependencies. For the latter we exploit the disintegration theorem. Intuitively, the disintegration theorem provides a way to decompose couplings into its components and their dependencies.

A congruence result with respect to trace-distribution relations already fails for subclasses of STSs (see Example 7.4.1 in [Seg95] and also [LSV07]). That is why, Part IV of this thesis only concentrates on relations induced by simulations and bisimulations.

**Publications of the author in the scope of this thesis.** The material of Part IV concerning the composition operator for STSs is published in the proceedings of the 43rd International Colloquium on Automata, Languages and Programming (ICALP) [GBK16]. The journal version of [GBK16] has been submitted [GB17] and additionally covers Part II on simulation, bisimulation, and trace-distribution relations. Part III of this thesis has been accepted for publication in the proceedings of the 16th International Conference on Hybrid Systems: Computation and Control (HSCC) [GB18]. In fact, the logical characterisation of the simulation preorder and the bisimulation equivalence provided in this thesis goes beyond [GB18] as the cited paper only covers purely stochastic Souslin STSs.

# 2 Probability measures on Polish spaces

This chapter recalls concepts from the literature towards a self-contained framework about Polish spaces, on which we rely on throughout this thesis. We also present observations as well as auxiliary lemmas that, to the best of our knowledge, extend existing literature. However, whose proofs usually do not involve fundamental new insights.

First, we provide a brief overview of important terms and notions regularly used in this thesis and sketch their significance.

**Polish spaces.** Common sets (equipped with their natural metrics) occurring in modelling stochastic systems are covered by Polish spaces, e.g., every countable set, the real numbers $\mathbb{R}$, the non-negative real numbers $\mathbb{R}_{\geq 0}$, and the set of all evaluations of a countable set of real-valued variables. The notion of a Polish space enables an advanced mathematical theory [Kec95, Arv98, AB06, Sri08, Gao08], providing powerful results such as the disintegration theorem or measurable-selection principles for set-valued functions. Besides this, Polish spaces constitute the key ingredient of descriptive set theory where certain classes of subsets of Polish spaces are studied, e.g., the class of all Borel sets as well as the class of all Souslin sets (see also [Kan95]). It turns out that Polish spaces yield a convenient and common setting for the study of stochastic system including uncountable state and action spaces (see also [Dob07, Pan09]).

**Souslin sets.** Concepts related to Souslin sets provide a crucial mathematical tool in this thesis. Souslin sets (also called analytic sets) are special subsets of Polish spaces that enjoy many favourable closure properties [Kec95, Bog07], in particular, the image of a Souslin set under a Borel function also constitutes a Souslin set. Since every Borel set is also a Souslin set, the notion of a Souslin set yields a powerful mathematical formalism for the subsequent chapters of this thesis. For instance, we investigate a Souslin condition for stochastic transition systems in Chapter 4. Thanks to the rich theory on Souslin sets, this requirement is general enough to cover many powerful modelling formalisms from the literature while maintaining strong mathematical properties for theoretical arguments.

**Measurable-selection principles.** At various points in this thesis the existence of certain Borel functions with appropriate properties is required, for instance, in the construction of schedulers for stochastic transition systems inducing a given trace distribution. In such a situation we typically rely on a measurable-selection principle [Wag80, HPV81, AB06, Bog07]. Such a principle provides sufficient conditions on a set of functions represented by a set-valued function that guarantees the existence of a Borel function in that set. The task of showing the existence of a specific Borel function then reduces to provide an appropriate set of functions and to show that this set satisfies the requirements of a measurable-selection principle. This procedure has been also successfully employed in many applied areas, e.g., concerning stochastic optimal control [DIY79, BS96].

**Barycentres and convex hulls.** Consider a set $P$ of probability measures on some Polish space. Intuitively, the convex hull of $P$ consists of exactly those probability measures that are obtained by, roughly speaking, a reweighting of probability measures in $P$. More precisely, such a reweighting is given by a barycentre of a probability measure that forms a distribution over the probability measures in $P$. Our notions regarding barycentres and convex hulls are inspired by [DM88] (see also [MS00]). It turns out that building the convex hull of sets of probability measures yields a closure operator when focusing on Souslin sets. The mathematical challenge is to prove that taking the convex hull of a set of probability measures twice gives the same result as if it would have been taken once. We provide a direct argument for this statement without a detour on related results in [DM88] for locally convex topological vector spaces.

**Couplings of probability measures.** A coupling is a probability measure on a product space that relates two probability measures in the same space. The crux is that besides a few exceptions there are various couplings between probability measures such that the probability measures under consideration can be related from different perspectives. The latter fact enables many elegant proof techniques in probability theory [Lin92, AGS05, LPW09, Vil09]. Besides this, couplings have versatile applications in different areas of (theoretical) computer science. For instance, a weight function in the context of simulation and bisimulation is a coupling being compatible with relation (see below). Besides this, behavioural metrics relying on the Kantorovich lifting for stochastic systems exploit couplings (see also [DD09] for a survey). Couplings haven been moreover recently employed for the formal verification of differentially private algorithms [BEG$^+$15, BGG$^+$16b, BGG$^+$16a, BGHS17].

**Weight functions.** Intuitively, a coupling that is additionally compatible with a given relation is called a weight function for this relation. The concept of weight functions is not new and has been studied in both the mathematics and computer-science community. For instance, the series of articles [Str65, KKO77, Edw78, Kel84, Ska93] study sufficient and necessary conditions for the existence of probability measures with given marginal distributions from a purely mathematical point of view. These contributions have many applications concerning, among others, stochastic inequalities, distances between probabilities measures, and comparisons of stochastic processes. Interestingly, in the computer-science community weight functions also play a key role for the comparison of the branching-time behaviour of stochastic systems by means of simulations and bisimulations. This discussion is postponed to the next chapter where we introduce stochastic transition systems and related notions. The concept of weight function also appears in the thesis [Jon90] in the context of probabilistic power domains.

**Smooth and weakly smooth relations.** Intuitively, considering smooth relations from the perspective of descriptive set theory where one studies the complexity of the definition of subsets of Polish spaces, a smooth relation are precisely those relations that are Borel reducible to the diagonal relation of some Polish space [Kec95, Sri08]. As diagonal relations are comparable simple relations, smooth relations enjoy many strong structural properties. For smooth as well as weakly smooth relations we prove a characterisation for the existence of specific weight functions the thesis benefits at various points. More precisely, we provide a characterisation of the corresponding weight lifting in terms of a lifting used, e.g., in the context of simulation and bisimulation for labelled Markov processes [Pan09]. Our argumentation for the mentioned characterisation requires no essential new mathematical ideas and basically uses Strassen's theorem on stochastic domination [Str65, KKO77, Kel84, Les10] as well as the techniques developed in the recent paper [FKP17]. Our investigations on smooth as well as weakly smooth relations also yield important links to related work as we discuss later in Chapter 7.

## 2.1 Setting the mathematical framework

This section introduces the overall mathematical framework of this thesis, in particular, recalls basic definitions and results for Polish spaces and related notions. The reader is supposed to be familiar with standard concepts from measure and probability theory as well as elementary material concerning metric and topological spaces [Bil95, Bil99, Fre01, Kal02, Bog07, Sch08]. The first chapters in [Pan09] provide an introduction to these topics adapted for computer scientists. We briefly elaborate on the basic elements of measure theory first, i.e., sigma algebras, (probability) measures, and measurable functions.

**Sigma algebras.** The notion of a sigma algebra is fundamental in measure theory, e.g., for specifying measurable information of sets, for defining measures quantifying these available information, and for declaring the events of interest in probability theory. Recall, a sigma algebra on a set $X$ is a family $\mathcal{B}$ of subsets of $X$ such that $\mathcal{B}$ is not empty, $\mathcal{B}$ is closed under complementation, and $\mathcal{B}$ is closed under countable unions. A measurable space is a set that is endowed with a sigma algebra. Elements of a sigma algebra are called measurable sets. In many situations a sigma algebra on a set $X$ is specified by a family of subsets of $X$. More precisely, for every family $\mathcal{G}$ of subsets of $X$ there exists a uniquely determined sigma algebra $\mathcal{B}$ on $X$ such that $\mathcal{G} \subseteq \mathcal{B}$ and so that for every sigma algebra $\mathcal{B}'$ on $X$ it holds $\mathcal{B} \subseteq \mathcal{B}'$. In other words, $\mathcal{B}$ is the smallest sigma algebra on $X$ that contains every set in $\mathcal{G}$. In this context, the family $\mathcal{G}$ is called a generator of the sigma algebra $\mathcal{B}$ and we say that the sigma algebra $\mathcal{B}$ is generated by $\mathcal{G}$.

**Probability measures.** As sketched before, a measure assigns a real value to every measurable set and hence, quantifies the available information specified by the sigma algebra. To be more precise, consider a measurable space $X$ and denote the associated sigma algebra by $\mathcal{B}$. A measure on $X$ with respect to $\mathcal{B}$ is a sigma additive function $\mu\colon \mathcal{B} \to [0,1]$ such that $\mu(\varnothing) = 0$. Note, we only work with measures where the domain is given by $[0,1]$, however, in general the domain may be the extended real number line $\mathbb{R} \cup \{-\infty, +\infty\}$. For every measurable set $B \in \mathcal{B}$ the value $\mu(B)$ intuitively represents the information content of $B$ concerning the measure $\mu$. For instance, for every $x \in X$ the very simple measure $\mathit{Dirac}[x]$, called Dirac measure concentrated at $x$, quantifies the information whether $x$ is contained in a given measurable set or not. Indeed, for every measurable set $B \subseteq X$ it holds $\mathit{Dirac}[x](B) = 1$ if $x \in B$ and $\mathit{Dirac}[x](B) = 0$ if $x \notin B$. In the context of probability theory, a probability measures assigns the likelihood of the occurrence of the events specified by the sigma algebra. Recall, a measure $\mu$ on a measurable space $X$ is a probability measure provided $\mu(X) = 1$. The set of all probability measures on a measurable space $X$ is denoted by $\mathit{Prob}[X]$.

**Carathéodory uniqueness and extension theorem.** We present two results named after Carathéodory topic of every basic measure theory course. Let $X$ be a measurable space whose sigma algebra is denoted by $\mathcal{B}$ and pick a generator $\mathcal{G}$ of $\mathcal{B}$. The precise statement of Carathéodory uniqueness theorem is as follows (see, e.g., Satz 5.1.1 in [Sch08]): if the generator $\mathcal{G}$ is closed under finite intersections, then for every measures $\mu\colon \mathcal{B} \to [0,1]$ and $\mu'\colon \mathcal{B} \to [0,1]$ one has $\mu = \mu'$ precisely when for every $B \in \mathcal{G}$ it holds $\mu(B) = \mu'(B)$. Hence, to justify that two probability measures are the same, it suffices to consider selected subsets of the whole sigma algebra. As generators of sigma algebras often admit rich structural properties, Carathéodory uniqueness theorem allows for various applications in many contexts. Carathéodory extension theorem is in the same spirit as the result before (see, e.g., Satz 5.3.3 in [Sch08]): for every sigma additive function $\tilde{\mu}\colon \mathcal{G} \to [0,1]$ with $\mu(\varnothing) = 0$, there exists a measure $\mu\colon \mathcal{B} \to [0,1]$ such that $\mu(B) = \tilde{\mu}(B)$ for every $B \in \mathcal{G}$ provided the following two conditions hold: the generator $\mathcal{G}$ is closed under finite intersections and moreover, for every $B_1, B_2 \in \mathcal{G}$ there are a natural number $n \in \mathbb{N} \setminus \{0\}$ and pairwise disjoint sets $B'_1, \ldots, B'_n \in \mathcal{G}$ such that $B_1 \setminus B_2 = B'_1 \cup \ldots \cup B'_n$. Here, the measure $\mu$ hence yields an extension of the function $\tilde{\mu}$ to the whole sigma algebra $\mathcal{B}$.

**Example 1.** Let $X_1$ and $X_2$ be two measurable spaces. Define $X = X_1 \times X_2$. The set $X$ equipped with the product sigma algebra forms a measurable space. It is well-known that the product sigma algebra is generated by the family $\mathcal{G}$ consisting of all the sets $A_1 \times A_2$ where $A_1 \subseteq X_1$ and $A_2 \subseteq X_2$ are measurable. According to Carathéodory uniqueness theorem, for every $\mu \in \mathit{Prob}[X_1 \times X_2]$, $\mu_1 \in \mathit{Prob}[X_1]$, and $\mu_2 \in \mathit{Prob}[X_2]$ it holds $\mu = \mu_1 \otimes \mu_2$ iff for every measurable sets $A_1 \subseteq X_1$ and $A_2 \subseteq X_2$ one has $\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$. Here, $\mu_1 \otimes \mu_2$ denotes the product measure of $\mu_1$ and $\mu_2$. Moreover, it is easy to see that the family of sets $\mathcal{G}$ also satisfies the requirements of Carathéodory extension theorem. Define $\tilde{\mu}\colon \mathcal{G} \to [0,1]$ by $\tilde{\mu}(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2)$ for every measurable sets $A_1 \subseteq X_1$ and $A_2 \subseteq X_2$. Then, the product measure $\mu_1 \otimes \mu_2$ can be defined as the uniquely determined extension of $\tilde{\mu}$. Here, the argument that $\tilde{\mu}$ is indeed sigma additive is part of every basic measure theory course. ⌟

**Polish spaces.** A metric on a set $X$ is a function $\mathit{dist}\colon X \times X \to \mathbb{R}_{\geq 0}$ that intuitively assigns a distance to every pair of elements in $X$, i.e., the smaller the value $\mathit{dist}(x_1, x_2)$ is, the closer the distance of the elements $x_1$ and $x_2$ are. With that intuition in mind, the conditions on a metric are natural requiring the identity of indiscernibles, symmetry, and the triangle inequality. Every metric gives rise to a topological space. More precisely, the topology on $X$ induced by a metric $\mathit{dist}$ on $X$ is defined as the topology on $X$ generated by the basis consisting of all open balls in $X$ with respect to $\mathit{dist}$.

A Polish space is a complete and separable metric space [Kec95]. In particular, by the separability of a Polish space $X$ with respect to the metric $\mathit{dist}$, there exists a countable dense subset $\mathit{Dense}$ of $X$, i.e., for every $\varepsilon \in \mathbb{R}_{>0}$ and $x \in X$ there is $x' \in \mathit{Dense}$ with $\mathit{dist}(x, x') < \varepsilon$. Given a topological space $X$ with topology $\mathcal{O}$, we say that $\mathcal{O}$ is Polish provided $\mathcal{O}$ is separable and there is a complete metric on $X$ that induces $\mathcal{O}$. In this thesis we rarely work directly with the defining properties of Polish spaces. In fact, the precise definitions of the underlying metrics of given Polish spaces are often irrelevant. Instead, many of our mathematical arguments are heavily based on the developed mathematical theory for Polish spaces and thus, the notion of Polish spaces can be used as a black box.

Fortunately, those metric spaces that appear in our studies of stochastic systems with uncountable state and action spaces are Polish:

**Example 2.** The following itemisation provides examples for Polish spaces:

(1) Every countable set $Q$ with the discrete metric yields a Polish space. Here, the distance between elements $q_1$ and $q_2$ of $Q$ is one precisely when $q_1 \neq q_2$. The induced topology is the largest topology containing all the subsets as open sets.

(2) The real number line $\mathbb{R}$ with the Euclidean metric, i.e., $dist(r_1, r_2) = |r_1 - r_2|$ for every $r_1, r_2 \in \mathbb{R}$, forms a Polish space. The separability of $\mathbb{R}$ follows from the well-known fact that the rational numbers $\mathbb{Q}$ are countable and dense in $\mathbb{R}$.

(3) Every closed and every open subset of a Polish space with the corresponding subspace metrics constitutes a Polish space. The closed and the open unit interval, i.e., the metric spaces $[0, 1]$ and $(0, 1)$, respectively, and the non-negative real numbers $\mathbb{R}_{\geq 0}$ hence provide examples for Polish spaces.

(4) Let $(X_i)_{i \in Index}$ be a family of pairwise disjoint Polish spaces with *Index* being a countable set. The union of these sets with the disjoint-union topology forms a Polish space where a subset $O$ is open iff for every $i \in Index$ the set $O \cap X_i$ is open in $X_i$. Hence, $Q \cup \mathbb{R}_{\geq 0}$ is a Polish space where $Q$ is a countable set disjoint from $\mathbb{R}_{\geq 0}$.

(5) Let $(X_i)_{i \in Index}$ be a family of Polish spaces such that *Index* is a countable set. Then the Cartesian product with the product topology constitutes a Polish space. It in particular follows that the Cantor space $\{0, 1\}^\omega$ and the Baire space $\mathbb{N}^\omega$ are a Polish spaces that have an important role in descriptive set theory [Kec95].

(6) For every Polish space $X$ the set $Prob[X]$ of all probability measures on $X$ constitutes a Polish space (see Theorem 17.23 in [Kec95]). The latter statement is not trivial and relies, e.g., on convergence properties of probability measures given by the prominent Portmanteau theorem (see Theorem 17.20 in [Kec95]).

┘

Relying on the previous examples and closure properties of Polish spaces, we think that the setting of Polish spaces is general enough for the study of stochastic systems with uncountable state and action spaces (see Section 3.1 and also [Dob07, Pan09]). In particular, by Example 2 (5) and (6), if the state space *Sta* and the action space *Act* of a stochastic system form Polish spaces, the set $Prob[Act \times Sta]$ of probability measures over action-state pairs yield a Polish space.

**Borel sigma algebras.** There is a natural sigma algebra associated to topological spaces. Throughout this thesis every topological space is equipped with the associated Borel sigma algebra if not stated otherwise. Recall, given a topological space $X$ whose topology is denoted by $\mathcal{O}$, the Borel sigma algebra on $X$, which we denote by $Borel[X]$ or $Borel[\mathcal{O}]$, is generated by the family of sets $\mathcal{O}$. As sigma algebras are closed under complementation, Borel sigma algebras are also generated by the family of the closed subsets of $X$. The elements of a Borel sigma algebra are called Borel sets. A function $f\colon X \to Y$ between metric spaces $X$ and $Y$ is called Borel provided the function $f$ is measurable with respect to the Borel sigma algebras on $X$ and $Y$, i.e, the preimage $f^{-1}(B_X)$ is a Borel set in $Y$ for every Borel set $B_X \subseteq X$.

**Example 3.** We revisit the Polish spaces presented in Example 2 and investigate the associated Borel sigma algebras.

(1) The Borel sigma algebra of a countable set $Q$ is the largest sigma algebra on $Q$ containing all subsets of $Q$ as Borel sets. It is easy to see that this sigma algebra is generated by the family $\{\{q\} \ ; q \in Q\}$ consisting of all the singleton sets.

(2) We consider the Borel sigma algebra of the real number line $\mathbb{R}$. It is well-known that this sigma algebra is for instance generated by the family consisting of all the intervals $[q_1, q_2)$ with rational endpoints $q_1, q_2 \in \mathbb{Q}$.

(3) Let $X$ be a Polish space and $C \subseteq X$ be a closed set. The Borel sigma algebra of the Polish $C$ and the induced subspace sigma algebra on $C$ in $X$ are the same, i.e., one has the identity $Borel[C] = \{B \in Borel[X] \ ; B \subseteq C\}$ (see Lemma 6.2.4 in [Bog07]). The same observation applies for the open subsets of $X$.

(4) Let $X$ and $Y$ be disjoint Polish spaces. Then $Borel[X \cup Y]$ is equal to the disjoint-union sigma algebra of $Borel[X]$ and $Borel[Y]$, i.e., a subset $B \subseteq X \cup Y$ is contained in $Borel[X \cup Y]$ iff one has $B \cap X \in Borel[X]$ and $B \cap Y \in Borel[Y]$.

(5) For every two Polish space $X$ and $Y$ one has that $Borel[X \times Y]$ and the product sigma algebra of $Borel[X]$ and $Borel[Y]$ are the same (see Lemma 6.4.2 in [Bog07]). In particular, $Borel[X \times Y]$ is generated by the family of all sets $B_X \times B_Y$ such that $B_X \in Borel[X]$ and $B_Y \in Borel[Y]$ (see also Example 1). The same holds for Cartesian products of a countable number of Polish spaces.

(6) Let $X$ be a Polish space. The Borel sigma algebra on $Prob[X]$ and the Giry sigma algebra on $Prob[X]$ are the same (see Theorem 17.24 in [Kec95]). Recalling [Gir82], the Giry sigma algebra on $Prob[X]$ is defined as the smallest sigma algebra on $Prob[X]$ such that the function $probeval_B \colon Prob[X] \to [0,1]$, $probeval_B(\mu) = \mu(B)$ becomes measurable for every Borel set $B \subseteq X$, i.e., the Giry sigma algebra is defined as an initial sigma algebra.

⌟

The presented examples of Polish spaces and their corresponding Borel sigma algebras are essentially the same from a measure-theoretic point of view: for every uncountable Polish spaces $X$ and $Y$ there exists a bijective function $Iso \colon X \to Y$ such that both functions $Iso$ and its inverse $Iso^{-1}$ are Borel functions (see Corollary 6.8.8 in [Bog07]). In this context, we refer to the function $Iso$ as Borel isomorphism (between $X$ and $Y$). Two Polish spaces are hence Borel isomorphic iff they have the same cardinality. In particular, as the real number line $\mathbb{R}$ is Polish, every uncountable Polish space has the cardinality of the continuum. This classifies all the Polish spaces up to Borel isomorphism. Moreover, to prove purely measure-theoretic results for Polish spaces, it suffices to consider a specific Polish space as a representative.

**Remark 4.** The Borel sigma algebra of every Polish space $X$ is generated by a family $\mathcal{G}$ of subsets of $X$ that satisfies the following three properties:

(1) $\mathcal{G}$ is countable.

(2) $\mathcal{G}$ is closed under complementation and finite intersections.

(3) $\mathcal{G}$ separates the points of $X$, i.e., for every $x_1, x_2 \in X$ one has

$$x_1 = x_2 \quad \text{iff} \quad \text{for every } B \in \mathcal{G} \text{ it holds } x_1 \in B \text{ iff } x_2 \in B.$$

Provided the Polish space $X$ is countable, the family $\mathcal{G} = \{B \subseteq X \,;\, B \text{ or } X \setminus B \text{ is finite}\}$ of all finite and cofinite subsets of $X$ satisfies the stated requirements. Assuming $X = [0,1]$, the family of all finite unions of intervals in $[0,1]$ with rational endpoints fulfills the given properties. As every uncountable Polish space is Borel isomorphic to $[0,1]$, we obtain a generator with the desired properties for every Polish space. ⌟

**Giry sigma algebras revisited.** Inspecting Example 3 (6) again, the definition of the Giry sigma algebra on $Prob[X]$ does not involve any topological properties of the Polish space $X$.

Hence, we can safely endow the set of all probability measures on an arbitrary measurable spaces (that is not necessarily induced by a Polish space) with the corresponding Giry sigma algebra.

**Remark 5.** Let $X$ and $Y$ be measurable spaces and $f\colon X \to Prob[Y]$ be a function. Let $\mathcal{G}_Y$ be a generator of the sigma algebra on $Y$ that is additionally closed under finite intersections. For every measurable set $B \subseteq Y$ define the function $f_B\colon X \to [0,1]$, $f_B(x) = f(x)(B)$. Then the following three statements are equivalent:

(1) $f$ is measurable.

(2) $f_B$ is measurable for every measurable set $B \subseteq Y$.

(3) $f_B$ is measurable for every $B \in \mathcal{G}_Y$.

The equivalence of (1) and (2) follows from properties of initial sigma algebras. In fact, the Giry sigma algebra is the initial sigma algebra with respect to the functions $probeval_B$ for every measurable set $B \subseteq X$. Here, the function $probeval_B$ is defined as in Example 3 (6). The equivalence of (2) and (3) can be proven by a standard application of Dynkin's $\pi$-$\lambda$ theorem (see, e.g., Theorem 136B in [Fre01]). ⌟

**Example 6.** Let $X$ and $Y$ be measurable space. Then the function $f\colon X \to Prob[X]$,

$$f(x) = Dirac[x]$$

and the function $g\colon Prob[X] \times Prob[Y] \to Prob[X \times Y]$,

$$g(\mu_X, \mu_Y) = \mu_X \otimes \mu_Y$$

are measurable. The latter can be shown relying on the characterization provided by Remark 5. We illustrate this by showing the claim for the function $g$. For every measurable sets $B_X \subseteq X$ and $B_Y \subseteq Y$ the function $g_{B_X \times B_Y}\colon Prob[X] \times Prob[Y] \to [0,1]$, $g_{B_X \times B_Y}(\mu_X, \mu_Y) = \mu_X(B_X) \cdot \mu_Y(B_Y)$ is measurable by the definition of the Giry sigma algebra and as the point-wise product of measurable functions yields a measurable function. The family consisting of all the sets $B_X \times B_Y$ with measurable $B_X \subseteq X$ and $B_Y \subseteq Y$ generates the sigma algebra on $X \times Y$ (see also Example 1) and hence, Remark 5 justifies that function $g$ is Borel. ⌟

**Sections of relations.** When considering probability measures on product spaces, the following notion for sections of sets turns out to be useful. Let $X$ and $Y$ be Polish spaces and $B \subseteq X \times Y$ be a set. For every $x \in X$ define

$$Section[B, x, \cdot] = \{y \in Y \,;\, \langle x, y \rangle \in R\}$$

and similarly, for every $y \in Y$ let

$$Section[B, \cdot, y] = \{x \in X \,;\, \langle x, y \rangle \in R\}.$$

Clearly, if one has $B = B_X \times B_Y$ for some subsets $B_X \subseteq X$ and $B_Y \subseteq Y$, for every $x \in X$ it holds $Section[B, x, \cdot] = B_Y$ if $x \in B_X$ and $Section[B, x, \cdot] = \varnothing$ if $x \notin B_X$. If the set $B$ is Borel in $X \times Y$, then the sets $Section[B, x, \cdot]$ and $Section[B, \cdot, y]$ are Borel in $X$ and $Y$, respectively, for every $x \in X$ and $y \in Y$ (see also Lemma 9.6.4 in [Sch08]).

**Semi-product measures.** Example 1 provides an overview on product measures. In turns out that product measures are an instance of the more general notion of semi-product measures. Let $X$ and $Y$ be measurable spaces, $\mu \in Prob[X]$ be a probability measure, and $f\colon X \to Prob[Y]$ be a measurable function. Recall, $Prob[Y]$ is equipped with the Giry sigma algebra. The semi-product measure (of $\mu_X$ and $f$) is given by the probability measure $\mu_X \rtimes f \in Prob[X \times Y]$ defined as follows (see Lemma 1.38 in [Kal02]): for every Borel set $B \subseteq X \times Y$ let

$$\mu_X \rtimes f(B) = \int f(x)(Section[B, x, \cdot])\, d\mu(x).$$

Thus, for every Borel sets $B_X \subseteq X$ and $B_Y \subseteq Y$ we obtain

$$\mu_X \rtimes f(B_X \times B_Y) = \int_{B_X} f(x)(B_Y)\, d\mu(x).$$

If $X$ and $Y$ are moreover countable, the integral is simplified as follows:

$$\mu_X \rtimes f(B_X \times B_Y) = \sum_{x \in B_X} \sum_{y \in B_Y} \mu_X(\{x\}) \cdot f(x)(\{y\}).$$

We write $f \ltimes \mu_X$ to indicate the analogously defined probability measure on $Y \times X$. Note, provided $\mu_Y \in Prob[Y]$ is a probability measure such that $f(x) = \mu_Y$ for every $x \in X$, we obtain the identity $\mu_X \rtimes f = \mu_X \otimes \mu_Y$.

**Post operator.** As before, let $X$ and $Y$ be measurable spaces, $\mu_X \in Prob[X]$ be a probability measure, and $f\colon X \to Prob[Y]$ be a measurable function. One can also think of the Borel function $f$ as taking the measure $\mu_X$ into a measure on $Y$. We formalize this by means of the following notion. Define the probability measure $Post[\mu_X, f]$ on $Y$ as the marginal of the semi-product measure $\mu_X \rtimes f$ concerning $Y$, i.e., we have $Post[\mu_X, f] \in Prob[Y]$ such that for every Borel sets $B_Y \subseteq Y$,

$$Post[\mu_X, f](B_Y) = \mu_X \rtimes f(X \times B_Y) = \int f(x)(B_Y)\, d\mu_X(x).$$

**Disintegration theorem.** Relying on the notion of semi-product measures, one can easily define probability measures on product spaces. Vice versa, the question which measures on a product space can be represented as a semi-product measures arises naturally. To be more precise, let $X$ and $Y$ be measurable spaces and $\mu \in Prob[X \times Y]$. Denote the marginal of $\mu$ concerning $X$ by $\mu_X$, i.e., we have $\mu_X \in Prob[X]$ and $\mu_X(B_X) = \mu(B_X \times Y)$ for all measurable sets $B_X \subseteq X$. Does there exists a measurable function $f\colon X \to Prob[Y]$ such that $\mu = \mu_X \rtimes f$? The disintegration theorem (see Exercise 17.35 in [Kec95] or Theorem 5.4 in [Kal02]) states that this question can be answered positively provided $X$ and $Y$ are Polish spaces, i.e., there indeed exist a Borel function $f\colon X \to Prob[Y]$ with $\mu = \mu_X \rtimes f$. Moreover, this function $f$ is almost surely uniquely determined: assuming a second Borel function $f'\colon X \to Prob[Y]$ with $\mu = \mu_X \rtimes f'$, there exists a Borel set $B_X \subseteq X$ such that $\mu_X(B_X) = 1$ and $f(x) = f'(x)$ for every $x \in B_X$.

The following example illustrates the disintegration principle for countable sets and also sketches a connection to the concept of conditioning. Note that disintegration and conditioning typically appear together in the literature (see, e.g., Chapter 5 in [Kal02]).

**Example 7.** Let $X$ and $Y$ be countable sets and $\mu \in Prob[X \times Y]$. To simplify the following discussion, assume that $\mu(\{x\} \times Y) > 0$ for every $x \in X$. As before, let $\mu_X \in Prob[X]$ be the probability measure defined by $\mu_X(\{x\}) = \mu(\{x\} \times Y)$ for every $x \in X$. We provide a Borel function $f\colon X \to Prob[Y]$ such that $\mu = \mu_X \rtimes f$. Define $f\colon X \to Prob[Y]$ as follows: for every $x \in X$ and $y \in Y$ let

$$f(x)(\{y\}) = \mu(X \times \{y\} \mid \{x\} \times Y),$$

i.e., $f(x)(\{y\})$ is the conditional probability of $X \times \{y\}$ given the event $\{x\} \times Y$. For every $x \in X$ and $y \in Y$ we therefore obtain

$$f(x)(\{y\}) = \frac{\mu((X \times \{y\}) \cap (\{x\} \times Y))}{\mu(\{x\} \times Y)} = \frac{\mu(\{x\} \times \{y\})}{\mu_X(\{x\})}.$$

Since $X$ is countable, the function $f$ is obviously Borel and moreover, it is easy to see that $\mu = \mu_X \rtimes f$. Indeed, for every $x \in X$ and $y \in Y$ we have

$$\mu_X \rtimes f(\{\langle x, y\rangle\}) = \mu_X(\{x\}) \cdot f(x)(\{y\}) = \mu(\{\langle x, y\rangle\}).$$

This confirms the disintegration theorem for the trivial case where the involved Polish spaces are countable. ⌟

**Fubini's theorem.** Every integral with respect to a probability measures on a product spaces can be computed using iterated integrals applying the prominent Fubini's theorem for Markov kernels (see Exercise 17.36 in [Kec95] or Theorem 5.4 in [Kal02]). The precise statement is as follows. Let $X$ and $Y$ be Polish spaces, $\mu_X \in \mathit{Prob}[X]$ be a probability measure as well as $f\colon X \to \mathit{Prob}[Y]$ and $g\colon X \times Y \to [0,1]$ be Borel functions. First of all, according to Fubini's theorem, for every $x \in X$ the function $g[x]\colon Y \to [0,1]$,

$$g[x](y) = \int g(x,y)\, df(x)(y).$$

is Borel. Moreover, for every Borel sets $B_X \subseteq X$ and $B_Y \subseteq Y$ Fubini's theorem provides the identity below:

$$\int_{B_X \times B_Y} g(x,y)\, d(\mu_X \rtimes f)(x,y) = \int_{B_X} \Big( \int_{B_Y} g(x,y)\, df(x)(y) \Big)\, d\mu_X(x).$$

Note that the integral on the righthand side is indeed well-defined since, as mentioned before, the function $g[x]$ is Borel for every $x \in X$.

**Pushforward functions.** Every probability measure on a domain of a measurable function can be transferred to the codomain by invoking the concept of a pushforward function. Let $X$ and $Y$ be measurable spaces. The pushforward function of a measurable function $f\colon X \to Y$ is defined as the function $f_\sharp\colon \mathit{Prob}[X] \to \mathit{Prob}[Y]$ given as follows: for every probability measure $\mu \in \mathit{Prob}[X]$ and measurable set $B_Y \subseteq Y$ let

$$f_\sharp(\mu)(B_Y) = \mu(f^{-1}(B_Y)).$$

It is easy to see that $f_\sharp(\mu)$ forms indeed a probability measure. In this context, one refers to the probability measure $f_\sharp(\mu)$ as a pushforward measure or as an image measure.

**Remark 8.** Let $X$ and $Y$ be measurable spaces and $f\colon X \to Y$ be a measurable function. Then the following statements hold for pushforward functions:

(1) The function $f_\sharp$ is Borel.

(2) Assuming $X$ and $Y$ are Polish spaces and if $f$ is surjective, then $f_\sharp$ is also surjective.

Whereas statement (1) is an easy consequence of Remark 5, a proof for statement (2) is much more involved (see Theorem 15.14 or Corollary 18.24 in [AB06]). ⌟

The notions are as before, i.e., $X$ and $Y$ are measurable functions, $\mu_X \in Prob[X]$ is a probability measure, and $f\colon X \to Y$ is a measurable function. At various points throughout this thesis, we rely on integration by substitution yielding a powerful tool to transform integrals with respect to pushforward measures (see Satz 9.4.1 in [Sch08]): for every Borel set $B_Y \subseteq Y$ and Borel function $g\colon Y \to [0,1]$ it holds

$$\int_{B_Y} g(y)\, df_\sharp(\mu_X)(y) = \int_{f^{-1}(B_Y)} g(f(x))\, d\mu_X(x).$$

Whereas the integral of the left hand side refers to the measurable space $Y$, the integral on the right hand side refers to the measurable space $X$.

**Outer measures.** A measure on a measurable space constitutes a function mapping every element of the corresponding sigma algebra to some real number between zero and one. Suppose a measurable space $X$ and a measure $\mu$ on $X$. Given an arbitrary (not necessarily measurable) set $M \subseteq X$, the expression $\mu(M)$ hence makes no sense. However, one can regard the outer measure of $M$ concerning $\mu$ instead. Relying on the concept of outer measures (see Section 1.5 in [Bog07]), every measure can be naturally extended to a function whose domain is given by the whole powerset. The idea is to over-approximate a subset of $X$ by measurable sets in $X$. More precisely, the outer-measure function (of $\mu$) is defined by the function $\mu^{\text{out}}\colon 2^X \to [0,1]$,

$$\mu^{\text{out}}(M) = \inf\{\mu(B)\ ;\ B \subseteq X \text{ measurable set such that } M \subseteq B\},$$

i.e., $\mu^{\text{out}}(M)$ is the greatest lower bound of all $\mu$ masses of measurable sets in $X$ subsuming the given set $M$. Note, as the set $X$ is measurable and $\mu(X) = 1$, the outer-measure function indeed assigns a real number between zero and one to every every subset of $X$. It is easy to see that $\mu(B) = \mu^{\text{out}}(B)$ for all measurable sets $B \subseteq X$ and therefore, the function $\mu^{\text{out}}$ yields an extension of $\mu$ to the whole powerset. However, the outer-measure function may be not sigma additive and hence, constitutes no measure anymore (with respect to the powerset sigma algebra). For instance, it is well-known that there does not exist a measure on the powerset of $[0,1]$ that extends the Lebesgue measure defined on the Borel sets of $[0,1]$. Recall, every Lebesgue measure is translation invariant and it is not possible to define a translation-invariant measure on $[0,1]$ (see, e.g., the proof of Theorem 2.9 in [Pan09]). However, the outer-measure function still enjoys the property of sigma subadditivity.

## 2.2 Souslin sets in Polish spaces

The 1905 paper [Leb05] of Lebesgue contains a mathematical mistake that may add to the list of the most interesting ones in the history of mathematics: it is claimed that a projection of a Borel subset in the plane $\mathbb{R} \times \mathbb{R}$ onto the real line $\mathbb{R}$ yields again a Borel subset. However, Souslin found a counterexample for this statement. Among others this had contributed to the development of the descriptive set theory and the accompanied notion of so-called analytic sets, which were later also called Souslin sets. Detailed historical information and facts on this and related topics are nicely presented in [Kan95].

This section recalls the notion of Souslin sets and presents results from the literature required for this thesis. Our exposition mainly follows Chapter 14 in [Kec95] and Chapter 6 in [Bog07]. In the latter part of the following material, observations referring to the outer measures of Souslin sets are proven. However, these lemmas could be skipped at a first reading and might be checked in detail later on demand.

**Definition 9.** Let $X$ be a Polish space. We call the set $M \subseteq X$ *Souslin* if $M = \varnothing$ or there exist a Polish space $Z$ and a continuous function $g\colon Z \to X$ such that $g(Z) = M$, i.e., the set $M$ can be represented as an image of a Polish space under a continuous function. ⌟

To obtain a feeling for Souslin sets, we solve one part of Exercise 14.3 in [Kec95] that requires a proof for the following statement: a subset $M$ of a Polish space $X$ is Souslin provided there exist a Polish space $Y$ and a Borel set $B \subseteq X \times Y$ such that

$$M = \{x \in X \text{ ; there exists } y \in Y \text{ with } \langle x, y \rangle \in B\},$$

i.e., the set $M$ is obtained by the projection of the set $B$ onto the Polish space $X$. Let us see why. The following argument moreover illustrates the usefulness of Polish spaces as well as their strong (topological) properties. Suppose the Polish space $Y$, the Borel set $B \subseteq X \times Y$, and the set $M \subseteq X$ are given as above. Our task is to show that $M$ is Souslin in $X$. Denote the topology on $X \times Y$ by $\mathcal{O}_{X,Y}$. By Theorem 13.1 in [Kec95], there exists a topology $\mathcal{O}$ on $X \times Y$ such that the following three statements hold:

$$\mathcal{O} \text{ is Polish}, \quad \mathcal{O}_{X,Y} \subseteq \mathcal{O}, \quad \text{and} \quad B \text{ is closed concerning the topology } \mathcal{O}.$$

Define the set $Z = B$. The set $Z$ is equipped with the induced subspace topology of $\mathcal{O}$. As a consequence of the two facts that $\mathcal{O}$ is Polish and $B$ is closed in $\mathcal{O}$, the topological space $Z$ forms a Polish subspace of $X \times Y$ (see Example 2 (3)). Introduce the function $f\colon Z \to X$, $f(x, y) = x$. It is easy to see that $f(Z) = M$. Since $\mathcal{O}_{X,Y} \subseteq \mathcal{O}$, it moreover follows that the function $f$ is continuous. Putting things together, the set $M$ is Souslin in $X$. Among other useful properties of Souslin sets, the following remark even shows that the proven sufficient criterion for a Souslin set is even necessary:

**Remark 10.** The following itemisation provides important properties of Souslin sets:

(1) Let $X$ be a Polish space. A set $M \subseteq X$ is Souslin iff there exists a Polish space $Y$ and a Borel set $B \subseteq X \times Y$ such that $M = \{x \in X \,;\, \text{there exists } y \in Y \text{ with } \langle x, y \rangle \in B\}$.

(2) Let $X$ be a Polish space. Every Borel set in $X$ is also Souslin in $X$. Vice versa, provided $X$ is uncountable, there are Souslin sets in $X$ that are not Borel in $X$.

(3) For every $n \in \mathbb{N}$ let $M_n$ be a Souslin set in the Polish space $X_n$. Then the Cartesian product $M_0 \times M_1 \times \ldots$ is Souslin in the product space $X_0 \times X_1 \times \ldots$.

(4) Let $X$ be a Polish space. Like the family of all Borel sets in $X$, the family of all Souslin sets in $X$ is closed under both countable intersections and countable unions.

(5) Let $X$ and $Y$ be Polish spaces and $f\colon X \to Y$ be a Borel function. For every Souslin set $M_Y \subseteq Y$ the preimage $f^{-1}(M_Y)$ is Souslin in $X$ and for every Souslin set $M_X \subseteq X$ the image $f(M_X)$ is Souslin in $Y$.

(6) Let $X$ and $Y$ be Polish spaces and $f\colon X \to Y$ be a function. Then the function $f$ is Borel iff the set $\mathit{Graph}[f]$ is Borel in $X \times Y$ iff the set $\mathit{Graph}[f]$ is Souslin in $X \times Y$. Here, we define $\mathit{Graph}[f] = \{\langle x, y \rangle \in X \times Y \,;\, f(x) = y\}$.

The characterisation of Souslin sets in (1) can be found in Exercise 14.3 in [Kec95]. Here, a proof for the difficult implication is presented in the discussion before the remark. Statement (2) yields a combination of the Theorems 13.7 and 14.2 in [Kec95]. Statement (3) follows easily from the definition of Souslin sets and closure properties of Polish spaces (see Example 2 (5) and also Lemma 6.6.5 in [Bog07]). The closure properties stated in (4) and (5) can be derived from Proposition 14.4 in [Kec95]. The characterisation of Borel functions in statement (6) is proven by Theorem 14.12. in [Kec95]. ⌟

According to the characterisation of (1) of the previous remark, Souslin sets are precisely those subsets that arise by the projections of a Borel subsets in product spaces or, in other words, that can be represented by means of an existential quantification over a Polish space. In fact, taking up our introductory discussion concerning a famous mistake of Lebesgue, statements (1) and (2) together justify the existence a Borel set $B$ in the plane $\mathbb{R} \times \mathbb{R}$ such that its projection onto the real line $\mathbb{R}$ is not Borel. Inspecting (4), one might ask whether the family of all Souslin sets in a Polish space forms a sigma algebra. This is not the case in general as Souslin sets are not closed under complementation. In fact, one can show that a subset $B$ of a Polish space $X$ is Borel precisely when both $B$ and its complement $X \setminus B$ are Souslin in $X$ (see Corollary 6.6.10 in [Bog07]). Recall, the image of a Borel set under a Borel functions is not Borel in general. However, the corresponding result for Souslin sets holds by statement (5). Summarising the previous discussions, Remark 10 provides many strong closure properties of Souslin sets.

**Remark 11.** Let $X$ be a Polish space, $\mu \in \mathit{Prob}[X]$ be a probability measure, and $M \subseteq X$ be a Souslin set. Then there are Borel sets $B_l, B_u \subseteq X$ with

$$B_l \subseteq M \subseteq B_u \quad \text{and} \quad \mu(B_l) = \mu(B_u).$$

The statement can be found in Theorem 7.4.1 in [Bog07]. This cited theorem uses notions investigated by Definition 1.5.1 and Corollary 1.5.8 in [Bog07]. ⌟

Intuitively, Remark 11 states that every Souslin set can be approximated by a lower and an upper Borel set with respect to a given probability measure. The powerful connection between Souslin sets in a Polish space and the corresponding Borel sets appear at various points throughout this thesis.

**Outer measures of Souslin sets.** Recalling Remark 10 (2), a Souslin set may be not Borel and hence, it is not appropriate to evaluate a probability measure $\mu$ on some Polish space $X$ concerning a Souslin set $M$, i.e., the expression $\mu(M)$ makes no sense in general. Instead, one can however regard outer measures of Souslin sets (see Section 2.1). The remainder of this sections studies properties of outer measures of Souslin sets. One may skip the following material at a first reading of this section.

**Lemma 12.** *Let $X$ be a Polish space, $M \subseteq X$ be a Souslin set, and $\mu \in \mathit{Prob}[X]$ be a probability measure. Then the following two statements are equivalent:*

(1) $\mu^{\text{out}}(M) = 1$.

(2) *There is a Borel set $B \subseteq X$ with $B \subseteq M$ and $\mu(B) = 1$.*

*Proof.* It is easy to derive statement (1) from statement (2). The reverse implication follows directly from Remark 11. Indeed, according to this remark, there are Borel sets $B_l, B_u \subseteq X$ such that $B_l \subseteq M \subseteq B_u$ and $\mu(B_l) = \mu(B_u)$. Assuming $\mu^{\text{out}}(M) = 1$, we obtain $\mu(B_u) = 1$ and therefore, it holds $\mu(B_l) = \mu(B_u)1$. □

**Lemma 13.** *Let $X$ and $Y$ be Polish spaces and $f\colon X \to Y$ be a Borel function, $M_X \subseteq X$ be a Souslin set, and $\mu \in Prob[X]$. Define the probability measure $\nu \in Prob[Y]$ by*

$$\nu = f_\sharp(\mu)$$

*as well as the set $M_Y \subseteq Y$ by*

$$M_Y = f(M_X).$$

*Then the following statement holds:*

$$\mu^{\mathrm{out}}(M_X) = 1 \quad \textit{implies} \quad \nu^{\mathrm{out}}(M_Y) = 1.$$

*Proof.* According to Remark 10 (5), the set $M_Y$ is Souslin in $Y$. Thanks to Remark 11, there hence exist Borel sets $B_{Y,l}, B_{Y,u} \subseteq Y$ with

$$B_{Y,l} \subseteq M_Y \subseteq B_{Y,u} \quad \text{and} \quad \nu(B_{Y,l}) = \nu(B_{Y,u}).$$

It moreover holds

$$M_X \subseteq f^{-1}(f(M_X)) = f^{-1}(M_Y) \subseteq f^{-1}(B_{Y,u}).$$

Assuming $\mu^{\mathrm{out}}(M_X) = 1$, it hence follows $\mu(f^{-1}(B_{Y,u})) = 1$ and therefore,

$$\nu(B_{Y,l}) = \nu(B_{Y,u}) = \mu(f^{-1}(B_{Y,u})) = 1.$$

From this one easily derives $\nu^{\mathrm{out}}(M_Y) = 1$. $\square$

Roughly speaking, the previous lemma yields a result to transfer Souslin sets from one Polish space to another Polish space using a Borel function while preserving specific properties referring to a given probability measure.

**Lemma 14.** *Let $X$ be a Polish space and $M \subseteq X$ be a Souslin set. Then the set $P$ is Souslin in $Prob[X]$ where*

$$P = \{\mu \in Prob[X]\ ;\ \mu^{\mathrm{out}}(M) = 1\}.$$

*Proof.* Since $M$ is a Souslin set in $X$, there are a Polish space $Z$ as well as a Borel function $f\colon Z \to X$ such that $f(Z) = M$. Relying on Remarks 8 (1) and 10 (5), it suffices to show the identity

$$P = f_\sharp(Prob[Z]).$$

For the remainder of this proof let $\mu \in Prob[X]$.

Assume $\mu \in f_\sharp(Prob[Z])$ first. Let $\nu \in Prob[Z]$ be a probability measure with $\mu = f_\sharp(\nu)$. It clearly holds $\nu(Z) = 1$. Relying on Lemma 13, we thus obtain $\mu^{\text{out}}(f(Z)) = 1$. By Lemma 12, there hence exists a Borel set $B \subseteq X$ with $B \subseteq f(Z)$ and $\mu(B) = 1$. Since $f(Z) = M$, it follows $B \subseteq f(Z) = M$ and therefore, one has $\mu^{\text{out}}(M) = 1$. This shows the first inclusion $f_\sharp(Prob[Z]) \subseteq P$.

The rest of this proof is devoted to the more intricate inclusion $P \subseteq f_\sharp(Prob[Z])$. To this end assume $\mu \in P$, i.e., it holds $\mu^{\text{out}}(M) = 1$. Our task is to argue that there exists a probability measure $\nu \in Prob[Z]$ such that $f_\sharp(\nu) = \mu$. According to Remark 11 and as we additionally have the identity $\mu^{\text{out}}(M) = 1$, there are Borel sets $B_l, B_u \subseteq X$ with

$$B_l \subseteq M \subseteq B_u \quad \text{and} \quad \mu(B_l) = \mu(B_u) = 1.$$

Define

$$Z' = f^{-1}(B_l) \quad \text{and} \quad X' = B_l$$

as well as the function $f' \colon Z' \to X'$, $f'(z') = f(z')$. It is easy to see that $f'$ is well-defined, i.e., $f(z') \in X'$ for all $z' \in Z'$. The sets $X'$ and $Z'$ are equipped with the sigma algebras $\mathcal{B}_{X'}$ and $\mathcal{B}_{Z'}$, respectively, where $\mathcal{B}_{X'}$ is the induced sigma algebra of $X$ and accordingly, $\mathcal{B}_{Z'}$ is the induced sigma algebra of $Z$. It follows that $f'$ is measurable (with respect to the selected sigma algebras $\mathcal{B}_{X'}$ and $\mathcal{B}_{Y'}$).

As $X'$ and $Z'$ are also Borel sets in $X$ and $Z$, respectively, we moreover obtain that $X'$ and $Z'$ form standard Borel spaces (see Definition 12.5 and Corollary 13.4 in [Kec95]). By the definition of standard Borel spaces, there exists a metrics on $X'$ such that the resulting metric space is even a Polish spaces whose induced Borel sigma algebra coincide with the sigma algebra $\mathcal{B}_{X'}$. Analogously, there is a metric on $Z'$ such that $Z'$ becomes a Polish space whose induced Borel sigma algebra agrees with the sigma algebra $\mathcal{B}_{Z'}$.

Summarising the previous discussions, the sets $X'$ and $Z'$ constitute Polish spaces and the function $f'$ with domain $X'$ and codomain $Z'$ is Borel.

Observing the statement $B_l \subseteq M = f(Z)$, it is easy to see that the function $f'$ is surjective. By Remark 8 (2), the pushforward function of $f'$ is also surjective. As a consequence, there exists $\nu' \in Prob[Z']$ such that

$$\mu' = (f')_\sharp(\nu').$$

Here, $\mu' \in Prob[X']$ denotes the restriction of the probability measure $\mu$ onto $X'$, i.e., it holds $\mu'(B') = \mu(B')$ for every Borel set $B' \subseteq X'$ (that also yields a Borel set in $X$). We have $\mu'(X') = \mu(X') = 1$ and thus, the function $\mu'$ represents indeed a probability measure.

Since $Z' = f^{-1}(X') = (f')^{-1}(X')$, it holds

$$\nu'(Z') = \nu'((f')^{-1}(X')) = \mu'(X') = 1.$$

We can thus extend $\nu'$ to a probability measure $\nu \in \mathit{Prob}[Z]$ in a natural way: for every Borel set $B_Z \subseteq Z$ define

$$\nu(B_Z) = \nu'(B_Z \cap Z').$$

It remains to show $f_\sharp(\nu) = \mu$. For every Borel set $B \subseteq X$ we have

$$f^{-1}(B) \cap Z' = f^{-1}(B \cap B_l) = (f')^{-1}(B \cap B_l)$$

and therefore, we obtain

$$f_\sharp(\nu)(B) = \nu'(f^{-1}(B) \cap Z') = \nu'((f')^{-1}(B \cap B_l)) = \mu'(B \cap B_l) = \mu(B \cap B_l).$$

As $\mu(B_l) = 1$, for every Borel set $B \subseteq X$ it holds $f_\sharp(\nu)(B) = \mu(B)$. □

Let $X$ be a Polish space. Relying on properties of a Giry sigma algebra (see Section 2.1), it is straightforward to prove that the set $\{\mu \in \mathit{Prob}[X] \; ; \; \mu(B) = 1\}$ is Borel in $\mathit{Prob}[X]$ for every Borel set $B \subseteq X$. Lemma 14 provides an adaption of this statement for Souslin sets involving the concept of outer-measure functions. The given proof of Lemma 14 relies on the powerful notion of standard Borel spaces. Here, a measurable space $X$ with sigma algebra $\mathcal{B}$ is called a standard Borel space provided there exists a metric on $X$ such that the resulting metric space is Polish and the associated Borel sigma-algebra agrees with $\mathcal{B}$ (see Definition 12.5 in [Kec95]). An important and useful aspect is that this existing metric is not uniquely determined in general and can be often chosen in such a way that certain additional properties are fulfilled (see Section 13.A in [Kec95]). This fact also underlies Corollary 13.4 in [Kec95], which is essential for the previously presented proof.

**Lemma 15.** *Let $X$ and $Y$ be Polish spaces, $f\colon X \to \mathit{Prob}[Y]$ be a Borel function, $\mu \in \mathit{Prob}[X]$ be a probability measure, and $M \subseteq X \times Y$ be a Souslin set. The two statements below are equivalent:*

(1) $(\mu \rtimes f)^{\mathrm{out}}(M) = 1$.

(2) *There exists a Borel set $B_X \subseteq X$ such that*

$$\mu(B_X) = 1 \quad \textit{and} \quad (f(x))^{\mathrm{out}}(\mathit{Section}[M, x, \cdot]) = 1 \textit{ for all } x \in B_X.$$

*Proof.* Define the probability measure $\nu \in Prob[X \times Y]$ by

$$\nu = \mu \rtimes f.$$

(1) *implies* (2). Assume $\nu^{\text{out}}(M) = 1$. According to Lemma 12, there exists a Borel set $B \subseteq X \times Y$ with

$$B \subseteq M \quad \text{and} \quad \nu(B) = 1.$$

The set $Section[B, x, \cdot]$ is Borel in $Y$ for all $x \in X$ (see Section 2.1). It holds

$$\int f(x)(Section[B, x, \cdot])\, d\mu(x) = \nu(B) = 1$$

and therefore,

$$\int 1 - f(x)(Section[B, x, \cdot])\, d\mu(x) = 0.$$

Relying on a basic result from measure theory (see Lemma 8.2.8 in [Sch08]), there hence exists a Borel set $B_X \subseteq X$ with

$$\mu(B_X) = 1 \quad \text{and} \quad f(x)(Section[B, x, \cdot]) = 1 \text{ for all } x \in B_X.$$

Since $B \subseteq M$, for every $x \in X$ it holds $Section[B, x, \cdot] \subseteq Section[M, x, \cdot]$ and therefore, for every $x \in B_X$ one has $(f(x))^{\text{out}}(Section[M, x, \cdot]) = 1$.

(2) *implies* (1). Suppose a Borel set $B_X \subseteq X$ as in statement (2). Applying Remark 11, there are Borel sets $B_l, B_u \subseteq X \times Y$ with

$$B_l \subseteq M \subseteq B_u \quad \text{and} \quad \nu(B_l) = \nu(B_u).$$

For every $x \in X$ one has $Section[M, x, \cdot] \subseteq Section[B_u, x, \cdot]$ and therefore, for every $x \in B_X$ it holds

$$f(x)(Section[B_u, x, \cdot]) = 1.$$

As $\mu(B_X) = 1$, we obtain

$$\nu(B_l) = \nu(B_u) = \int f(x)(Section[B_u, x, \cdot])\, d\mu(x) = 1.$$

It follows $\nu^{\text{out}}(M) = 1$ and hence, we are done. $\square$

Consequently, the outer measure of the set $M$, which is a subset of the plane $X \times Y$, is one precisely when the vertical outer measures of the one-dimensional slices $\mathit{Section}[M, x, \cdot]$ are almost surely one. The characterisation of the previous lemma is also in the spirit of the disintegration theorem (see Section 2.1). Moreover, the usefulness of Lemma 15 also results from the disintegration theorem. Indeed, thanks to the disintegration theorem, every probability measure $\nu$ on the product space $X \times Y$ can be represented in terms of a semi-product measure, i.e., there are a probability measure $\mu \in \mathit{Prob}[X]$ as well as a Borel function $f\colon X \to \mathit{Prob}[Y]$ such that $\nu = \mu \rtimes f$.

## 2.3 Measurable-selection principles

At various points throughout this thesis we are confronted with the task of showing the existence of Borel functions with specific properties. To do so, we typically rely on measurable-selection principles [Wag80, HPV81, AB06, Bog07]. The general approach concerning the application of a measurable-selection principle is illustrated by means of the following example investigating a basic mathematical problem.

**Example 16.** Suppose two Polish spaces $X$ and $Y$ as well as a surjective function $g\colon Y \to X$. Of course, using the surjectivity, it is easy to provide a right inverse of the function $g$ (accepting the axiom of choice), i.e., a function $f\colon X \to Y$ with $g(f(x)) = x$ for all $x \in X$. The problem becomes more difficult if one wants a right inverse that is Borel in addition as the property of being a Borel function refers to specific subsets of $X$ and $Y$. To overcome this issue, one may proceed as follows. First of all, we introduce a set-valued function $F$ mapping every element in $X$ to a subset of $Y$ such that for every $x \in X$,

$$F(x) = \{y \in Y\,;\, g(y) = x\}.$$

Intuitively, for every $x \in X$ the set $F(x)$ consists of all the possible values of a function $f\colon X \to Y$ evaluated at $x$ such that $g(f(x)) = x$. We have that there exists an Borel right inverse of $g$ precisely when $F$ admits a Borel selection, i.e., a Borel function $f\colon X \to Y$ such that $f(x) \in F(x)$ for all $x \in X$. At this point, measurable-selection principles come into play as they provide conditions on $F$ guaranteeing the existence of such a Borel selection. Obviously, a necessary condition on $F$ is that the set $F(x)$ is not empty for all $x \in X$ which is, however, fulfilled in our example since the function $g$ is required to be surjective. ⌟

For every Polish spaces $X$ and $Y$ we use $F\colon X \rightsquigarrow Y$ as a shorthand notion for $F\colon X \to 2^Y$, i.e., $F$ is a set-valued function assigning a subset of $Y$ to every element of $X$. Every set-valued function $F\colon X \to 2^Y$ induces a relation $Rel[F]$ over $X$ and $Y$ as follows:

$$Rel[F] = \{\langle x, y\rangle \in X \times Y \; ; \; y \in F(x)\}.$$

Vice versa, every relation $R$ over $X$ and $Y$ can be translated into a set-valued function $F_R$ such that $Rel[F_R] = R$. The presentation of measurable-selection principles in the literature typically involves set-valued functions rather than relations. The reason becomes clear below when discussing prominent measurable-selection principles below.

**Definition 17.** Let $X$ and $Y$ be Polish spaces and $F\colon X \rightsquigarrow Y$ be a set-valued function. We call a function $f\colon X \to Y$ a *Borel selection of $F$* provided $f$ is a Borel function and

$$f(x) \in F(x) \text{ for all } x \in X,$$

i.e., it holds $Graph[f] \subseteq Rel[F]$. Moreover, for every probability measure $\mu \in Prob[X]$ a function $f\colon X \to Y$ is called a *Borel $\mu$-selection of $F$* provided $f$ is a Borel function and there exists a Borel set $B \subseteq X$ such that

$$\mu(B) = 1 \quad \text{and} \quad f(x) \in F(x) \text{ for all } x \in B,$$

i.e., one has $Graph[f] \cap (B \times Y) \subseteq Rel[F]$. ⌟

Our brief survey on measurable-measurable selection principles starts with the most prominent theorem in this area:

**Theorem 18** (Theorem 6.9.3 in [Bog07], Kuratowski and Ryll-Nardzewski)**.** *Let $X$ and $Y$ be Polish spaces and $F\colon X \rightsquigarrow Y$ be a set-valued function. Then there exists a Borel selection of $F$ if the following two statements hold:*

(1) *The set $\{x \in X \; ; \; F(x) \cap O_Y \neq \varnothing\}$ is Borel in $X$ for every open set $O_Y \subseteq Y$.*

(2) *The set $F(x)$ is not empty and closed in $Y$ for every $x \in X$.*

⌟

**Remark 19.** In fact, the theorem of Kuratowski and Ryll-Nardzewski is even stronger than the previously stated theorem (see Corollary 6.9.3 in [Bog07]). Consider two Polish spaces $X$ and $Y$ and a set-valued function $F\colon X \rightsquigarrow Y$ that satisfies conditions (1) and (2) of Theorem 18. Then there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of Borel functions such that for every $x \in X$ the following statement holds:

$$\text{the set } \{f_n(x) \; ; \; n \in \mathbb{N}\} \text{ is dense in } F(x),$$

i.e., the set $F(x)$ and the topological closure of the set $\{f_n(x)\ ;\ n \in \mathbb{N}\}$ in $X$ are the same, in particular, for every $n \in \mathbb{N}$ the function $f_n$ yields a Borel selection of $F$. In this context, the sequence $(f_n)_{n\in\mathbb{N}}$ is also called a *Castaing representation of* $F$. For every $x \in X$ such that the set $F(x)$ is finite we have the identity $F(x) = \{f_n(x)\ ;\ n \in \mathbb{N}\}$. ⌟

Defining an appropriate sigma algebra on the powerset of $Y$, condition (1) in Theorem 18 provides a measurability requirement on the set-valued function $F$. To be more precise, define the hit sigma algebra on the powerset $2^Y$ as the smallest sigma algebra on $2^Y$ containing the set $\{M_Y \subseteq Y\ ;\ M_Y \cap O_Y \neq \varnothing\}$ for every open set $O_Y \subseteq B_Y$. Intuitively, $\{M_Y \subseteq Y\ ;\ B_Y \cap O_Y \neq \varnothing\}$ consists of those subsets of $Y$ that hit the given set $O_Y$. Then, condition (1) is equivalent to the requirement that $F$ is measurable with respect to the hit sigma algebra. Indeed, for every open set $O_Y \subseteq Y$ it holds

$$F^{-1}(\{B_Y \subseteq Y\ ;\ B_Y \cap O_Y \neq \varnothing\}) = \{x \in X\ ;\ F(x) \cap O_Y \neq \varnothing\}.$$

Condition (2) in Theorem 18 constrains the possible image sets of a set-valued functions. As a sanity check consider a set-valued function $F\colon X \rightsquigarrow Y$ such that the set $F(x)$ is a singleton for every $x \in X$, i.e., $F$ can be viewed as a function $f\colon X \to Y$, in particular, it holds $F(x) = \{f(x)\}$ for every $x \in X$. As the open sets in $Y$ generate the Borel sigma algebra on $Y$ (see Section 2.1), it is easy to see that $F$ satisfies condition (1) precisely when the function $f$ is Borel. Consequently, $F$ has a Borel selection iff $f$ is a Borel function. In this sense, the conditions (1) and (2) can be seen as adequate measurability notions for set-valued functions. Moreover, the conditions in Theorem 18 are similar to the requirements on the stochastic model called non-deterministic labelled Markov processes [DTW12, Wol12] that we precisely discuss in Chapter 7.

**Example 20.** We continue our discussions of Example 16. Relying on Theorem 18, we present a condition on a surjective function $g$ such that there exists a Borel right inverse. Recall, the domain and the codomain of the function $g$ are given by $Y$ and $X$, respectively. For every open set $O_Y \subseteq Y$ it holds

$$\{x \in X\ ;\ F(x) \cap O_Y \neq \varnothing\} = \{x \in X\ ;\ g(y) = x \text{ for some } y \in O_Y\} = g(O_Y).$$

Since the function $g$ is surjective, basic knowledge of an undergraduate analysis course suffices to show that the set $\{x \in X\ ;\ F(x) \cap O_Y \neq \varnothing\}$ is Borel in $X$ provided both the topological space $Y$ is compact and the function $g$ is continuous. Moreover, for every $x \in X$ one has

$$F(x) = g^{-1}(\{x\}).$$

and consequently, the set $F(x)$ is closed in $Y$ provided the function $g$ is continuous. Thanks to the Theorem 18, summarising our previous discussions, every surjective continuous function with compact domain admits a Borel right inverse. ⌟

In the context of this thesis, we almost only work with Borel selections with respect to a probability measure. In fact, the probability measure of interest is naturally given in many situations and moreover, known measurable-selection principles involve surprisingly weak assumptions on the given set-valued function.

**Theorem 21** (Theorem 6.9.5 in [Bog07])**.** *Let $X$ and $Y$ be Polish spaces, $\mu \in Prob[X]$ be a probability measure, and $F\colon X \rightsquigarrow Y$ be a set-valued function. Then there exists a Borel $\mu$-selection of $F$ if the following two properties hold:*

(1) *The set $Rel[F]$ is Souslin in $X \times Y$.*

(2) *There is a Borel set $B \subseteq X$ such that*

$$\mu(B) = 1 \quad \textit{and} \quad F(x) \neq \varnothing \textit{ for all } x \in B.$$

⌟

At the end of this section we provide some more detailed comments on Theorem 6.9.5 in [Bog07], which acts as a reference for Theorem 21. Observe, provided there exists a Borel $\mu$-selection of a given set-valued function, condition (2) in Theorem 21 is necessarily fulfilled. Thanks to the rich theory on Souslin sets and their various closure properties (see Section 2.2, in particular, Remark 10), condition (1) of the previously stated measurable-selection principle fits well with the mathematical framework of this thesis. The theory to be developed throughout the next chapters shows that this Souslin requirement appears natural in the context of modelling and analysis of stochastic transition systems.

**Example 22.** We continue our discussions concerning Example 16. Assume that the given function $g$ is Borel. According to Remark 10 (6), the set $Graph[g]$ is Souslin in $Y \times X$. Since $Rel[F] = Graph[g]^{-1}$, we have that the set $Rel[F]$ is Souslin in $X \times Y$. Thanks to Theorem 21, for every $\mu \in Prob[X]$ there exists a Borel function $f\colon X \to Y$ and a Borel set $B \subseteq X$ such that $\mu(B) = 1$ and $g(f(x)) = x$ for all $x \in B$. Consequently, summarising the previous discussions, every surjective Borel function admits a Borel right inverse with respect to a given probability measure. ⌟

**Remark 23.** Consider Polish spaces $X$ and $Y$ and let $F\colon X \rightsquigarrow Y$ be a set-valued function. If $F$ satisfies the conditions (1) and (2) of Theorem 18, then the set $Rel[F]$ is Borel in $X \times Y$ by Theorem 18.6 in [AB06], in particular, the set $Rel[F]$ is Souslin in $X \times Y$ by Remark 10 (2). Consequently, every set-valued function satisfying the requirements of Theorem 18 also fulfills the conditions of the measurable-selection principle given by Theorem 21.

Following the introductory discussions of Chapter 18 in [AB06], a set-valued function $F$ may not satisfy condition (1) of Theorem 18 even if the two conditions below are fulfilled: the set $Rel[F]$ is Borel in $X \times Y$ and the set $F(x)$ is closed in $Y$ for every $x \in X$. Consequently, the measurability assumptions on $F$ in Theorem 21 are strictly weaker than the requirements in Theorem 18 (see also Examples 2.1 (i) and (ii) in [Wag80]). ⌟

In view on Remark 23, we also refer to the article [Iof79] that studies necessary and sufficient conditions on a set-valued function $F\colon X \rightsquigarrow Y$ such its induced relation $Rel[F]$ forms a Souslin set in $X \times Y$.

**Comments on Theorem 21.** We summarise concepts and references to derive Theorem 21 in the stated form. In fact, the argument needs some care as well as insights concerning functions that are measurable with respect to sigma algebras generated by Souslin sets. We emphasis that Theorem 21 is a very important tool throughout this thesis that is why the following discussion is conducted carefully.

The precise formulation of Theorem 6.9.5 in [Bog07] is recalled first. For this purpose let $X$ and $Y$ be Polish spaces. Moreover, denote the sigma algebra on $X$ that is generated by all the Souslin sets in $X$ by $\mathcal{B}_{X,\mathrm{Sou}}$. In general, recalling the discussions in Section 2.2, the family consisting of all Souslin sets in an uncountable Polish space does not form a sigma algebra. Consider a set-valued function $F'\colon X \rightsquigarrow Y$ such that $Rel[F']$ is Souslin in $X \times Y$ and $F'(x) \neq \varnothing$ for every $x \in X$. Then Theorem 6.9.5 in [Bog07] states the existence of a function $f'\colon X \to Y$ satisfying

$$f'(x) \in F'(x) \text{ for every } x \in X$$

and

$$(f')^{-1}(B_Y) \in \mathcal{B}_{X,\mathrm{Sou}} \text{ for every Borel set } B_Y \subseteq Y.$$

In what follows we derive Theorem 21 from Theorem 6.9.5 in [Bog07]. To this end let $X$ and $Y$ be Polish spaces and $F\colon X \rightsquigarrow Y$ be a set-valued function such that $Rel[F]$ is Souslin in $X \times Y$. Moreover, let $\mu \in Prob[X]$ be a probability measure and $B \subseteq X$ be a Borel set so that $\mu(B) = 1$ and $F(x) \neq \varnothing$ for all $x \in B$.

The first step is straightforward. Let $y$ be an arbitrary element of $Y$ and define $F' \colon X \rightsquigarrow Y$,

$$F'(x) = \begin{cases} F(x), & \text{if } x \in B, \\ \{y\}, & \text{if } x \notin B. \end{cases}$$

We have $\mathit{Graph}[F'] = (\mathit{Graph}[F] \cap (B \times Y)) \cup ((X \setminus B) \times \{y\})$. According to Remark 10, since $\mathit{Graph}[F]$ is Souslin in $X \times Y$, it follows that $\mathit{Graph}[F']$ is Souslin in $X \times Y$. Moreover, for every $x \in X$ the set $F'(x)$ is not empty. We can hence apply the previously recapitulated Theorem 6.9.5 in [Bog07] and thus, there exists a function $f' \colon X \to Y$ with $f'(x) \in F'(x)$ for every $x \in X$ and such that for every Borel set $B_Y \subseteq Y$ it holds $(f')^{-1}(B_Y) \in \mathcal{B}_{X,\mathrm{Sou}}$. To finish the argument, it suffices to show that there are a Borel function $f \colon X \to Y$ and a Borel set $B' \subseteq X$ such that $\mu(B') = 1$ and $f(x) = f'(x)$ for every $x \in B'$. Indeed, it then follows $\mu(B \cap B') = 1$ and $f(x) \in F(x)$ for every $x \in B \cap B'$, i.e., $f$ is a Borel $\mu$-selection of the set-valued function $F$.

Let $\mathcal{B}_{X,\mu}$ be the smallest sigma algebra on $X$ that consists of all sets $M \subseteq X$ with the following property: there are Borel sets $B_l, B_u \subseteq X$ such that $B_l \subseteq M \subseteq M_u$ and $\mu(B_l) = \mu(B_u)$. According to Remark 11, the sigma algebra $\mathcal{B}_{X,\mu}$ consists of every Souslin set in $X$ and therefore, it holds $\mathcal{B}_{X,\mathrm{Sou}} \subseteq \mathcal{B}_{X,\mu}$. The function $f'$ is hence measurable with respect to the newly introduced sigma algebra $\mathcal{B}_{X,\mu}$

Without loss of generality we can safely assume that $Y = \mathbb{R}$. Indeed, if $Y$ is countable, the elements of $Y$ can be easily identified by a real number and moreover, provided $Y$ is uncountable, one can rely on the fact that every uncountable Polish space is Borel isomorphic to $\mathbb{R}$ (see Section 2.1). Since $Y = \mathbb{R}$, we are in the setting of Proposition 2.1.11 in [Bog07] (see also Definitions 1.5.1 and 2.1.10 and Corollary 1.5.8 in [Bog07]): as a consequence, there exists a Borel function $f \colon X \to Y$ and a Borel set $B' \subseteq X$ such that $\mu(B') = 1$ and $f(x) = f'(x)$ for every $x \in B'$. From this we can finally derive Theorem 21.

## 2.4 Barycentres and convex hulls

In the context of stochastic transition systems introduced in the next chapter, the material of this section yields the basis for the definition of the combined-transition relation that in turn provides the basic ingredient for the notion of schedulers. Assume a measurable space $X$ as well as probability measures $\mu_0, \mu_1, \mu_2, \ldots$ on $X$. A convex combination of these probability measures is typically defined point-wise. More precisely, a probability measure $\mu$ on $X$ is called a convex combination if there are real numbers $r_0, r_1, r_2, \ldots$ between zero and one such that

$$r_0 + r_1 + r_2 + \ldots = 1$$

and so that every for measurable set $B \subseteq X$ it holds

$$\mu(B) = r_0 \cdot \mu_0(B) + r_1 \cdot \mu_1(B) + r_2 \cdot \mu_2(B) + \ldots .$$

The following material goes one step further and provides a generalisation of the sketched concept. In fact, we want to consider convex combinations of a potentially uncountable number of probability measures. Based on this extended notion, we introduce the concept of convex hulls of sets of probability measures and show that this notion yields a closure operator on the family of all Souslin sets. In particular, it is shown that taking the convex hull of a Souslin set of probability measures twice yields the same result as if the convex-hull operator was applied once.

The following material is inspired by the notions and results of Chapters X.29 and XI.33 in [DM88]. Inspecting the setting in [DM88], the mentioned results are not directly applicable for our purposes as the cited book concentrates on locally convex topological vector spaces. In fact, relying on deep and fundamental results from functional analysis, one can identify the set of all probability measures $Prob[X]$ on a Polish space $X$ with a convenient topological vector space (see Section 15.1, in particular, Theorem 15.1, in [AB06]). Using this identification of spaces, one can then transfer notions and results from vector spaces to our setting. This indeed yields an alternative and equivalent approach to the following material. However, to make the result of this thesis more accessible and self-contained, our presentation avoids this involved machinery and develops direct arguments sufficient for our mathematical setting. Interestingly, the proof of this section's main result uses a measurable-selection principle (see Section 2.3).

**Definition 24.** Let $X$ be a measurable space. Define the function

$$Barycen\colon Prob[Prob[X]] \to Prob[X]$$

as follows: for every $\beta \in Prob[Prob[X]]$ and measurable set $B \subseteq X$ let

$$Barycen(\beta)(B) = \int \mu(B)\, d\beta(\mu).$$

For every $\beta \in Prob[Prob[X]]$ we refer to $Barycen(\beta)$ as the *barycentre of $\beta$*. ⌟

The barycentre $\mathit{Barycen}(\beta)$ of a probability measure $\beta \in \mathit{Prob}[\mathit{Prob}[X]]$ is intuitively derived by the following two-step experiment: in a first step a probability measure $\mu$ is sampled according to $\beta$ and in a second step an element of $X$ is sampled according to $\mu$. The notion in Definition 24 is well-defined. To see this, we first observe that for every Borel set $B \subseteq X$ the function $\mathit{probeval}_B \colon \mathit{Prob}[X] \to [0,1]$, $\mathit{probeval}_B(\mu) = \mu(B)$ is Borel (see the definition of the Giry sigma algebra in Section 2.1). Moreover, for every probability measure $\beta \in \mathit{Prob}[\mathit{Prob}[X]]$ the definition of $\mathit{Barycen}(\beta)$ indeed yields a probability measure on $X$. The latter is a consequence of the monotone convergence theorem (also known as Beppo Levi's theorem, see Folgerung 8.2.4 in [Sch08]).

**Example 25.** Let $X$ be a Polish space and $\mu_0, \mu_1, \mu_2, \ldots \in \mathit{Prob}[X]$ be probability measures. Consider real numbers $r_0, r_1, r_2, \ldots \in [0,1]$ with $r_0 + r_1 + r_2 + \ldots = 1$. Define the probability measure $\beta \in \mathit{Prob}[\mathit{Prob}[X]]$ by

$$\beta(\{\mu_0\}) = r_0, \quad \beta(\{\mu_1\}) = r_1, \quad \beta(\{\mu_2\}) = r_2, \quad \ldots.$$

Then for every Borel set $B \subseteq X$ it holds

$$\mathit{Barycen}(\beta)(B) = r_0 \cdot \mu_0(B) + r_1 \cdot \mu_1(B) + r_2 \cdot \mu_2(B) + \ldots.$$

Consequently, $\mathit{Barycen}(\beta)$ corresponds to the introductory sketched point-wise definition of convex combinations of a countable number of probability measures. ⌟

The previous example shows that Definition 24 yields a conservative extension of the classical notion of a convex combination. The convex hull of a subset in an Euclidean space is defined as the set of all convex combinations of points in that subset. The corresponding notion for sets of probability measures uses the previously introduced notion of barycentres:

**Definition 26.** Let $X$ be a measurable space. For every set $P \subseteq \mathit{Prob}[X]$ the *convex hull of P* is defined by

$$\mathit{Conv}[P] = \mathit{Barycen}(\{\beta \in \mathit{Prob}[\mathit{Prob}[X]] \; ; \; \beta^{\mathrm{out}}(P) = 1\}),$$

i.e., the convex hull consists of precisely those probability measures that can be represented as an appropriate barycentre. A set $P \subseteq \mathit{Prob}[X]$ is called *convex* if $P = \mathit{Conv}[P]$. ⌟

One typically expects that the notion of a convex hull forms a closure operator, i.e., an operator that is extensive, monotonically increasing, and idempotent. In the remainder of this section we show that all these properties are fulfilled for the notion as introduced in Definition 26 when restricting to the Souslin subsets of the Polish space under consideration. Our main theorem is as follows:

**Theorem 27.** *Let $X$ be a Polish space. The convex hull $Conv[Q]$ of every Souslin set $Q \subseteq Prob[X]$ is Souslin in $Prob[X]$. Moreover, the following three statements hold:*

(1) *$Q \subseteq Conv[Q]$ for every Souslin set $Q \subseteq Prob[X]$.*

(2) *$Q \subseteq Q'$ implies $Conv[Q] \subseteq Conv[Q']$ for every Souslin sets $Q, Q' \subseteq Prob[X]$.*

(3) *$Conv[Q] = Conv[Conv[Q]]$ for every Souslin set $Q \subseteq Prob[X]$.*

*In particular, the notion of convex hulls of probability measures induces a closure operator on the Souslin subsets of $Prob[X]$.* ⌟

Statements (1) and (2) can be easily derived from the definitions and even hold for arbitrary subsets of $Prob[X]$. Indeed, for every set $Q \subseteq Prob[X]$ and probability measure $\mu \in Q$ the probability measure $\beta \in Prob[Prob[X]]$ with $\beta = Dirac[\mu]$ satisfies $\beta^{\mathrm{out}}(Q) = 1$ and thus, we derive $\mu \in Conv[Q]$, which immediately yields statement (1). The monotonicity of the outer-measure function directly shows statement (2). Moreover, putting (1) and (2) together, we also obtain the first inclusion in (3), i.e., the inclusion $Conv[Q] \subseteq Conv[Conv[Q]]$ for every set $Q \subseteq Prob[X]$. The reverse inclusion is intricate and our proof below indeed needs the Souslin requirement for the subset of probability measures under consideration. Let $Q \subseteq Prob[X]$ be a Souslin set. To show the inclusion $Conv[Conv[Q]] \subseteq Conv[Q]$, the challenging task is to argue that for every $\beta \in Prob[Prob[X]]$ with

$$\beta^{\mathrm{out}}(Conv[Q]) = 1$$

there exists $\beta' \in Prob[Prob[X]]$ with

$$(\beta')^{\mathrm{out}}(Q) = 1 \quad \text{and} \quad Barycen(\beta) = Barycen(\beta').$$

The desired probability measure $\beta'$ intuitively represents a reweighting of the elements in $Q$ that is additionally compatible with the given $\beta$ by means of their respective barycentres. At this point of the proof, we need the Souslin assumption on the set $Q$ to rely on the measurable-selection principle given by Theorem 21 for a specific set-valued function depending on $Q$.

**Proof of section's main result.** For our proof of Theorem 27 we need the following basic observation referring to the function introduced in Definition 24.

**Lemma 28.** *For every Polish space the function Barycen is Borel.*

*Proof.* Let $X$ be a Polish space. Applying Theorem 15.13 in [AB06], for every Borel set $B \subseteq X$ the function $f_B \colon Prob[Prob[X]] \to [0,1]$,

$$f_B(\beta) = Barycen(\beta)(B)$$

is Borel. The claim thus immediately follows from Remark 5. We remark that the referred Theorem 15.13 in [AB06] even shows that for every Polish space $Y$ and Borel function $f \colon Y \to [0,1]$ the function $g_f \colon Prob[Y] \to [0,1]$,

$$g_f(\mu_Y) = \int f(y)\, d\mu_Y(y)$$

is Borel. $\square$

*Proof of Theorem 27.* Let $Q \subseteq Prob[X]$ be a Souslin set. Applying Lemmas 14 and 28 as well as Remark 10 (5), the set $Conv[Q]$ is Souslin in $Prob[X]$. Given the discussions after the statement of Theorem 27, it remains to show the inclusion $Conv[Conv[Q]] \subseteq Conv[Q]$. To this end let $\mu \in Prob[X]$ be a probability measure such that $\mu \in Conv[Conv[Q]]$. Our task is to show $\mu \in Conv[Q]$. By the definition of convex hulls, there exists a probability measure $\beta \in Prob[Prob[X]]$ with

$$\beta^{\mathrm{out}}(Conv[Q]) = 1 \quad \text{and} \quad \mu = Barycen(\beta).$$

Define the set-valued function $F \colon Prob[X] \rightsquigarrow Prob[Prob[X]]$,

$$F(\mu') = \{\beta' \in Prob[Prob[X]] \,;\, Barycen(\beta') = \mu' \text{ and } (\beta')^{\mathrm{out}}(Q) = 1\}.$$

Then there exists a Borel $\beta$-selection of $F$. Indeed, the assumptions of Theorem 21 are fulfilled that is shown in the next two proof paragraphs.

Let us first justify that the set $Rel[F]$ is Souslin in $Prob[X] \times Prob[Prob[X]]$. According to Lemma 28, the function $Barycen$ is Borel. By Remark 10 (6), the set $Graph[Barycen]$ is hence Borel in $Prob[Prob[X]] \times Prob[X]$. Moreover, thanks to Lemma 14, the set $\{\beta' \in Prob[Prob[X]] \,;\, (\beta')^{\mathrm{out}}(Q) = 1\}$ is Souslin in $Prob[Prob[X]]$. Putting things together, the set $Rel[F]$ is Souslin in $Prob[X] \times Prob[Prob[X]]$.

Exploiting $\beta^{\mathrm{out}}(Conv[Q]) = 1$, Lemma 12 yields a Borel set $P' \subseteq Prob[X]$ such that $P' \subseteq Conv[Q]$ and $\beta(P') = 1$. For every $\mu' \in P'$ the set $F(\mu')$ is not empty that can be seen as follows: since $\mu' \in P'$ and $P' \subseteq Conv[Q]$, there exists $\beta' \in Prob[Prob[X]]$ with $(\beta')^{\mathrm{out}}(Q) = 1$ and $Barycen(\beta') = \mu'$.

We are hence in the situation of Theorem 21. Consequently, there exist a Borel function $f\colon Prob[X] \to Prob[Prob[X]]$ and Borel set $P_\beta \subseteq Prob[X]$ such that

$$\beta(P_\beta) = 1 \quad \text{and} \quad f(\mu') \in F(\mu') \text{ for all } \mu' \in P_\beta.$$

For the following definition recall the concept of the post operator from Section 2.1. Introduce the probability measure $\beta' \in Prob[Prob[X]]$,

$$\beta' = Post[\beta, f],$$

i.e., for every Borel set $P \subseteq Prob[X]$ we have

$$\beta'(P) = \int f(\mu')(P)\, d\beta(\mu')$$

To conclude $\mu \in Conv[Q]$, it remains to show $Barycen(\beta') = \mu$ and $(\beta')^{\text{out}}(Q) = 1$. We attend the latter claim first.

Let $P \subseteq Prob[Prob[X]]$ be a Borel set such that $Q \subseteq P$. For every $\mu' \in P_\beta$ it holds $(f(\mu'))^{\text{out}}(Q) = 1$ that in turn implies $f(\mu')(P) = 1$. Since $\beta(P_\beta) = 1$, it follows

$$\beta'(P) = \int f(\mu')(P)\, d\beta(\mu') = \int_{P_\beta} 1\, d\beta(\mu') = \beta(P_\beta) = 1.$$

This yields $(\beta_f)^{\text{out}}(Q) = 1$.

It remains to show $Barycen(\beta') = \mu$. Let $B \subseteq X$ be a Borel set. Applying Fubini's theorem (see Section 2.1), we obtain

$$\int \mu'(B)\, d\beta'(\mu') = \int \left( \int \mu''(B)\, df(\mu')(\mu'') \right) d\beta(\mu').$$

From this we directly derive

$$Barycen(\beta')(B) = \int Barycen(f(\mu'))(B)\, d\beta(\mu').$$

For every $\mu' \in P_\beta$ it holds $f(\mu') \in F(\mu')$ and therefore, one has $Barycen(f(\mu')) = \mu'$. Using $\beta(P_\beta) = 1$, we consequently obtain

$$Barycen(\beta')(B) = \int \mu'(B)\, d\beta(\mu') = Barycen(\beta)(B) = \mu(B).$$

This finally completes our argument. □

## 2.5 Couplings of probability measures

Intuitively, a coupling places two a-priori unrelated probability measures in the same probability space by exhibiting an adequate witness distribution over pairs:

**Definition 29.** Let $X$ and $Y$ be measurable spaces and $\mu_X \in \mathit{Prob}[X]$ and $\mu_Y \in \mathit{Prob}[Y]$ be probability measures. A probability measure $\mu \in \mathit{Prob}[X \times Y]$ is called a *coupling of* $(\mu_X, \mu_Y)$ if the following two statements hold:

(1) $\mu(B_X \times Y) = \mu_X(B_X)$ for every measurable set $B_X \subseteq X$.

(2) $\mu(X \times B_Y) = \mu_Y(B_Y)$ for every measurable set $B_Y \subseteq Y$.

Obviously, the product measure of $\mu_X$ and $\mu_Y$, i.e., the measure given by $\mu_X \otimes \mu_Y$, represents a coupling of $(\mu_X, \mu_Y)$ to which we refer as the *independent coupling of* $(\mu_X, \mu_Y)$. ⌟

It is important to realise that the requirements for a coupling do not explicitly constrain the probabilities of the form $\mu(B_X \times B_Y)$ where $B_X \subseteq X$ and $B_Y \subseteq Y$ are measurable sets such that $B_X \neq X$ and $B_Y \neq Y$. This fact basically constitutes the crux of many applications of couplings. Intuitively, conditions (1) and (2) in the previous definition only require that the marginals of a coupling coincide with the given probability measures. In other words, defining the two Borel functions $\rho_X \colon X \times Y \to X$, $\rho_X(x, y) = x$ and $\rho_Y \colon X \times Y \to Y$, $\rho_Y(x, y) = y$, a probability measure $\mu \in \mathit{Prob}[X \times Y]$ is a coupling of $(\mu_X, \mu_Y)$ precisely when it holds

$$(\rho_X)_\sharp(\mu) = \mu_X \quad \text{and} \quad (\rho_Y)_\sharp(\mu) = \mu_Y.$$

In case the sets $X$ and $Y$ are countable, a probability measure $\mu$ is a coupling of $(\mu_X, \mu_Y)$ iff for every $x \in X$ and $y \in Y$ one has the following two identities

$$\sum_{y' \in Y} \mu(\{x\} \times \{y'\}) = \mu_X(\{x\}) \quad \text{and} \quad \sum_{x' \in X} \mu(\{x'\} \times \{y\}) = \mu_Y(\{y\}).$$

**Example 30.** Consider two (unfair) coins $\mathrm{Coin}_X$ and $\mathrm{Coin}_Y$. We assume that $\mathrm{Coin}_X$ lands on its head with probability $p_X$ and accordingly, that $\mathrm{Coin}_Y$ lands on its head with probability $p_Y$. The latter informal description is formalised by means of the two measurable spaces $X = \{\mathrm{Head}_X, \mathrm{Tail}_X\}$ and $Y = \{\mathrm{Head}_Y, \mathrm{Tail}_Y\}$ as well as the corresponding probability measures $\mu_X \in \mathit{Prob}[X]$ and $\mu_Y \in \mathit{Prob}[Y]$ with

$$\mu_X(\{\mathrm{Head}_X\}) = p_X \quad \text{and} \quad \mu_Y(\{\mathrm{Head}_Y\}) = p_Y.$$

Clearly, it follows $\mu_X(\{\mathrm{Tail}_X\}) = 1 - p_X$ and $\mu_Y(\{\mathrm{Tail}_Y\}) = 1 - p_Y$.

We consider the experiment of tossing these coins simultaneously now. For the moment, there are no further information concerning the interplay of the given coins. In particular, the events stating that the coins $\mathrm{Coin}_X$ and $\mathrm{Coin}_Y$ land on their heads, respectively, may be not (stochastically) independent of each other. If only this information is given, every coupling of $(\mu_X, \mu_Y)$ represents a convenient mathematical model for the sketched random experiment. Indeed, if $\mu$ is a coupling of $(\mu_X, \mu_Y)$, then it holds $\mu(\{\mathrm{Head}_X\} \times Y) = p_X$ and $\mu(X \times \{\mathrm{Head}_Y\}) = p_Y$, i.e., the occurrence of the event that $\mathrm{Coin}_X$ lands on its head happens with probability $p_X$ and accordingly for the second coin $\mathrm{Coin}_Y$.

In the present case, exploiting the finiteness of the sets $X$ and $Y$, a probability measure $\mu$ on $X \times Y$ forms a coupling of $(\mu_X, \mu_Y)$ precisely when the following four identities hold:

$$a + b = p_X, \quad c + d = 1 - p_X, \quad a + c = p_Y, \quad b + d = 1 - p_Y$$

where $a, b, c, d \in [0, 1]$ are defined by

$$\begin{aligned} a &= \mu(\{\mathrm{Head}_X\} \times \{\mathrm{Head}_Y\}), & b &= \mu(\{\mathrm{Head}_X\} \times \{\mathrm{Tail}_Y\}), \\ c &= \mu(\{\mathrm{Tail}_X\} \times \{\mathrm{Head}_Y\}), & d &= \mu(\{\mathrm{Tail}_X\} \times \{\mathrm{Tail}_Y\}). \end{aligned}$$

For instance, assuming $p_X \geq p_Y$, one can consider a coupling of $(\mu_X, \mu_Y)$ such that the event $\{\langle \mathrm{Head}_X\ \mathrm{Tail}_Y \rangle\}$ happens with probability zero intuitively meaning that it is never the case that $\mathrm{Coin}_X$ lands on its head while $\mathrm{Coin}_Y$ lands on its tail. The coins are in particular not tossed independently of each other here. Using the notions as before, the following assignment for $a$, $b$, $c$, and $d$ yields the uniquely defined coupling:

$$a = p_Y, \quad b = p_X - p_Y, \quad c = 0, \quad d = 1 - p_X.$$

Note, $b \geq 0$ is implied by the assumption $p_X \geq p_Y$. ⌟

The previous example clarifies that even for discrete spaces the number of existing couplings may be uncountable. A coupling of two probability measures is uniquely determined by the corresponding product measure provided one of the two given probability measures is a Dirac distribution:

**Remark 31.** Let $X$ and $Y$ be measurable spaces as well as $\mu_X \in \mathit{Prob}[X]$ and $\mu_Y \in \mathit{Prob}[Y]$ be probability measures. Suppose $y \in Y$ with $\mu_Y = \mathit{Dirac}[y]$. Then there is precisely one coupling of $(\mu_X, \mu_Y)$, namely the product measure $\mu_X \otimes \mu_Y$. This can be seen as follows.

Let $\mu$ be a coupling of $(\mu_X, \mu_Y)$. For every measurable sets $B_X \subseteq X$ and $B_Y \subseteq Y$ such that $y \notin B_Y$ it holds $\mu(B_X \times B_Y) = 0$ since

$$\mu(B_X \times B_Y) \leq \mu(X \times B_Y) = \mu_Y(B_Y) = \mathit{Dirac}[y](B_Y) = 0.$$

Moreover, for every measurable sets $B_X \subseteq X$ and $B_Y \subseteq Y$ where $y \in B_Y$ we obtain $\mu(B_X \times B_Y) = \mu_X(B_X)$ as

$$\mu(B_X \times B_Y) = \mu(B_X \times X) - \mu(B_X \times (Y \setminus B_Y)) = \mu(B_X \times X) = \mu_X(B_X).$$

Putting things together, for every measurable sets $B_X \subseteq X$ and $B_Y \subseteq Y$ we derive the identity $\mu(B_X \times B_Y) = \mu_X(B_X) \cdot \mu_Y(B_Y)$ that justifies the claim $\mu = \mu_X \otimes \mu_Y$. ⌟

We finish this section with the following prominent application of couplings in the area of optimal transport that goes beyond the content of this thesis. The example below provides another view on couplings also useful for an illustration of the concept of weight functions in the next section:

**Example 32.** Let $X$ and $Y$ be measurable spaces $\mu_X \in \mathit{Prob}[X]$ and $\mu_Y \in \mathit{Prob}[Y]$ be probability measures, and $\mathit{cost}\colon X \times Y \to [0,1]$ be a measurable function. In the study of optimal transportation and allocation of resources, the set $X$ may refer to produced products, the set $Y$ stands for consumers, and the value $\mathit{cost}(x,y)$ can be interpreted as the cost of moving one unit of a product to the consumer. Given the distributions $\mu_X$ and $\mu_Y$ on $X$ and $Y$, respectively, the objective is to find a transference plan such that the transportation costs are minimised. Every transference plan is given by a coupling of $(\mu_X, \mu_Y)$ that leads to the following formulation of the Monge-Kantorovich minimisation problem considering an infimum over expected costs (see [Vil09] for an excellent overview on historical notes and related literature):

$$\inf \Big\{ \int \mathit{cost}(x,y)\, d\mu(x,y) \ ;\ \mu \text{ is a coupling of } (\mu_X, \mu_Y) \Big\}.$$

In the theory of optimal transport one is interested in conditions such that the infimum is attained and moreover, such that the minimising coupling, also called optimal transference plan, satisfies certain side constraints. One also investigates conditions under which an existing optimal transference plan is uniquely determined. ⌟

## 2.6 Weight lifting of relations

The weight lifting of a relation to be defined next extends a relation $R \subseteq X \times Y$ over measurable spaces $X$ and $Y$ to a relation $R^{\mathrm{wgt}} \subseteq \mathit{Prob}[X] \times \mathit{Prob}[Y]$ over the corresponding sets of probability measures $Prob[X]$ and $Prob[Y]$. Intuitively, two probability measures $\mu_X$ and $\mu_Y$ are related by the relation $R^{\mathrm{wgt}}$ provided there exists a coupling for $(\mu_X, \mu_Y)$ that is additionally compatible with the relation $R$. In this context such a coupling is called a weight function:

**Definition 33.** Let $X$ and $Y$ be a measurable spaces and $R \subseteq X \times Y$ be a relation. For every probability measures $\mu_X \in Prob[X]$ and $\mu_X \in Prob[Y]$ we call a probability measure $W \in Prob[X \times Y]$ a *weight function for* $(\mu_X, R, \mu_Y)$ if the following two statements hold:

(1) $W$ is a coupling of $(\mu_X, \mu_Y)$.

(2) There exists a measurable set $R' \subseteq X \times Y$ with $R' \subseteq R$ and $W(R') = 1$.

The *weight lifting* (*of* $R$) is the relation $R^{\mathrm{wgt}} \subseteq Prob[X] \times Prob[Y]$ defined as follows: for every $\mu_X \in Prob[X]$ and $\mu_Y \in Prob[Y]$,

$$\langle \mu_X, \mu_Y \rangle \in R^{\mathrm{wgt}} \quad \text{and} \quad \text{there is a weight function for } (\mu_X, R, \mu_Y).$$

┘

The relation $R^{\mathrm{wgt}}$ can be indeed seen as a lifting of $R$ since for every $x \in X$ and $y \in Y$ we have the following equivalence (see also Remark 31):

$$\langle x, y \rangle \in R \quad \text{iff} \quad \langle Dirac[x], Dirac[y] \rangle \in R^{\mathrm{wgt}}.$$

The latter statement also shows that for every relation $R^w \subseteq Prob[X] \times Prob[Y]$ such that $R^w$ is a weight lifting of some relation over $X$ and $Y$ there exists exactly one $R \subseteq X \times Y$ with $R^{\mathrm{wgt}} = R^w$ and moreover, it holds $R = \{\langle x, y \rangle \in X \times Y\,;\, \langle Dirac[x], Dirac[y] \rangle \in R^w\}$. Thus, the definition of the weight lifting $R^{\mathrm{wgt}}$ relies on no other information other than the given relation $R$.

In Definition 33 the relation $R$ is not required to be measurable in $X \times Y$. That is why condition (2) involves the existence of a measurable set in $X \times Y$ that is a subset of $R$. If the set $R$ is measurable in $X \times Y$, then condition (2) is obviously fulfilled precisely when one has $W(R) = 1$. According to Lemma 12, assuming that $X$ and $Y$ are Polish spaces and that the set $R$ is Souslin $X \times Y$, condition (2) can be replaced by the requirement

$$W^{\mathrm{out}}(R) = 1.$$

Provided the set $X$ is countable, the condition (2) is equivalent to the following statement: for every $x \in X$ and $y \in Y$,

$$W(\{\langle x, y \rangle\}) > 0 \quad \text{implies} \quad \langle x, y \rangle \in R,$$

i.e., the support of the probability measure $W$ is subsumed by the relation $R$.

Assuming the set $R$ is Borel in $X \times Y$, the existence of a weight function can be also characterised in terms of the Monge-Kantorovich minimisation problem briefly introduced in Example 32 (see also [DD09]). Define the Borel function $cost_R \colon X \times Y \to [0,1]$ as follows: for every $x \in X$ and $y \in Y$ let

$$cost_R(x,y) = 0 \text{ if } \langle x,y\rangle \in R \quad \text{and} \quad cost_R(x,y) = 1 \text{ if } \langle x,y\rangle \notin R.$$

Intuitively, the defined cost function states that a pair in $X \times Y$ causes no costs provided it is contained in the relation $R$ and cost one otherwise. For every probability measures $\mu_X \in Prob[X]$ and $\mu_Y \in Prob[Y]$ it holds $\langle \mu_X, \mu_Y\rangle \in R^{\mathrm{wgt}}$ precisely when there exists a coupling $W$ of $(\mu_X, \mu_Y)$ such that

$$\int cost_R(x,y)\, dW(x,y) = 0,$$

in particular, this also means that the infimum of the Monge-Kantorovich minimisation problem is attained.

The remainder of this section regards the case where $X = Y$. However, notions and observations can be adapted for the case where $X$ and $Y$ are not the same. Throughout this thesis, we are typically confronted with the following question: given a relation $R$ over some measurable space $X$ and two probability measures $\mu_a$ and $\mu_b$ on $X$, does there exist a weight function for $(\mu_a, R, \mu_b)$? In what follows we provide simple necessary conditions for the existence of a corresponding weight function. For this let us recall the following standard notions first. Let $X$ be a set and $R \subseteq X \times X$ be a relation. A subset $B \subseteq X$ is called *upper $R$-stable* provided

$$R \cap (B \times X) \subseteq X \times B.$$

Similar, a subset $B \subseteq X$ is *called $R$-stable* if

$$R \cap (B \times X) = R \cap (X \times B),$$

i.e., the set $B$ is both upper $R$-stable and upper $R^{-1}$-stable. If $R$ is a preorder, an upper $R$-stable set is also called an upward closed set or simply an upset. Assuming the relation $R$ constitutes an equivalence on $X$, a set $B \subseteq X$ is $R$-stable precisely when $B$ can be written as an union of equivalence classes concerning $R$. The latter is equivalent to the requirement $B = \bigcup_{x \in B} [x]_R$ where for every $x \in B$ the equivalence class concerning $R$ that contains $x$ is denoted by $[x]_R$.

**Remark 34.** Let $X$ be a measurable space, $\mu_a, \mu_b \in \mathit{Prob}[X]$, and $R \subseteq X \times X$ be a relation such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Then the following two statements hold:

(1) $\mu_a(B) \leq \mu_b(B)$ for every upper $R$-stable Borel set $B \subseteq X$.

(2) $\mu_a(B) = \mu_b(B)$ for every $R$-stable Borel set $B \subseteq X$.

The argument is straightforward. Let $W$ be a weight function for $(\mu_a, R, \mu_b)$. Moreover, suppose a Borel set $R' \subseteq X \times X$ with $R' \subseteq R$ and $W(R') = 1$. For every Borel set $B \subseteq X$ with $R \cap (B \times X) \subseteq X \times B$ it also holds $R' \cap (B \times X) \subseteq X \times B$ and therefore,

$$\mu_a(B) = W(B \times X) = W(R' \cap (B \times X)) \leq W(X \times B) = \mu_b(B).$$

This justifies statement (1). Statement (2) follows analogously since for every set $B \subseteq X$ with $R \cap (B \times X) = R \cap (X \times B)$ one has $R' \cap (B \times X) = R' \cap (X \times B)$. ⌟

**Example 35.** Let $X$ be a Polish space. Consider the diagonal relation $\mathit{Diag}$ on $X$, i.e.,

$$\mathit{Diag} = \{\langle x_a, x_b \rangle \in X \times X \,;\, x_a = x_b\}.$$

By Theorem 6.5.7 in [Bog07] and Remark 4, the set $\mathit{Diag}$ is Borel in $X \times X$. It holds

$$\mathit{Diag}^{\mathrm{wgt}} = \{\langle \mu_a, \mu_b \rangle \in \mathit{Prob}[X] \times \mathit{Prob}[X] \,;\, \mu_a = \mu_b\},$$

i.e., the weight lifting $\mathit{Diag}^{\mathrm{wgt}}$ and the diagonal relation on $\mathit{Prob}[X]$ are the same. Let us provide arguments for that claim. Consider two probability measures $\mu_a, \mu_b \in \mathit{Prob}[X]$. Define the Borel function $\xi\colon X \to X \times X$, $\xi(x) = \langle x, x \rangle$. Provided $\mu_a = \mu_b$, it is easy to see that $\xi_\sharp(\mu_a)$ yields a weight function for $(\mu_a, \mathit{Diag}, \mu_b)$. Therefore, the diagonal relation on $\mathit{Prob}[X]$ is a subset of the weight lifting $\mathit{Diag}^{\mathrm{wgt}}$. Every Borel set $B \subseteq X_a$ is $\mathit{Diag}$-stable and therefore, assuming $\langle \mu_a, \mu_b \rangle \in \mathit{Diag}^{\mathrm{wgt}}$, we obtain $\mu_a(B) = \mu_b(B)$ by Remark 34. Hence, $\mathit{Diag}^{\mathrm{wgt}}$ is a subset of the diagonal relation on $\mathit{Prob}[X]$. ⌟

## 2.7 Smooth and weakly smooth relations

Inspecting Remark 34 again, one may ask whether the proven necessary conditions give rise to a characterisation of the weight lifting of a relation. More precisely, the following question arises naturally: considering probability measures $\mu_a$ and $\mu_b$ in $\mathit{Prob}[X]$ such that $\mu_a(B) = \mu_b(B)$ for every $R$-stable Borel set $B \subseteq X$, does there exist a weight function for $(\mu_a, R, \mu_b)$? This section presents a condition on $R$ for which the latter question can be answered positively. More precisely, we focus on smooth relations [Dob07, Sri08, Gao08] (also known as countably-separated relations) as well as the closely related concept of a weakly smooth relations. Intuitively, smooth as well as weakly smooth relations can be represented by means of a countable number of sets that can be viewed as test sets to determine whether two elements are in the relation or not:

**Definition 36.** Let $X$ be a Polish space. A relation $R \subseteq X \times X$ is called *weakly smooth* provided there exists a countable family $\mathcal{C}$ of Borel subsets of $X$ such that

$$R = \{\langle x_a, x_b \rangle \in X \times X \,;\, x_a \in B \text{ implies } x_b \in B \text{ for every } B \in \mathcal{C}\}.$$

In this context, we refer to $\mathcal{C}$ as a *witness of the weakly smoothness of* $R$. Similar, a relation $R \subseteq X \times X$ is called *smooth* if there exists a countable family $\mathcal{C}$ of Borel subsets of $X$ with

$$R = \{\langle x_a, x_b \rangle \in X \times X \,;\, x_a \in B \text{ iff } x_b \in B \text{ for every } B \in \mathcal{C}\}.$$

Here, $\mathcal{C}$ is called a *witness of the smoothness of* $R$. ⌟

Every weakly smooth relation is necessarily a preorder and accordingly, every smooth relation is an equivalence. As the notions suggests, every smooth relation is also weakly smooth. Indeed, considering a smooth relation $R \subseteq X \times X$ over some Polish space $X$ as well as a family of sets $\mathcal{C}$ serving as a witness, it is easy to see that the family of sets $\mathcal{C} \cup \{X \setminus B \,;\, B \in \mathcal{C}\}$ constitutes a witness of the weakly smoothness of $R$. The definition of a smooth relation is borrowed from Section 5.1 in [Sri08] (see also [Dob07, Gao08]). Besides this, Definition 36 is also inspired by the general technique for relating (bi)simulation and logic for labelled Markov processes [Des99, DEP02, DGJP03, DP03, FKP17] (see also Chapter 5 of this thesis). The usefulness of the introduced notions heavily originates from the fact that the family $\mathcal{C}$ of Borel subsets is required to be countable. Roughly speaking, although the cardinality of a (weakly) smooth relation may be uncountable, the relation under consideration is uniquely determined by a countable number of test sets.

**Remark 37.** Let $X$ be a Polish space and $R \subseteq X \times X$ be a relation. Then $R$ is smooth precisely when there are a Polish space $Y$ and a Borel function $f \colon X \to Y$ such that

$$R = \{\langle x_a, x_b \rangle \in X \times X \,;\, f(x_a) = f(x_b)\}.$$

Although the claimed equivalence is folklore (see, e.g., Exercise 5.1.10 in [Sri08]), let us present the standard argument. Assume that the relation $R$ is smooth and let $\mathcal{C}$ be a family of subsets of $X$ that serves as a witness. Let $B_0, B_1, \ldots$ be Borel sets such that $\mathcal{C} = \{B_0, B_1, \ldots\}$.

Consider the Polish space $Y = \{0,1\}^\omega$ (see Example 2 (5)). Define the Borel function $f\colon X \to Y$ as follows: for every $x \in X$ and $n \in \mathbb{N}$ let

$$f(x)[n] = 1 \text{ if } x \in B_n \quad \text{and} \quad f(x)[n] = 0 \text{ if } x \notin B_n.$$

It is easy to see that for every $x_a, x_b \in X$ it holds $\langle x_a, x_b\rangle \in R$ iff $f(x_a) = f(x_b)$.

Let $Y$ be a Polish space and $f\colon X \to Y$ be a Borel function such that the following identity holds: $R = \{\langle x_a, x_b\rangle \in X \times X\,;\, f(x_a) = f(x_b)\}$. Consider a countable generator $\mathcal{G}_Y$ of the Borel sigma algebra on $Y$ that separates the points of $Y$ (see Remark 4). Define the countable family $\mathcal{C}$ of Borel subsets of $X$ by

$$\mathcal{C} = \{f^{-1}(B_Y)\,;\, B_Y \in \mathcal{G}_Y\}.$$

It is easy to see that for every $x_a, x_b \in X$ it holds $\langle x_a, x_b\rangle \in R$ precisely when for every $B_Y \in \mathcal{G}_Y$ one has $x_a \in f^{-1}(B_Y)$ iff $x_b \in f^{-1}(B_Y)$. Hence, the relation $R$ is smooth. ⌟

By the previous remark, a relation $R$ over a set $X$ is smooth provided there is a Borel assignment $f$ of every element in $X$ to an element of some Polish space $Y$ such that two elements of $X$ are related by $R$ precisely when they are assigned to the same element in $Y$. Viewing the diagonal relation on some Polish space as a rather simple relation, we obtain another indication for the usefulness of smooth relations. In fact, many properties of a diagonal relation on a Polish space can be transferred to smooth relations.

**Remark 38.** Let $X$ be a Polish space and $R \subseteq X \times X$ be a relation. Provided $R$ is smooth or weakly smooth, the set $R$ is Borel in $X \times X$ (and hence also Souslin in $X \times X$ by Remark 10 (2)). It suffices to prove this claim for weakly smooth relations. Assume that the relation $R$ is weakly smooth in $X$ and let $\mathcal{C}$ be a witness of the weakly smoothness of $R$. Since the family $\mathcal{C}$ is countable and as we have

$$R = \bigcap_{B \in \mathcal{C}} \big((X \times B) \cup ((X \setminus B) \times X)\big),$$

it directly follows that the set $R$ is Borel in $X \times X$. ⌟

The remainder of this section provides characterisations of the weight liftings of smooth as well as of weakly smooth relations corresponding to the introductory sketched motivation of this section.

**Theorem 39.** *Let $X$ be a Polish space, $\mu_a, \mu_b \in Prob[X]$ be probability measures, and $R \subseteq X \times X$ be a weakly smooth relation. Let $\mathcal{C}$ be a witness of the weakly smoothness of $R$ such that $\mathcal{C}$ is closed under finite intersections and finite unions. The three statements below are equivalent:*

(1) $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$.

(2) $\mu_a(B) \leq \mu_b(B)$ *for every upper $R$-stable Borel set $B \subseteq X$.*

(3) $\mu_a(B) \leq \mu_b(B)$ *for every $B \in \mathcal{C}$.*

⌟

A precise proof can be found below. Remark 34 (1) shows that statement (2) of the previous theorem is necessarily fulfilled for every $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. The reverse direction, i.e., the implication from (2) to (1), is an application of Strassen's theorem on stochastic domination [Str65, KKO77, Kel84, Les10]. The difficulty we had is that Strassen's theorem on stochastic domination focuses on relations that yield closed sets in the product space. However, as in Polish spaces Borel functions can be turned into continuous ones (see Chapter 13 in [Kec95]), one can adapt the topology on the Polish space $X$ in an appropriate way such that the relation $R$ under consideration is closed in the product topology induced by this new topology. One can then apply Strassen's theorem on stochastic domination to finish the argument showing that (1) is implied by (2).

The crux of statement (3) in the previous theorem is that the family $\mathcal{C}$ is countable, in particular, it suffices to consider a countable number of test sets, namely those contained in $\mathcal{C}$, to determine whether there exists a weight function for $(\mu_a, R, \mu_b)$. To show the equivalence of (2) and (3), we extract a clever argument from the recent paper [FKP17] using the positive monotone class theorem and the positive unique structure theorem (see Section 5.1 in [FKP17]).

**Theorem 40.** *Let $X$ be a Polish space, $\mu_a, \mu_b \in Prob[X]$ be probability measures, and $R \subseteq X \times X$ be a smooth relation. Moreover, let $\mathcal{C}$ be a witness of the smoothness of $R$ such that $\mathcal{C}$ is closed under finite intersections. Then the following three statements are equivalent:*

(1) $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$.

(2) $\mu_a(B) = \mu_b(B)$ *for every $R$-stable Borel sets $B \subseteq X$.*

(3) $\mu_a(B) = \mu_b(B)$ *for every $B \in \mathcal{C}$.*

⌟

The previous theorem for smooth relations is exactly in the same shape as Theorem 39 with the important difference that the family is only required to be closed under finite intersection. The equivalence of (1) and (2) follows directly from Theorem 39. To prove the equivalence of (2) and (3), we again adapt an argument from the paper [FKP17] relying on Dynkin's $\pi$-$\lambda$ theorem and the unique structure theorem (see Section 4.1 in [FKP17]). The following corollary can be also found in [Lov12b] (see also [Lov12a]):

**Corollary 41.** *Let $X$ be a Polish space and $R \subseteq X \times X$ be a relation. Suppose a Polish space $Y$ and a Borel function $f\colon X \to Y$ with $R = \{\langle x_a, x_b\rangle \in X \times X\,;\, f(x_a) = f(x_b)\}$. For every probability measures $\mu_a, \mu_b \in Prob[X]$ the following equivalence holds:*

$$\langle \mu_a, \mu_b\rangle \in R^{\mathrm{wgt}} \quad \textit{iff} \quad f_\sharp(\mu_a) = f_\sharp(\mu_b).$$

*In particular, the relation $R^{\mathrm{wgt}}$ is smooth.*

*Proof.* For every Borel set $B_Y \subseteq Y$ it is easy to see that the set $f^{-1}(B_Y)$ is $R$-stable. Thus, according to Theorem 40, the identity $f_\sharp(\mu_a) = f_\sharp(\mu_b)$ follows from $\langle \mu_a, \mu_b\rangle \in R^{\mathrm{wgt}}$.

We regard the reverse implication. Assume $f_\sharp(\mu_a) = f_\sharp(\mu_b)$. Let $\mathcal{C}_Y$ be a countable generator of the Borel sigma algebra on $Y$ that is closed under finite intersections and separates the points of $Y$ (see Remark 4). Define

$$\mathcal{C}_X = \{f^{-1}(B_Y)\,;\, B_Y \in \mathcal{C}_Y\}.$$

Then the family $\mathcal{C}_X$ is countable, a subset of the Borel sigma algebra on $X$, and closed under finite intersections. As $\mathcal{C}_Y$ separates the points in $Y$, it is easy to see that $\mathcal{C}_X$ is a witness of the smoothness of $R$. It moreover holds $\mu_a(B_X) = \mu_b(B_X)$ for every $B_X \in \mathcal{C}_X$. Putting things together, we can apply Theorem 40 and hence, it holds $\langle \mu_a, \mu_b\rangle \in R^{\mathrm{wgt}}$.

Since $Prob[X]$ yields a Polish space (see Example 2 (6)) and as the pushforward function is Borel (see Remark 8 (1)), it moreover follows that the relation $R^{\mathrm{wgt}}$ is smooth. □

**Proof of section's main results.** It remains to prove Theorems 39 and 40.

**Lemma 42.** *Let $X$ be a Polish space and $R$ be a weakly smooth relation. Then there are a Polish space $Y$, a partial order $\trianglelefteq$ on $Y$ being a closed set in $Y \times Y$, and a Borel function $f\colon X \to Y$ satisfying the following identity*

$$R = \{\langle x_a, x_b\rangle \in X \times X\,;\, f(x_a) \trianglelefteq f(x_b)\}.$$

*Proof.* The argument is basically the same as in Remark 37. Consider a witness $\mathcal{C}$ of the weakly smoothness of $R$. Let $B_0, B_1, \ldots$ be Borel subsets of $X$ such that $\mathcal{C} = \{B_0, B_1, \ldots\}$. Define the Polish space $Y = \{0,1\}^\omega$ (see Example 2 (5)). The partial order $\trianglelefteq$ on $Y$ is defined point-wise: for every $y_a, y_b \in Y$ let

$$y_a \trianglelefteq y_b \quad \text{and} \quad \text{for every } n \in \mathbb{N} \text{ it holds } y_a[n] \le y_b[n].$$

It is easy to see that the set $\trianglelefteq$ is closed in $Y \times Y$. Define the Borel function $f\colon X \to Y$ as follows: for every $x \in X$ and $n \in \mathbb{N}$ let $f(x)[n] = 1$ if $x \in B_n$ and $f(x)[n] = 0$ if $x \notin B_n$. Then for every $x_a, x_b \in X$ one has $\langle x_a, x_b \rangle \in R$ iff $f(x_a) \trianglelefteq f(x_b)$. □

*Proof of Theorem 39.* (1) *iff* (2). Remark 34 yields the implication from (1) to (2). Our argument for the reverse implication, i.e., (2) implies (1), exploits a result from probability theory concerning the existence of probability measures on product spaces with given marginals: the following argument relies on Theorem 11 in [Str65] (see also Theorem 1 in [KKO77], Proposition 3.12 in [Kel84], and Theorem 2.5 in [Les10]).

By Lemma 42, there are a Polish space $Y$, a partial order $\trianglelefteq$ being a closed set in $Y \times Y$, and a Borel function $f\colon X \to Y$ such that $R = \{\langle x_a, x_b \rangle \in X \times X \,;\, f(x_a) \trianglelefteq f(x_b)\}$. Note, as the relation $\trianglelefteq$ is reflexive, the set $R$ is not empty.

We introduce a suitable topology $\mathcal{O}$ on $X$ such that the set $R$ is closed in the product topology $\mathcal{O} \otimes \mathcal{O}$. Denote the Polish topology on $X$ by $\mathcal{O}'$. By Theorem 13.11 in [Kec95], there is a Polish topology $\mathcal{O}$ on $X$ such that $\mathit{Borel}[\mathcal{O}'] = \mathit{Borel}[\mathcal{O}]$ and $\mathcal{O}' \subseteq \mathcal{O}$ as well as so that function $f$ is continuous with respect to $\mathcal{O}$. Defining the function $g\colon X \times X \to Y \times Y$, $g(x_a, x_b) = \langle f(x_a), f(x_b) \rangle$, it follows $R = g^{-1}(\trianglelefteq)$. From this we derive that the set $R$ is closed in $\mathcal{O} \otimes \mathcal{O}$ as, using the inclusion $\mathcal{O}' \subseteq \mathcal{O}$, the function $g$ is continuous with respect to the product topology $\mathcal{O} \otimes \mathcal{O}$.

Putting things together, we are in the situation of Theorem 11 in [Str65]. Indeed, using this theorem and as $\mathit{Borel}[\mathcal{O}'] = \mathit{Borel}[\mathcal{O}]$, there exists a weight function for $(\mu_a, R, \mu_b)$ if for every upper $R$-stable Borel set $B \subseteq X$ one has $\mu_a(B) \leq \mu_b(B)$. Statement (1) hence follows from (2).

(2) *iff* (3). As $\mathcal{C}$ is a witness of the weakly smoothness of $R$, it is easy to see that every $B \in \mathcal{C}$ is upper $R$-stable. Therefore, statement (3) immediately follows from (2). The remaining implication from (3) to (2) uses the argument for Theorem 17 in [FKP17]. More precisely, we rely on Theorems 15 and 16 in [FKP17] called positive monotone class theorem and positive unique structure theorem, respectively. We rely on the following additional notion. By $\mathit{SigmaLattice}[\mathcal{C}]$ we denote the smallest family of subsets of $X$ that contains every set of the family $\mathcal{C}$ and moreover, that is closed under both countable unions and countable intersections.

The implication from (3) to (2) is shown by contraposition. To this end consider an upper $R$-stable Borel set $B' \subseteq X$ such that $\mu_a(B') > \mu_b(B')$. As the family $\mathcal{C}$ is countable and a witness of the weakly smoothness of $R$, we can apply Theorem 16 in [FKP17] that yields $B' \in \mathit{SigmaLattice}[\mathcal{C}]$. Introduce the family $\mathcal{M}$ of Borel subsets of $X$ by

$$\mathcal{M} = \{B \subseteq X \text{ Borel} \; ; \mu_a(B) \leq \mu_b(B)\}.$$

Since every measure is continuous from above and below, it follows that $\mathcal{M}$ is a monotone class on $X$, i.e., the family $\mathcal{M}$ is closed under unions of increasing chains and under intersections of decreasing chains. As $\mathcal{C}$ is closed under finite intersections and finite unions, we are in the situation of Theorem 15 in [FKP17] providing the following implication:

$$\mathcal{C} \subseteq \mathcal{M} \quad \text{implies} \quad \mathit{SigmaLattice}[\mathcal{C}] \subseteq \mathcal{M}.$$

Since $B' \in \mathit{SigmaLattice}[\mathcal{C}]$ and $B' \notin M$, we hence obtain $\mathcal{C} \nsubseteq \mathcal{M}$. As a consequence, there exists $B'' \in \mathcal{C}$ such that $\mu_a(B'') > \mu_b(B'')$. Putting things together, this shows to contraposition of the implication from (3) to (2). □

*Proof of Theorem 40.* (1) *iff* (2). Remark 34 justifies the implication from (1) to (2). As $R$ is smooth, the relation $R$ is in particular weakly smooth. Moreover, since $R$ is a symmetric relation, every upper $R$-stable Borel set in $X$ is $R$-stable. The implication from (2) to (1) hence follows immediately from Theorem 39.

(2) *iff* (3). Since $\mathcal{C}$ is a witness of the smoothness of $R$, every $B \in \mathcal{C}$ satisfies the identity $R \cap (B \times X) = R \cap (X \times B)$ and therefore, statement (3) immediately follows from (2). As in the proof of Theorem 39, the remaining implication from (3) to (2) exploits the argument for Theorem 13 in [FKP17]. To be more precise, we use Theorems 11 and 12 in [FKP17] called Dynkin's $\pi$-$\lambda$ theorem and unique structure theorem, respectively. Here, the former theorem is a prominent theorem in basic-measure theory (see also, e.g., Theorem 136B in [Fre01]). In what follows, $\mathit{SigmaAlgebra}[\mathcal{C}]$ denotes the smallest sigma algebra on $X$ that contains every set in $\mathcal{C}$.

The implication from (3) to (2) is proven by contraposition. Consider a $R$-stable Borel set $B' \subseteq X$ such that $\mu_a(B') = \mu_b(B')$. As the family $\mathcal{C}$ is countable and a witness of the smoothness of $R$, Theorem 12 in [FKP17] justifies $B' \in \mathit{SigmaAlgebra}[\mathcal{C}]$. Introduce the family $\mathcal{D}$ of Borel subsets of $X$ by

$$\mathcal{D} = \{B \subseteq X \text{ Borel} \; ; \mu_a(B) = \mu_b(B)\}.$$

It follows that $\mathcal{D}$ is a Dynkin system, i.e., $\mathcal{D}$ is not empty, closed under complementation, and closed under unions of a countable number of pairwise disjoint sets. As $\mathcal{C}$ is closed under finite intersections, we are in the situation of Theorem 11 in [FKP17] stating the following implication:

$$\mathcal{C} \subseteq \mathcal{D} \quad \text{implies} \quad \mathit{SigmaAlgebra}[\mathcal{C}] \subseteq \mathcal{D}.$$

As $B' \in \mathit{SigmaAlgebra}[\mathcal{C}]$ and $B' \notin \mathcal{D}$, we hence obtain $\mathcal{C} \nsubseteq \mathcal{D}$. There hence exists $B'' \in \mathcal{C}$ with $\mu_a(B'') \neq \mu_b(B'')$. This finally justifies the contraposition of the implication from (3) to (2). $\square$

It would be interesting whether one could drop or weaken the requirements on the relation $R$ in Theorems 40 (and also Theorem 39) for the relation under consideration. In this context we remark that personal communications with the authors of [BBLM14] yield that (the proof of) Proposition 13 in [BBLM14] is flawed. More precisely, it turns out that a function constructed in the proof in [BBLM14] is not well-defined. Moreover, adapting a counterexample provided in the article [Swa96], Proposition 13 in [BBLM14] is actually false for arbitrary measurable spaces. However, the claim might be still true for the interesting Polish-space case. In fact, the mentioned proposition for the Polish-space case would yield a strong improvement of Theorem 40.

# 3 Stochastic transition systems

The main model considered in this thesis is called stochastic transition system (STS) and can be seen as a conservative generalisation of the generative stochastic model in the classification of [GSS95] for discrete probabilistic systems and of probabilistic automata [Seg95]. However, while [GSS95, Seg95] focus on systems with countable state spaces, we investigate stochastic systems with uncountable state and action spaces. The definition of STSs is closely related to the model in [CSKN05, Cat05]. Indeed, the formalism studied in the latter cited articles corresponds to the notion of simple STSs in this thesis. Further prominent models with uncountable state spaces are given by labelled Markov processes [BDEP97, Des99, DEP02, Pan09] and non-deterministic labelled Markov processes [DTW12, Wol12]. These models as well as stochastic automata [D'A99, DK05] and other classes of stochastic hybrid systems are covered by STSs. For more details on this subject, however, we refer to Chapter 7 since a profound discussion should include those subclasses of STSs that are studied in later chapters of this thesis.

One challenge we had to address for the definition of simulation and bisimulation is to conservatively extend the notion of generative bisimulation in [GSS95] (see also [Tin07]) to our setting addressing STSs with uncountable state and action spaces. Relying on the action-lifting of a given relation, we indeed obtain suitable notions for simulation and bisimulation. In Chapter 7 we argue that these notions and corresponding standard definitions for, e.g., (non-deterministic) labelled Markov processes, are the same. Moreover, based on a two-step view on sampling an action-state pair, we obtain an intuitive characterisation of the introduced concepts.

As a main result of this chapter we provide a proof of Theorem A presented in Chapter 1 stating that the simulation preorder and the bisimulation equivalence form transitive relations. Our argument proceeds as follows: it is shown that the composition of two simulations also yields a simulation and accordingly, that the composition of two bisimulations constitutes again a bisimulation. The proof of the latter statement requires the construction of a weight function that intuitively represents a merging of two given weight functions. For this we exploit the gluing lemma (see, e.g., Lemma 5.3.2 in [AGS05] or Section 1 in [Vil09]), which states that two given probability measures on product spaces can be glued together along a common marginal. The assumption that the state space and the action space are Polish turns out to be crucial for our argument, in particular, as the gluing lemma does not hold for arbitrary measurable spaces [Swa96].

The main result of the early contribution [BDEP97] on labelled Markov processes is the result that bisimulation indeed yields an equivalence equivalence relation where the definition of bisimulation relies on categorical ideas and zig-zag morphisms. In a later article [DEP02] the authors provide a complete logical characterisation for the latter mentioned notion of bisimulation that directly implies transitivity of the induced bisimulation relation. In [DGJP03] the authors recapitulate this early approach. The recent contribution [FKP17] shows that simulation preorder on labelled Markov processes is also transitive that extends a corresponding result in [Des99] for a subclass of labelled Markov processes. The chapter's main contribution continuous this research for a much broader class of models and moreover, as we describe in Chapter 7, covers the previous mentioned results. Whereas in [Des99, DEP02, DGJP03, FKP17] the transitivity follows directly from the logical characterisation, our proof does not involve any logical characterisation of the simulation preorder or the bisimulation equivalence.

To reasoning about the linear-time behaviour of STSs, trace distributions as well as the associated concepts of schedulers and path measures are introduced. Intuitively, in every state a scheduler chooses a probability measure that can be represented as a convex combination of enabled distributions in that given state. The latter is precisely formalised by the notion of combined-transition relation. Relying on the theory developed in Section 2.4, we show that the combined-transition relation is closed under combining probability measures under a side constraints that fits to the thesis setting as we see in the subsequent chapters. This contribution provides a generalisation of Proposition 4.2.1 in [Seg95] where a corresponding statement is shown for discrete probabilistic automata.

## 3.1 Modelling stochastic transition systems

The state and the action space of the basic stochastic model investigated throughout this thesis are Polish spaces. We refer to Section 2.1 for a brief overview on Polish spaces and their rich mathematical theory. However, for the moment, it suffices to recall Example 2 where important Polish spaces appearing in modelling stochastic systems are presented such as countable sets as well as countable Cartesian products of the real number line.

**Definition 43.** A *stochastic transition system* (*STS*) is a triple

$$\mathcal{T} = (Sta, Act, \rightarrow)$$

consisting of the following elements:

(1) *Sta* is a Polish space (*state space*),

(2) *Act* is a Polish space (*action space*), and

(3) $\rightarrow \subseteq Sta \times Prob[Act \times Sta]$ is a relation (*transition relation*).

⌟

Before we present examples for STSs, let us introduce the following notions associated to an STS $\mathcal{T} = (Sta, Act, \rightarrow)$. As usual, for every $s \in Sta$ and $\varphi \in Prob[Act \times Sta]$ we write $s \rightarrow \varphi$ rather than $\langle s, \varphi \rangle \in \rightarrow$. Accordingly, for every $s \in Sta$, $act \in Act$, and $\mu \in Prob[Sta]$ we use $s \rightarrow \langle act, \mu \rangle$ as an alternative notation for $s \rightarrow Dirac[act] \otimes \mu$. For every state $s \in Sta$ we moreover define the set

$$Enabled[s] = \{\varphi \in Prob[Act \times Sta] \ ; s \rightarrow \varphi\}.$$

Similar, for every state $s \in Sta$ and action $act \in Act$ let

$$Enabled[s, act] = \{\mu \in Prob[Sta] \ ; s \rightarrow \langle act, \mu \rangle\}.$$

For every $s \in Sta$ and $act \in Act$, to indicate the STS under consideration, we also write $Enabled[\mathcal{T}, s]$ and $Enabled[\mathcal{T}, s, act]$ instead of $Enabled[s]$ and $Enabled[s, act]$, respectively. The STS $\mathcal{T}$ is called *non-blocking* if for every $s \in Sta$ the set $Enabled[s]$ is not empty. Similar, we call the STS $\mathcal{T}$ *point-wise non-blocking* if for every $s \in Sta$ and $act \in Act$ the set $Enabled[s, act]$ is not empty. The STS $\mathcal{T}$ is said to be *image-finite* if for every state $s \in Sta$ the set $Enabled[s]$ is finite. Moreover, $\mathcal{T}$ is called *point-wise image-finite* provided for every state $s \in Sta$ and action $act \in Act$ the set $Enabled[s, act]$ is finite.

We refer to the STS $\mathcal{T}$ as *simple* provided for every transition $s \rightarrow \varphi$ there exists an action $act \in Act$ such that $\varphi(\{act\} \times Sta) = 1$. In other words, every transition in a simple STS involves only one single action: for every every transition $s \rightarrow \varphi$ and every action $act \in Act$ one either observes the action $act$ with probability one, i.e., $\varphi(\{act\} \times Sta) = 1$, or with probability zero, i.e., $\varphi(\{act\} \times Sta) = 0$. Note, the model of simple STSs conservatively extends the concept of simple probabilistic automata [Seg95].

**Real-valued variables.** We use variables in order to represent states of an STSs with a descriptive name. Besides this, in compositional modelling, variables serve as a communication mechanism since different STSs may reference to the same shared variable. Let *Var* be a finite and non-empty set of real-valued variables. Denote the set of all evaluations over *Var* by $\mathit{Eval}[\mathit{Var}]$, i.e., the set $\mathit{Eval}[\mathit{Var}]$ consists of all the functions $e$ with domain *Var* and codomain $\mathbb{R}$. Intuitively, if $e \in \mathit{Eval}[\mathit{Var}]$ and $v \in \mathit{Var}$, then $e(v)$ represents the value of the variable $v$. Obviously, $\mathit{Eval}[\mathit{Var}]$ can be identified with a finite Cartesian product of the real number line and hence, $\mathit{Eval}[\mathit{Var}]$ is a Polish space according to Example 2 (5).

Considering a finite non-empty subset $V \subseteq \mathit{Var}$, for every $e \in \mathit{Eval}[\mathit{Var}]$ the variable evaluation $e_{|V} \in \mathit{Eval}[V]$ is defined by $e_{|V}(v) = e(v)$ for all $v \in V$, i.e., $e_{|V}$ denotes the restriction of $e$ onto $V$. Similar, for every probability measure $\eta \in \mathit{Prob}[\mathit{Eval}[\mathit{Var}]]$ the restriction of $\eta$ onto $V$ is given by the probability measure $\eta_{|V} \in \mathit{Prob}[\mathit{Eval}[V]]$ defined as follows: for every Borel set $E \subseteq \mathit{Eval}[V]$ let $\eta_{|V}(E) = \eta(\{e \in \mathit{Eval}[\mathit{Var}]\ ;\ e_{|V} \in E\})$. Thus, $\eta_{|V}$ is the pushforwad measure of $\eta$ with respect to the Borel function assigning $e \in \mathit{Eval}[\mathit{Var}]$ to $e_{|V} \in \mathit{Eval}[V]$ (see also Section 2.1).

Let $n \in \mathbb{N} \setminus \{0\}$ and $v_1, \ldots, v_n$ be variables such that $\mathit{Var} = \{v_1, \ldots, v_n\}$. The set of all conditions over *Var* is denoted by $\mathit{Cond}[\mathit{Var}]$. More precisely, $\mathit{Cond}[\mathit{Var}]$ is the smallest family of subsets of $\mathit{Eval}[\mathit{Var}]$ that is closed under all boolean connectives and that consists of every subset *cond* of $\mathit{Eval}[\mathit{Var}]$ with the following property: there are rational numbers $q_1, \ldots, q_n, q \in \mathbb{Q}$ and a comparison operator $\bowtie \in \{<, \leq, =, \geq, >\}$ with

$$\mathit{cond} = \{e \in \mathit{Eval}[\mathit{Var}]\ ;\ q_1 \cdot e(v_1) + \ldots + q_n \cdot e(v_n) \bowtie q\}.$$

Here, the set *cond* is usually represented in a symbolic way by $q_1 \cdot v_1 + \ldots + q_n \cdot v_n \bowtie q$. Clearly, the family $\mathit{Cond}[\mathit{Var}]$ is countable and consists only of Borel subsets of $\mathit{Eval}[\mathit{Var}]$. For every conditions $\mathit{cond}_1$ and $\mathit{cond}_2$ we write $\mathit{cond}_1 \wedge \mathit{cond}_2$ rather than $\mathit{cond}_1 \cap \mathit{cond}_2$ and accordingly, for the other boolean connectives (negation $\neg$, disjunction $\vee$, implication $\rightarrow, \ldots$). We write $e \models \mathit{cond}$ if the variable evaluation $e \in \mathit{Eval}[\mathit{Var}]$ satisfies the condition $\mathit{cond} \in \mathit{Cond}[\mathit{Var}]$, i.e., when it holds $e \in \mathit{cond}$. For instance, if $n \geq 2$, then for every variable evaluation $e \in \mathit{Eval}[\mathit{Var}]$,

$$e \models (v_1 \geq 2.718) \wedge (v_2 \leq 3.141) \quad \text{iff} \quad e(v_1) \geq 2.718 \text{ and } e(v_2) \leq 3.141.$$

We finally remark that the set $\mathit{Cond}[\mathit{Var}]$ separates the points of $\mathit{Eval}[\mathit{Var}]$, i.e., for every $e_1, e_2 \in \mathit{Eval}[\mathit{Var}]$ where $e_1 \neq e_2$ there exists a condition $\mathit{cond} \in \mathit{Cond}[\mathit{Var}]$ such that $e_1 \models \mathit{cond}$ while $e_2 \not\models \mathit{cond}$. Indeed, if $e_1 \neq e_2$, then there exists $v \in \mathit{Var}$ with $e_1(v) \neq e_2(v)$ and thus, there is $\bowtie \in \{<, >\}$ and $q \in \mathbb{Q}$ such that $e_1(v) \bowtie q \bowtie e_2(v)$. Summarising facts, the family $\mathit{Cond}[\mathit{Var}]$ forms a generator of the Borel sigma algebra on $\mathit{Eval}[\mathit{Var}]$ that satisfies the conditions in Remark 4 (see also Example 3 (5)).

**Example: simple cooling system for a server.** As examples for describing stochastic systems with STSs, we provide different refinement of an STS modelling a simple cooling system for a server. The following presentation is inspired by [HH15] providing a classification of modelling formalisms concerning their expressive power.

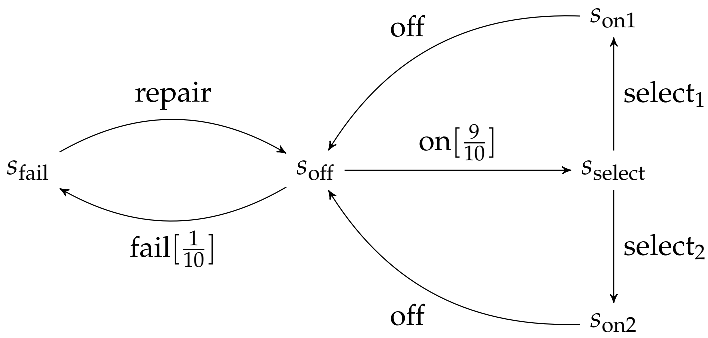


Figure 3.1: Simple cooling system for a server.

**Example 44.** To maintain the functionality of hardware components of a server, the cooling system dissipates the heat produced. In order to save energy costs, a cooling system can be switched on in phases where the server is busy and off when the server is idling. With a small probability, a switching on the system may fail. Moreover, to be adaptive for different uses of the server, the cooling system implements two cooling strategies. The STS $(Sta, Act, \to)$ depicted in Figure 3.1 models the simple described cooling system where the state space is given by

$$Sta = \{s_{\text{fail}}, s_{\text{off}}, s_{\text{on1}}, s_{\text{on2}}, s_{\text{select}}\}$$

and the action space is formalised by

$$Act = \{\text{fail}, \text{off}, \text{on}, \text{repair}, \text{select}_1, \text{select}_2\}.$$

Intuitively, the state where the cooling system is off is represented by the state $s_{\text{off}}$. When switching on the cooling system in this state, a failure occurs with the small probability $1/10$, i.e., the state $s_{\text{fail}}$ is entered and a repair is necessary before trying to turn on the system the next time. Switching on the cooling system in state $s_{\text{off}}$ is successful with probability $9/10$ where state $s_{\text{select}}$ is entered. Thus, to be precise, we have

$$s_{\text{off}} \to \varphi \quad \text{where} \quad \varphi(\{\text{fail}, s_{\text{fail}}\rangle\}) = \frac{1}{10} \text{ and } \varphi(\{\text{on}, s_{\text{select}}\rangle\}) = \frac{9}{10}.$$

The illustrated STS is hence not simple. In the state $s_{\text{select}}$ there is a non-deterministic choice between two cooling strategies. Depending on the user's choice, the system either enters the state $s_{\text{on1}}$ or the state $s_{\text{on2}}$, i.e.,

$$s_{\text{select}} \to \langle \text{select}_1, \mathit{Dirac}[s_{\text{on1}}] \rangle \quad \text{and} \quad s_{\text{select}} \to \langle \text{select}_2, \mathit{Dirac}[s_{\text{on2}}] \rangle.$$

Both states $s_{\text{on1}}$ and $s_{\text{on2}}$ represent the phase where the cooling system is working, i.e., where the heat produced by the server is dissipated. After a while, to safe energy costs, the cooling system switches off itself depending on the selected cooling strategy. ⌟

The STS in the previous example has a finite state space and a finite action space. Including additional features for a refinement of the model such as timing behaviour leads to a more complex STS where the action space may be uncountable:

**Example 45.** The discussions of Example 44 are continued by including information concerning the timing behaviour. In particular, we want to include the fact that the time between entering and leaving the state $s_{\text{on1}}$, i.e., the time where the cooling system dissipates the heat produced, is exponential distributed with rate $r_1$. Accordingly, the sojourn time in state $s_{\text{on2}}$ is exponentially distributed with rate $r_2$. For this purpose, we extend the underlying action space as follows:

$$\mathit{Act}' = \mathbb{R}_{\geq 0} \times \mathit{Act}$$

where $\mathit{Act}$ is as in Example 44. Note, the set $\mathit{Act}'$ forms a Polish space by Example 2 (5). Intuitively, executing the action $\langle t, \text{off} \rangle$ in state $s_{\text{on1}}$ means that the cooling system switches off itself after it has dissipated heat for $t$ time units. The transition relation $\to'$ for the new STS needs also be adapted. For instance, there is the transition $s_{\text{on}_1} \to' \varphi_1$ where $\varphi_1 \in \mathit{Prob}[\mathit{Act}' \times \mathit{Sta}]$ is determined by the exponential distribution with rate $r_1$ and thus, for every interval interval $[\underline{t}, \overline{t}] \subseteq \mathbb{R}_{\geq 0}$,

$$\varphi(([\underline{t}, \overline{t}] \times \{\text{off}\}) \times \{s_{\text{off}}\}) = \int_{\underline{t}}^{\overline{t}} r_1 \cdot e^{-r_1 \cdot t} \, dt = r_1 \cdot (e^{-r_1 \cdot \underline{t}} - e^{-r_1 \cdot \overline{t}}),$$

i.e., the real number $\varphi(([\underline{t}, \overline{t}] \times \{\text{off}\}) \times \{s_{\text{off}}\})$ indicates the probability of entering state $s_{\text{off}}$ with a delay of at least $\underline{t}$ and at most $\overline{t}$ time units. In all the other states of the cooling system, there are no further timing information available. Thus, there is a non-deterministic choice in the corresponding states e.g., for every $t \in \mathbb{R}_{\geq 0}$ we have the transition

$$s_{\text{off}} \to \varphi_t \quad \text{where} \quad \varphi_t(\{\langle \langle t, \text{fail} \rangle, s_{\text{fail}} \rangle\}) = \frac{1}{10} \text{ and } \varphi_t(\{\langle \langle t, \text{on} \rangle, s_{\text{select}} \rangle\}) = \frac{9}{10}.$$

This means that the cooling system may be switched off for an arbitrary amount of time. Note that the set $\mathit{Enabled}[s_{\text{off}}]$ is uncountable. ⌟

The action space of the STS presented in the previous example is uncountable. Moreover, the example requires sampling from continuous distributions as well as non-deterministic choices from an uncountable number of alternatives. However, the state space is still finite. Including dependencies in the model caused by variables representing physical quantities such as the internal temperature of the server leads to STSs with uncountable state spaces. This is illustrated by the following example:

**Example 46.** We proceed our discussions concerning Examples 44 and 45. Assume that the cooling system is equipped with a sensor measuring the internal temperature of the server. We include a variable representing the actual internal temperature of the server. The continuous evolution of this variable can be modelled by differential equations or differential inclusions for the different modes of the server and the cooling system. Guards for actions can be also included, e.g., specifying that the cooling system can be only switched on if the temperature increases above a certain threshold. Using invariants in locations, it can be moreover modelled that the cooling system needs to be switched off if the temperature decreases below a given threshold. These features lead to the model of hybrid systems called hybrid automata. We refer to [Hen96] for an excellent survey on hybrid automata, in particular, including a nice and simple example of a thermostat that can be easily adapted for the running example of this thesis. ⌟

To model temperature fluctuations and sensor inaccuracies, the formalism sketched in the previous example can be further extended by including, e.g, stochastic jumps and stochastic differential equations. STSs obtained by a (stochastic) hybrid automaton as sketched in the previous example in particular include uncountable state spaces representing actual values of real-valued variables. To model and analyse such complex stochastic systems, a general modelling formalism as introduced by Definition 43 is hence necessary.

## 3.2 Simulation and bisimulation using weight functions

Two states $s_a$ and $s_b$ are bisimilar to each other if they exhibit the same stepwise behaviour with respect to the basic observables given by the set of actions. A state $s_b$ simulates another state $s_a$ provided $s_b$ can mimic all the stepwise behaviour of $s_a$, however, there are may transitions of $s_b$ that cannot be performed by $s_a$. Simulation and bisimulation yield local notions to reason about the branching-time structure of states in the sense that only states and their successor distributions are taken into account.

For the following material it is appropriate to recall the definition for the weight lifting of a relation from Section 2.6. STSs include distributions over action-state pair that is why we need the following additional notion. For every Polish spaces *Sta* and *Act* as well as every relation $R \subseteq Sta \times Sta$ the *action lifting (of R concerning Act)* is given by the relation $R^{Act} \subseteq (Act \times Sta) \times (Act \times Sta)$ defined as follows: for every actions $act_a, act_b \in Act$ and states $s_a, s_b \in Sta$,

$$\langle \langle act_a, s_a \rangle, \langle act_b, s_b \rangle \rangle \in R^{Act} \quad \text{iff} \quad act_a = act_b \text{ and } \langle s_a, s_b \rangle \in R.$$

The relation $R$ is extended to a relation over $Act \times Sta$ by exploiting the diagonal relation on the action space $Act$. Intuitively, the definition of the action lifting of a relation guarantees that probability measures $\varphi_a, \varphi_b \in Prob[Act \times Sta]$ with $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}$ admit the same observable. To be more precise, let $\alpha_a$ and $\alpha_b$ be the probability measures on $Act$ obtained by projecting $\varphi_a$ and $\varphi_b$ on the action space $Act$, respectively, i.e., for every Borel set $A \subseteq Act$ one has $\alpha_a(A) = \varphi_a(A \times Sta)$ and $\alpha_b(A) = \varphi_b(A \times Sta)$. By the definition of the action lifting, for every set $A \subseteq Act$ it is easy to see that the set $A \times Sta$ is $R^{Act}$-stable. According to Remark 34, it hence holds $\alpha_a = \alpha_b$ provided $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}$. This discussion is deepened in Section 3.3 where we consider the relation $(R^{Act})^{\mathrm{wgt}}$ from another point of view.

**Definition 47.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS. We call a relation $R \subseteq Sta \times Sta$ on the state space a *simulation (for $\mathcal{T}$)* provided for every states $s_a, s_b \in Sta$ and probability measure $\varphi_a \in Prob[Act \times Sta]$ the following implication holds: if

$$\langle s_a, s_b \rangle \in R \quad \text{and} \quad s_a \rightarrow \varphi_a,$$

then there exists a probability measure $\varphi_b \in Prob[Act \times Sta]$ such that

$$s_b \rightarrow \varphi_b \quad \text{and} \quad \langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}.$$

A relation $R \subseteq Sta \times Sta$ is said to be a *bisimulation (for $\mathcal{T}$)* provided both relations $R$ and its inverse $R^{-1}$ are simulations. ⌟

$$\begin{array}{ccc} s_a & R & s_b \\ \downarrow & & \\ \varphi_a & & \end{array} \qquad \text{can be completed to} \qquad \begin{array}{ccc} s_a & R & s_b \\ \downarrow & & \downarrow \\ \varphi_a & (R^{Act})^{\mathrm{wgt}} & \varphi_b \end{array}$$

Figure 3.2: Condition for a simulation $R$.

The condition on a simulation is illustrated by Figure 3.2. An analogous illustration can be given for a bisimulation. The notions of simulation and bisimulation induce the following relations over the state space of the STS under consideration. Consider an STS $\mathcal{T} = (Sta, Act, \rightarrow)$. The *simulation preorder* (*for* $\mathcal{T}$) and the *bisimulation equivalence* (*for* $\mathcal{T}$) are the binary relation $\preceq$ and $\simeq$ on the state space *Sta*, respectively, given as follows: for every states $s_a, s_b \in Sta$,

$$\begin{aligned} s_a \preceq s_b \quad &\text{iff} \quad \text{there exists a simulation } R \text{ with } \langle s_a, s_b \rangle \in R, \\ s_a \simeq s_b \quad &\text{iff} \quad \text{there exists a bisimulation } R \text{ with } \langle s_a, s_b \rangle \in R. \end{aligned}$$

Note that the simulation preorder $\preceq$ is the coarsest simulation for $\mathcal{T}$, i.e., the relation $\preceq$ is a simulation for $\mathcal{T}$ and moreover, every simulation for $\mathcal{T}$ is a subset $\preceq$. An analogue statement holds for bisimulations. The latter two claims follow directly from the easy observation that the concepts of action and weight liftings induce a monotonically increasing operator on the subsets of $Sta \times Sta$, more precisely, for every relations $R_1$ and $R_2$ over *Sta* we have that $R_1 \subseteq R_2$ implies $((R_1)^{Act})^{\mathrm{wgt}} \subseteq ((R_2)^{Act})^{\mathrm{wgt}}$.

**Remark 48.** Let *Sta* and *Act* be countable sets, $R \subseteq Sta \times Sta$ be a relation, as well as $\varphi_a, \varphi_b \in Prob[Act \times Sta]$. For every state $s \in Sta$ define $[s]_R = \{s' \in Sta \,;\, \langle s, s' \rangle \in R\}$. Moreover, let $\mathcal{C}$ be the smallest family of subsets of *Sta* consisting the set $[s]_R$ for every $s \in Sta$ and that is additionally closed under finite intersections and finite unions. Obviously, if the relation $R$ is an equivalence, then the family $\mathcal{C}$ and the family of all the finite unions of equivalence classes concerning $R$ are the same. With the introduced notions we have following two statements:

(1) Assuming the relation $R$ is a preorder, the following equivalence holds:

$$\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}} \quad \text{iff} \quad \begin{aligned}[t] &\text{for every action } act \in Act \text{ and every } S \in \mathcal{C}, \\ &\varphi_a(\{act\} \times S) \leq \varphi_b(\{act\} \times S). \end{aligned}$$

(2) If the relation $R$ is an equivalence, then we have the statement below:

$$\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}} \quad \text{iff} \quad \begin{aligned}[t] &\text{for every action } act \in Act \text{ and every } S \in \mathcal{C}, \\ &\varphi_a(\{act\} \times S) = \varphi_b(\{act\} \times S). \end{aligned}$$

We provide a proof for claim (1). The following argument relies on Theorem 39. For this purpose define $\mathcal{C}' = \{\{act\} \times S \,;\, act \in Act \text{ and } S \in \mathcal{C}\}$. Since the sets *Act* and

*Sta* are finite, it easily follows that the family $\mathcal{C}'$ is countable. Moreover, assuming the relation $R$ be preorder, it is easy to see that for every action-state pairs $s'_a, s'_b \in Act \times Sta$ it holds $\langle s'_a, s'_b \rangle \in R^{Act}$ iff for every $S' \in \mathcal{C}'$ we have that $s'_a \in S'$ implies $s'_b \in S'$. Putting things together, the relation $R^{Act}$ is weakly smooth and the family $\mathcal{C}'$ yields a witness. Consequently, claim (1) follows directly from Theorem 39. A proof for statement (2) can be obtained in the same way using Theorem 40 instead of Theorem 39. Moreover, assuming the relation $R$ is an equivalence, it suffices to consider the quotient space of the relation $R$ instead of the family $\mathcal{C}$ in statement (2) by Theorem 40. ⌟

In the introduction of this chapter we mentioned that STSs can be seen as an generalisation of the generative stochastic model in the classification of [GSS95]. In this context, the previous remark shows that Definition 47 yields a conservative generalisation of generative bisimulation defined in [GSS95] (see also [Tin07]).

**Remark 49.** Let *Sta* and *Act* be Polish spaces, $R \subseteq Sta \times Sta$ be a relation, as well as $\varphi_a, \varphi_b \in Prob[Act \times Sta]$. Consider $act_a, act_b \in Act$ and $\mu_a, \mu_b \in Prob[Sta]$ with

$$\varphi_a = Dirac[act_a] \otimes \mu_a \quad \text{and} \quad \varphi_b = Dirac[act_b] \otimes \mu_b.$$

Then we have the following equivalence:

$$\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}} \quad \text{iff} \quad act_a = act_b \text{ and } \langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}.$$

The claim is an immediate consequence of Remark 31 as well as Example 35. ⌟

According to the previous remark, the notion of simulation in Definition 47 can be rephrased as follows for a simple STS $\mathcal{T} = (Sta, Act, \rightarrow)$. A relation $R$ is a simulation precisely when for every states $s_a, s_b \in Sta$, action $act \in Act$, and probability measure $\mu_a \in Prob[Sta]$ such that

$$\langle s_a, s_b \rangle \in R \quad \text{and} \quad s_a \rightarrow \langle act, \mu_a \rangle$$

there exists $\mu_b \in Prob[Sta]$ with

$$s_b \rightarrow \langle act, \mu_b \rangle \quad \text{and} \quad \langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}.$$

A similar characterisation can be given for bisimulation. Consequently, the classical notions for simulation and bisimulation for simple STSs can be derived from our approach using the action-lifting of relation.

## 3.3 Two-step view on distributions over action-state pairs

As the transition relation of STSs include proper distribution over action-state pairs, we rely on the action lifting $R^{Act}$ of a relation $R$ in order to obtain appropriate simulation and bisimulation notions. Recall, probability measures $\varphi_a, \varphi_b \in Prob[Act \times Sta]$ are related in terms of the relation $R$ provided one has $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}$. We present an alternative approach of relating $\varphi_a$ and $\varphi_b$ next. It is then shown that this alternative approach enables a characterisation of the introduced simulation and bisimulation notions. The following discussion can be seen as additional material providing a deeper understanding for the action lifting of a relation.

To explain the starting point of the following discussions, let $\mathcal{T}$ be an STS whose state and action space are given by $Sta$ and $Act$, respectively. Moreover, pick a probability measure $\varphi \in Prob[Act \times Sta]$. Sampling an action-state pair according to $\varphi$ in an execution of $\mathcal{T}$ can be also intuitively realised by the following two-step procedure: first of all, an action $act$ is sampled according to $\alpha$ where the probability measure $\alpha$ is obtained by projecting $\varphi$ onto the action space, i.e., for every Borel set $A \subseteq Act$ it holds $\alpha(A) = \varphi(A \times Sta)$. After that, a successor state $s'$ is sampled according to $f(act)$ where, roughly speaking, the probability measure $f(act)$ on $Sta$ results from $\varphi$ by conditioning with respect to the previously sampled action $act$ (see disintegration theorem in Section 2.1 and also Example 7). More precisely, $f$ is a Borel function with domain $Act$ and codomain $Prob[Sta]$ such that $\varphi = \alpha \rtimes f$, i.e., for every Borel sets $A \subseteq Act$ and $S \subseteq Sta$ one has

$$\varphi(A \times S) = \int_A f(act)(S)\, d\alpha(act).$$

This procedure is illustrated by means of the following simple example:

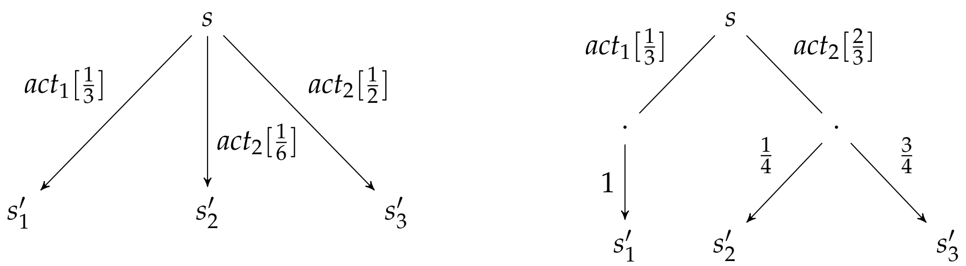


Figure 3.3: Two illustrations of the same probability measure on $Act \times Sta$.

**Example 50.** Consider the sets $Sta = \{s, s'_1, s'_2, s'_3\}$ and $Act = \{act_1, act_2\}$ as well as define the probability measure $\varphi \in Prob[Act \times Sta]$ by

$$\begin{aligned}
&\varphi(\{act_1\} \times \{s'_1\}) = 1/3,\\
&\varphi(\{act_2\} \times \{s'_2\}) = 1/6,\\
&\varphi(\{act_2\} \times \{s'_3\}) = 1/6.
\end{aligned}$$

The left-hand side of Figure 3.3 provides an illustration of this distribution on $Act \times Sta$. Let $\alpha \in Prob[Act]$ be the probability measure obtained by projecting $\varphi$ onto the action space, i.e.,

$$\begin{aligned}
&\alpha(\{act_1\}) = 1/3,\\
&\alpha(\{act_2\}) = 1/6 + 1/2 = 2/3.
\end{aligned}$$

Assuming $f\colon Act \to Prob[Sta]$ is a function such that $\varphi = \alpha \rtimes f$, then the following three identities hold (see also Example 7):

$$\begin{aligned}
&f(act_1)(\{s'_1\}) = \varphi(Act \times \{s'_1\} \mid \{act_1\} \times Sta) = 1,\\
&f(act_2)(\{s'_2\}) = \varphi(Act \times \{s'_2\} \mid \{act_2\} \times Sta) = 1/4,\\
&f(act_2)(\{s'_3\}) = \varphi(Act \times \{s'_3\} \mid \{act_2\} \times Sta) = 3/4.
\end{aligned}$$

The right-hand side of Figure 3.3 sketches the two-step view on the probability measure $\varphi$ where an action is sampled with respect to the probability measure $\alpha$ first and then a state is sampled with respect to a probability measure determined by the function $f$. ⌟

**Definition 51.** Let $Act$ and $Sta$ be Polish spaces, $R \subseteq Sta \times Sta$ be a relation, as well as $\varphi_a, \varphi_b \in Prob[Act \times Sta]$ be probability measures. Let $\alpha_a, \alpha_b \in Prob[Act]$ be probability measures and $f_a\colon Act \to Prob[Sta]$ and $f_b\colon Act \to Prob[Sta]$ be Borel functions with

$$\varphi_a = \alpha_a \rtimes f_a \quad \text{and} \quad \varphi_b = \alpha_b \rtimes f_b.$$

We say that probability measures $\varphi_a, \varphi_b \in Prob[Act \times Sta]$ *are related by $R$ from the two-step point of view*, denoted by $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\text{two}}$, if the following two statements hold:

(1) $\alpha_a = \alpha_b$.

(2) There are a Borel set $A \subseteq Act$ such that

$$\alpha_a(A) = 1 \qquad \text{and} \quad \langle f_a(act), f_b(act) \rangle \in R^{\text{wgt}} \text{ for every } act \in A.$$

⌟

A representation of the probability measures $\varphi_a$ and $\varphi_b$ in terms of a semi-product measure, i.e., $\varphi_a = \alpha_a \rtimes f_a$ and $\varphi_b = \alpha_b \rtimes f_b$, always exists thanks to the disintegration theorem (see Section 2.1). Recall, the disintegration theorem also yields that the functions $f_a$ and $f_b$ are almost surely uniquely determined, i.e., provided $f'_a \colon Act \to Prob[Sta]$ is a Borel function with $\varphi_a = \mu_a \rtimes f'_a$, there exists a Borel set $A_a \subseteq Act$ such that $\alpha_a(A_a) = 1$ and $f_a(act) = f'_a(act)$ for every action $act \in A_a$. The same fact applies for the function $f_b$. Roughly speaking, instead of requiring a single weight function for $(\varphi_a, R^{Act}, \varphi_b)$ as in the last section, Definition 51 relates $\varphi_a$ and $\varphi_b$ with two steps that correspond to the previously sketched two-step view on distributions over action-state pairs. Indeed, the distributions $\alpha_a$ and $\alpha_b$ on the action space are compared in a first step. After that, all the relevant distributions on the state space are considered. With regard to this discussion, we think that Definition 51 is intuitive. Fortunately, the approaches of this and the previous section fit together as the following theorem shows:

**Theorem 52.** *Let Act and Sta be Polish spaces, $\varphi_a, \varphi_b \in Prob[Act \times Sta]$ be probability measures, and $R \subseteq Sta \times Sta$ be a Souslin set. Then we have the following equivalence:*

$$\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}} \quad \textit{iff} \quad \langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{two}}.$$

⌟

This theorem completely characterises the weight lifting of the action lifting of a relation that yields a Souslin set in the product space. This Souslin condition is in particular fulfilled if the Polish space *Sta* is countable. The next chapters of this thesis moreover show that Souslin sets yield a convenient setting for the analysis of stochastic systems with uncountable state and action spaces. The statement $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}$ refers to the existence of one single weight function being a probability measure on $(Act \times Sta) \times (Act \times Sta)$. In comparison to that, the statement $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{two}}$ includes multiple weight functions yielding probability measures on $Sta \times Sta$. Roughly speaking, whereas the proof of one direction of the equivalence requires the merging of a possible uncountable number of weight functions, the argument for the other direction is accompanied by a decomposition of one weight function into a number of weight function.

**Proof of section's main result.** Our argument for Theorem 52 requires the following two auxiliary lemmas.

**Lemma 53.** *Let $X$ and Act be Polish spaces and $R \subseteq X \times X$ be a Souslin set. Then the relation $R^{Act}$ is Souslin in $(Act \times X) \times (Act \times X)$.*

*Proof.* Since $Act$ is required to be a Polish space, the diagonal relation on $Act$, i.e., the relation defined by $\{\langle act_a, act_b\rangle \in Act \times Act \,;\, act_a = act_b\}$, is Borel and hence Souslin in $Act \times Act$ (see also Remark 10 and Example 35). From this the claim is obtained easily. $\square$

**Lemma 54.** *Let $X$ and $Y$ be Polish spaces, $R \subseteq X \times X$ be a Souslin set, and $f_a \colon Y \to Prob[X]$ and $f_b \colon Y \to Prob[X]$ be Borel functions. Moreover, let $\mu_Y \in Prob[Y]$ be a probability measure such that there exists a Borel set $B_Y \subseteq Y$ with*

$$\mu_Y(B_Y) = 1 \quad \textit{and} \quad \langle f_a(y), f_b(y)\rangle \in R^{\mathrm{wgt}} \textit{ for every } y \in Y.$$

*Then there exists a Borel function $g \colon Y \to Prob[X \times X]$ and a Borel set $B'_Y \subseteq Y$ such that the following two properties are fulfilled:*

(1) $\mu_Y(B'_Y) = 1$.

(2) *$g(y)$ is a weight function for $(f_a(y), R, f_b(y))$ for every $y \in B'_Y$.*

*Proof.* The claim follows by a standard application of the measurable-selection principle stated in Theorem 21. Indeed, defining the set-valued function $F \colon Y \rightsquigarrow Prob[X \times X]$,

$$F(y) = \{W \in Prob[X \times X] \,;\, W \text{ is weight function for } (f_a(y), R, f_b(y))\},$$

it suffices to justify that there exists a Borel $W$-selection of $F$.

Thanks to the assumptions of the lemma, there exists a Borel set $B_Y \subseteq Y$ such that $\mu_Y(B_Y) = 1$ and $F(y) \neq \varnothing$ for all $y \in B_Y$. By Theorem 21, it remains to show that the set $Rel[F]$ is Souslin in $Y \times Prob[X \times X]$. For reasons of clarity abbreviate

$$Z = Prob[X] \times Prob[X] \times Prob[X \times X].$$

Define the Borel function $\zeta \colon Y \times Prob[X \times X] \to Z$,

$$\zeta(y, W) = \langle f_a(y), f_b(y), W\rangle.$$

Introduce $M \subseteq Z$ by

$$M = \{\langle \mu_a, \mu_b, W\rangle \in Z \,;\, W \text{ is a weight function for } (\mu_a, R, \mu_b)\}$$

It is easy to observe the identity

$$Rel[F] = \zeta^{-1}(M).$$

Define the Borel functions $\zeta_a\colon X \times X \to X$, $\zeta_a(x_a, x_b) = x_a$ and $\zeta_b\colon X \times X \to X$, $\zeta_b(x_a, x_b) = x_b$. Thanks to Lemma 12, for every probability measures $\mu_a, \mu_b \in Prob[X]$ and $W \in Prob[X \times X]$ we have that $W$ is a weight function for $(\mu_a, R, \mu_b)$ precisely when the following three statements hold:

$$\langle W, \mu_a \rangle \in Graph[(\zeta_a)_\sharp], \quad \langle W, \mu_b \rangle \in Graph[(\zeta_b)_\sharp], \quad \text{and} \quad W^{\text{out}}(R) = 1.$$

By Remarks 8 (1) and 10 (6), the set $M$ is hence Souslin in $Z$. According to Remark 10 (5), we finally obtain that the set $Rel[F]$ is Souslin in $Y \times Prob[X \times X]$. As pointed our earlier, this completes a proof. □

*Proof of Theorem 52.* Throughout this proof let $\alpha_a, \alpha_b \in Prob[Act]$ be the probability measures defined as follows: for every Borel set $A \subseteq Act$ let

$$\alpha_a(A) = \varphi_a(A \times Sta) \quad \text{and} \quad \alpha_b(A) = \varphi_b(A \times Sta).$$

Moreover, let $f_a\colon Act \to Prob[Sta]$ and $f_b\colon Act \to Prob[Sta]$ be Borel functions such that

$$\varphi_a = \alpha_a \rtimes f_a \quad \text{and} \quad \varphi_b = \alpha_b \rtimes f_b.$$

According to the Disintegration theorem (see Section 2.1), functions $f_a$ and $f_b$ with the stated properties indeed exist.

*Implication from left to right.* Assume that $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\text{wgt}}$. Thanks to Remark 34, for every Borel set $A \subseteq Act$ it holds $\alpha_a(A) = \alpha_b(A)$ as the set $A$ is $R^{Act}$-stable. This yields the identity $\alpha_a = \alpha_b$. It remains to show statement (2) of Definition 51. For this purpose let $W$ be a weight function for $(\varphi_a, R^{Act}, \varphi_b)$. Moreover, introduce the Borel function $\zeta\colon (Act \times Sta) \times (Act \times Sta) \to Act \times (Sta \times Sta)$,

$$\zeta(act_a, s_a, act_b, s_b) = \langle act_a, s_a, s_b \rangle.$$

Define the probability measure $W' \in Prob[Act \times Sta \times Sta]$ by

$$W' = \zeta_\sharp(W).$$

It is easy to see that for all Borel sets $A \subseteq Act$ it holds $W'(A \times Sta \times Sta) = \alpha_a(A)$. According to the Disintegration theorem (see Section 2.1), there hence exists a Borel function $g\colon Act \to Sta \times Sta$ such that

$$W' = \alpha_a \rtimes g.$$

We prove that there exists a Borel set $A_{\alpha_a} \subseteq Act$ such that the following two properties hold: $\alpha_a(A_{\alpha_a}) = 1$ and for every action $act \in A_{\alpha_a}$ the probability measure $g(act)$ is a weight function for $(f_a(act), R, f_b(act))$. This then shows statement (2) of Definition 51. The following argument proceeds in three steps.

Let $\mathcal{G}$ be a countable generator of the Borel sigma algebra on *Sta* that is additionally closed under finite intersections (see Remark 4). Suppose $\mathcal{G} = \{S_0, S_1, S_2, \ldots\}$ for some Borel sets $S_0, S_1, S_2, \ldots \subseteq Sta$.

*Step* (*a*). We first prove that there exists a Borel set $A' \subseteq Act$ with $\alpha_a(A') = 1$ and so that for every $act \in A'$ and Borel set $S \subseteq Sta$ one has $f_a(act)(S) = g(act)(S \times Sta)$.

For every $n \in \mathbb{N}$ and Borel set $A \subseteq Act$ one easily derives the identity

$$\varphi_a(A \times S_n) = W((A \times S_n) \times (Act \times Sta)) = W'(A \times (S_n \times Sta))$$

and therefore, it holds

$$\int_A f_a(act)(S_n)\, d\alpha_a(act) = \int_A g(act)(S_n \times Sta)\, d\alpha_a(act).$$

By a basic result from measure theory (see Folgerung 9.2.5 in [Sch08]), for every $n \in \mathbb{N}$ there exists a Borel set $A'_n \subseteq Act$ such that $\alpha_a(A'_n) = 1$ and $f_a(act)(S_n) = g(act)(S_n \times Sta)$ for all $act \in A'_n$. Define the Borel set $A' \subseteq Act$ by $A' = \bigcap_{n \in \mathbb{N}} A'_n$. It follows

$$\alpha_a(A') = 1$$

and moreover, relying on Carathéodory extension theorem (see Section 2.1), for every action $act \in A'$ and every Borel set $S \subseteq Sta$ it holds

$$f_a(act)(S) = g(act)(S \times Sta).$$

*Step* (*b*). This proof step is similar to (a). More precisely, using an analogous argument as in (a), we show that there exists a Borel set $A'' \subseteq Act$ with $\alpha_a(A'') = 1$ and such that for every $act \in A''$ and Borel set $S \subseteq Sta$ one has $f_b(act)(S) = g(act)(Sta \times S)$.

Using that the probability measure $W$ is a weight function for $(\varphi_a, R^{Act}, \varphi_b)$, there exists a Borel set $R' \subseteq (Act \times Sta) \times (Act \times Sta)$ such that

$$W(R') = 1 \quad \text{and} \quad R' \subseteq R^{Act}.$$

For every $\langle\langle act_a, s_a\rangle, \langle act_b, s_b\rangle\rangle \in R'$ one has $act_a = act_b$. Consequently, one easily derives that for every $n \in \mathbb{N}$ and Borel set $A \subseteq Act$ it holds $((Act \times Sta) \times (A \times S_n)) \cap R' = ((A \times Sta) \times (Act \times S_n)) \cap R'$ and therefore, we obtain

$$\varphi_b(A \times S_n) = W((Act \times Sta) \times (A \times S_n)) = W'(A \times (Sta \times S_n)).$$

For every $n \in \mathbb{N}$ and Borel set $A \subseteq Act$ one thus has

$$\int_A f_b(act)(S_n)\, d\alpha_a(act) = \int_A g(act)(Sta \times S_n)\, d\alpha_a(act).$$

As in (a), relying on a basic result from measure theory (see Folgerung 9.2.5 in [Sch08]), for every $n \in \mathbb{N}$ there exists a Borel set $A''_n \subseteq Act$ such that $\alpha_a(A''_n) = 1$ and $f_b(act)(S_n) = g(act)(Sta \times S_n)$ for all $act \in A''_n$. Let $A'' \subseteq Act$ be the Borel set given by $A'' = \bigcap_{n\in\mathbb{N}} A''_n$. It follows

$$\alpha_a(A'') = 1$$

and moreover, exploiting Carathéodory extension theorem (see Section 2.1), for every action $act \in A''$ and every Borel set $S \subseteq Sta$ one has

$$f_b(act)(S) = g(act)(Sta \times S).$$

*Step (c).* In this proof step, we show that there exists a Borel set $A''' \subseteq Act$ satisfying the following two properties: $\alpha_a(A''') = 1$ and for every $act \in A'''$ there is a Borel set $R'_{act} \subseteq Sta \times Sta$ with $g(act)(R'_{act}) = 1$ and $R'_{act} \subseteq R$.

As in step (b), let $R' \subseteq (Act \times Sta) \times (Act \times Sta)$ be a Borel set such that

$$W(R') = 1 \quad \text{and} \quad R' \subseteq R^{Act}.$$

Define the Borel function $\xi'\colon (Act \times Sta) \times (Act \times Sta) \to Sta \times Sta$,

$$\xi'(\langle act_a, s_a\rangle, \langle act_b, s_b\rangle) = \langle s_a, s_b\rangle$$

We justify $(W')^{\mathrm{out}}(Act \times \xi'(R')) = 1$ next. For every Borel set $B \subseteq Act \times (Sta \times Sta)$ such that $Act \times \xi'(R') \subseteq B$ it holds $R' \subseteq \xi^{-1}(Act \times \xi'(R')) \subseteq \xi^{-1}(B)$ and therefore, one has the identity

$$W'(B) = W(\xi^{-1}(B)) \geq W(R') = 1.$$

This shows $(W')^{\mathrm{out}}(Act \times \xi'(R')) = 1$.

By Remark 10 (5), the set $Act \times \xi'(R')$ is Souslin in $Act \times (Sta \times Sta)$. According to Lemma 15, there hence exists a Borel set $A''' \subseteq Act$ with

$$\alpha_a(A''') = 1$$

and such that for every $act \in A'''$ it holds

$$(g(act))^{\mathrm{out}}(\xi'(R')) = 1.$$

According to Lemma 12, for every action $act \in A'''$ there exists a Borel set $R'_{act} \subseteq Sta \times Sta$ such that $g(act)(R'_{act}) = 1$ and $R'_{act} \subseteq \xi'(R') \subseteq R$.

*Finishing the argument.* Putting the three steps (a), (b), and (c) together, for every action $act \in A' \cap A'' \cap A'''$ the probability measure $g(act)$ constitutes indeed a weight function for $(f_a(act), R, f_b(act))$. Since $\alpha_a(A' \cap A'' \cap A''') = 1$, we finally derive the implication from left to right of the equivalence to be proven.

*Implication from right to left.* Assume that $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{two}}$. It follows $\alpha_a = \alpha_b$. Let $A \subseteq Act$ be a Borel set and $g\colon Act \to Prob[Sta \times Sta]$ be a Borel function as in Definition 51 (2), i.e., $\alpha_a(A) = 1$ and for every $act \in A$ the probability measure $g(act)$ is a weight function for $(f_a(act), R, f_b(act))$. Introduce the Borel function $\xi\colon Act \times (Sta \times Sta) \to (Act \times Sta) \times (Act \times Sta)$,

$$\xi(act, s_a, s_b) = \langle act, s_a, act, s_b \rangle.$$

Define $W \in Prob[(Act \times Sta) \times (Act \times Sta)]$ by

$$W = \xi_\sharp(\alpha_a \rtimes g).$$

We prove that $W$ is a weight function for $(\varphi_a, R^{Act}, \varphi_b)$.

It is shown that $W$ is a coupling of $(\varphi_a, \varphi_b)$ first. For every Borel sets $A \subseteq Act$ and $S \subseteq Sta$ it holds

$$\begin{aligned} & W((A \times S) \times (Act \times Sta)) \\ = & \int_A g(act)(S \times Sta)\, d\alpha_a(act) \\ = & \int_A f_a(act)(S)\, d\alpha_a(act) \\ = & \varphi_a(A \times S). \end{aligned}$$

For every Borel set $B \subseteq Act \times Sta$ it follows $W(B \times (Act \times Sta)) = \varphi_a(B)$ thanks to Carathéodory uniqueness theorem (see Section 2.1). Using the fact $\alpha_a = \alpha_b$, one analogously obtains the corresponding identity $W((Act \times Sta) \times B) = \varphi_b(B)$ for every Borel set $B \subseteq Act \times Sta$.

Thanks to Lemmas 12 and 53, it remains to show $W^{\text{out}}(R^{Act}) = 1$. To this end let $R' \subseteq (Act \times Sta) \times (Act \times Sta)$ be a Borel set such that $R^{Act} \subseteq R'$. We justify $W(R') = 1$. For every action $act \in Act$ define the relation $R'_{act} \subseteq Sta \times Sta$ by

$$R'_{act} = \{\langle s_a, s_b\rangle \in Sta \times Sta\ ;\ \langle act, s_a, act, s_b\rangle \in R'\}.$$

For every $act \in Act$, as $R'_{act}$ can be viewed as a sections of the relation $R'$, it directly follows that the set $R'_{act}$ is Borel in $Sta \times Sta$ (see Section 2.1). Moreover, for every $act \in Act$ it holds $R \subseteq R'_{act}$ and $Section[\xi^{-1}(R'), act, \cdot] = R'_{act}$. We hence obtain

$$W(R') = \int g(act)(R'_{act})\, d\alpha_a(act) = 1.$$

From this we finally derive the claim. $\square$

## 3.4 Transitivity of simulation and bisimulation relations

We show that simulation preorder and bisimulation equivalence indeed form transitive relations. The latter facts are fundamental for convenient notions in the context of abstractions and equivalences of stochastic systems. The main result of this section is as follows:

**Theorem 55.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS. Then the following two statements hold:*

(1) *The relation $\preceq$ is a preorder on Sta.*

(2) *The relation $\simeq$ is an equivalences on Sta.*

*In particular, both relations $\preceq$ and $\simeq$ are transitive.* ⌟

We emphasise that Theorem 55 holds for every STS, i.e., the result does not involve any measurability restrictions. As already pointed out in the introduction of the chapter, the challenging part of the proof is to establish the transitivity of the relations under consideration. Our argument for transitivity shows that the product of two simulations and of two bisimulations also yields a simulation and bisimulation, respectively. To be more precise, let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and consider two simulations $R_{ab}$ and $R_{bc}$. Define $R_{ac} \subseteq Sta \times Sta$ as the composition of the two relations $R_{ab}$ and $R_{bc}$, i.e., for every states $s_a, s_c \in Sta$ it holds

$$\langle s_a, s_c\rangle \in R_{ac} \quad \text{iff} \quad \text{there is } s_b \in Sta \text{ with } \langle s_a, s_b\rangle \in R_{ab} \text{ and } \langle s_b, s_c\rangle \in R_{bc}.$$

Obviously, if $R_{ac}$ is a simulation, then it directly follows that $\preceq$ is transitive. To show that $R_{ac}$ is a simulation, pick states $s_a, s_c \in \mathit{Sta}$ such that $\langle s_a, s_c\rangle \in R_{ac}$. Let $\varphi_a \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ be so that $s_a \to \varphi_a$. Relying on the definition of $R_{ac}$ and since the relations $R_{ab}$ and $R_{bc}$ are simulations, there are a state $s_b \in \mathit{Sta}$ and probability measures $\varphi_b \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ and $\varphi_c \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ such that the following four statements are fulfilled:

$$s_b \to \varphi_b, \quad \langle \varphi_a, \varphi_b\rangle \in ((R_{ab})^{\mathit{Act}})^{\mathrm{wgt}}, \quad s_c \to \varphi_c, \quad \langle \varphi_b, \varphi_c\rangle \in ((R_{bc})^{\mathit{Act}})^{\mathrm{wgt}}.$$

To conclude that $R_{ac}$ is a simulation, it suffices to argue that there exist a weight function for $(\varphi_a, (R_{ac})^{\mathit{Act}}, \varphi_c)$. For this we exploit the so-called Gluing lemma (see, e.g., Lemma 5.3.2 in [AGS05] or Section 1 in [Vil09]): the intuitive idea is to glue together weight functions for $(\varphi_a, (R_{ab})^{\mathit{Act}}, \varphi_b)$ and $(\varphi_b, (R_{bc})^{\mathit{Act}}, \varphi_c)$ along the common marginal given by $\varphi_b$. This is sketched in detail below.

**Proof of section's main result.** Let $X_a$, $X_b$, and $X_c$ be Polish spaces. For every relations $R_{ab} \subseteq X_a \times X_b$ and $R_{bc} \subseteq X_b \times X_c$ the *composition* (*of* $R_{ab}$ *and* $R_{bc}$) is given by the relation $R_{ab} \diamond R_{bc}$ over $X_a$ and $X_c$ as follows: for every $x_a \in X_a$ and $x_c \in X_c$,

$$\langle x_a, x_c\rangle \in R_{ab} \diamond R_{bc} \quad \text{iff} \quad \text{there is } x_b \in X_b \text{ with } \langle x_a, x_b\rangle \in R_{ab} \text{ and } \langle x_b, x_c\rangle \in R_{bc}.$$

**Lemma 56.** *Let $X_a$, $X_b$, and $X_c$ be a Polish space as well as $R_{ab} \subseteq X_a \times X_b$ and $R_{bc} \subseteq X_b \times X_c$ be Souslin sets. Then the set $R_{ab} \diamond R_{bc}$ is Souslin in $X_a \times X_c$.*

*Proof.* The argument is standard: defining the Borel function $\zeta\colon X_a \times X_b \times X_c \to X_a \times X_c$, $\zeta(x_a, x_b, x_c) = \langle x_a, x_c\rangle$, we immediately obtain the identity

$$\zeta((R_{ab} \times X_c) \cap (X_a \times R_{bc})) = R_{ab} \diamond R_{bc}$$

and hence, Remark 10 (5) yields the claim. $\square$

**Example 57.** Take the notions from Lemma 56 and assume that $R_{ab}$ and $R_{bc}$ are even Borel sets in $X_a \times X_b$ and $X_b \times X_c$, respectively. Interestingly, it turns out that the set $R_{ab} \diamond R_{bc}$ is not Borel in $X_a \times X_c$ in general. To see this let us investigate the following simple example. Suppose a Borel set $R \subseteq \mathbb{R} \times \mathbb{R}$ such that the set $M$ is not Borel in $\mathbb{R}$ (see Remarks 10 (1) and (2)). Here, the relation $M$ is defined as the subset of $\mathbb{R}$ obtained by projecting $R$ onto its first component, i.e., it holds $M = \{r_a \in \mathbb{R}\ ;\ \text{there is } r_b \in \mathbb{R} \text{ with } \langle r_a, r_b\rangle \in R\}$. Define

$$X_a = X_b = X_c = \mathbb{R}, \quad R_{ab} = R, \quad \text{and} \quad R_{bc} = \mathbb{R} \times \mathbb{R}.$$

For every $r_a, r_c \in \mathbb{R}$ it holds

$$\langle r_a, r_c \rangle \in R_{ab} \diamond R_{bc} \quad \text{iff} \quad \text{there is } r_b \in \mathbb{R} \text{ with } \langle r_a, r_b \rangle \in R \quad \text{iff} \quad r_a \in M,$$

and therefore, $R_{ab} \diamond R_{bc} = M \times \mathbb{R}$. The set $R_{ab} \diamond R_{bc}$ is hence not Borel in $\mathbb{R} \times \mathbb{R}$. ⌟

Let us recapitulate the gluing lemma next (see, e.g., Lemma 5.3.2 in [AGS05] or Section 1 in [Vil09]). For this purpose consider three Polish spaces $X_a$, $X_b$, and $X_c$ and moreover, let $W_{ab} \in Prob[X_a \times X_b]$ and $W_{bc} \in Prob[X_b \times X_c]$ be two probability measures. Provided the projections of $W_{ab}$ and $W_{bc}$ onto $X_b$ coincide, i.e., for every Borel set $B_b \subseteq X_b$,

$$W_{ab}(X_a \times B_b) = W_{bc}(B_b \times X_c),$$

the gluing lemma guarantees the existence of a probability measure on $X_a \times X_b \times X_c$ that is compatible with $W_{ab}$ and $W_{bc}$ in terms of the respective projections. More precisely, there exists a probability measure $\mu \in Prob[X_a \times X_b \times X_c]$ such that for all Borel sets $B_{ab} \subseteq X_a \times X_b$ and $B_{bc} \subseteq X_b \times X_c$ it holds

$$\mu(B_{ab} \times X_c) = W_{ab}(B_{ab}) \quad \text{and} \quad \mu(X_a \times B_{bc}) = W_{bc}(B_{bc}).$$

Intuitively, the probability measure $\mu$ is obtained by gluing together the probability measures $W_{ab}$ and $W_{bc}$ along their common marginal.

The gluing lemma as presented before is a standard application of the disintegration theorem (see Section 2.1). For the sake of completeness let us present this argument. Suppose $X_a, X_b, X_c, W_{ab}$, and $W_{bc}$ are given as before so that $W_{ab}(X_a \times B_b) = W_{bc}(B_b \times X_c)$ for all Borel sets $B_b \subseteq X_b$. Define $\mu_b \in Prob[X_b]$ as follows: for every Borel set $B_b \subseteq X_b$ let

$$\mu_b(B_b) = W_{ab}(X_a \times B_b).$$

For every Borel set $B_b \subseteq X_b$ we obviously have

$$\mu_b(B_b) = W_{bc}(B_b \times X_c).$$

According to the disintegration theorem, there exist Borel functions $g_{ba} \colon X_b \to Prob[X_a]$ and $g_{bc} \colon X_b \to Prob[X_c]$ such that

$$W_{ab} = g_{ba} \ltimes \mu_b \quad \text{and} \quad W_{bc} = \mu_b \rtimes g_{bc}.$$

Define $g \colon X_b \to Prob[X_a \times X_c]$ by

$$g(x_b) = g_{ba}(x_b) \otimes g_{bc}(x_b).$$

By Example 6, the function $g$ is Borel. Hence, we can safely define $\mu \in Prob[X_a \times X_b \times X_c]$ as follows: for every Borel sets $B_a \subseteq X_a$, $B_b \subseteq X_b$, and $B_c \subseteq X_c$ let

$$\mu(B_a \times B_b \times B_c) = \mu_b \rtimes g(B_b \times B_a \times B_c).$$

Here, $\mu$ is well-defined according to Carathéodory extension theorem (see Section 2.1). Moreover, for every Borel sets $B_a \subseteq X_a$ and $B_b \subseteq X_b$ it holds

$$\mu(B_a \times B_b \times X_c) = \mu_b \rtimes g(B_b \times B_a \times X_c) = \mu_b \rtimes g_{ba}(B_b \times B_a) = W_{ab}(B_a \times B_b)$$

and similar, for every Borel sets $B_b \subseteq X_b$ and $B_c \subseteq X_c$ one has

$$\mu(X_a \times B_b \times B_c) = \mu_b \rtimes g(B_b \times X_a \times X_c) = \mu_b \rtimes g_{bc}(B_b \times B_c) = W_{bc}(B_b \times B_c).$$

Thus, Carathéodory uniqueness theorem completes a proof for the gluing lemma.

It is interesting to note that the gluing lemma does not hold for arbitrary measurable spaces: the article [Swa96] provides an example including measurable spaces $Y_a$, $Y_b$, and $Y_c$ as well as probability measures on $Y_a \times Y_b$ and $Y_b \times Y_c$ that have a common marginal, however, that cannot be glued together. For the following arguments involving the gluing lemma, it is hence essential to work with Polish spaces.

**Theorem 58.** *Let $X_a$, $X_b$, and $X_c$ be Polish spaces as well as $R_{ab} \subseteq X_a \times X_b$ and $R_{bc} \subseteq X_b \times X_c$ be relations. Then, the following two statements hold:*

(1) $(R_{ab})^{\mathrm{wgt}} \diamond (R_{bc})^{\mathrm{wgt}} \subseteq (R_{ab} \diamond R_{bc})^{\mathrm{wgt}}$.

(2) $(R_{ab})^{\mathrm{wgt}} \diamond (R_{bc})^{\mathrm{wgt}} \supseteq (R_{ab} \diamond R_{bc})^{\mathrm{wgt}}$ *if the two sets $R_{ab}$ and $R_{bc}$ are Souslin in $X_a \times X_b$ and $X_b \times X_c$, respectively.*

*Proof.* Denote the relation product of $R_{ab}$ and $R_{bc}$ by $R_{ac}$, i.e., $R_{ac} = R_{ab} \diamond R_{bc}$. Moreover, suppose probability measures $\mu_a \in Prob[X_a]$ and $\mu_c \in Prob[X_c]$.

*Ad* (1). Suppose a probability measure $\mu_b \in Prob[X_b]$ satisfying $\langle \mu_a, \mu_b \rangle \in (R_{ab})^{\mathrm{wgt}}$ and $\langle \mu_b, \mu_c \rangle \in (R_{bc})^{\mathrm{wgt}}$. Let $W_{ab}$ and $W_{bc}$ be weight functions for $(\mu_a, R_{ab}, \mu_b)$ and $(\mu_b, R_{bc}, \mu_c)$, respectively. Moreover, let $R'_{ab} \subseteq X_a \times X_b$ and $R'_{bc} \subseteq X_b \times X_c$ be Borel sets such that

$$R'_{ab} \subseteq R_{ab}, \quad R'_{bc} \subseteq R_{bc}, \quad W_{ab}(R'_{ab}) = 1, \quad \text{and} \quad W_{bc}(R'_{bc}) = 1.$$

In what follows we provide a weight function for $(\mu_a, R_{ac}, \mu_c)$. For this purpose define the Borel functions $\zeta_{ab} \colon X_a \times X_b \times X_c \to X_a \times X_b$, $\zeta_{ac} \colon X_a \times X_b \times X_c \to X_a \times X_c$, and $\zeta_{bc} \colon X_a \times X_b \times X_c \to X_b \times X_c$ as follows: for every $x_a \in X_a$, $x_b \in X_b$, and $x_c \in X_c$ let

$$\zeta_{ab}(x_a, x_b, x_c) = \langle x_a, x_b \rangle,$$

$$\zeta_{ac}(x_a, x_b, x_c) = \langle x_a, x_c \rangle,$$
$$\zeta_{bc}(x_a, x_b, x_c) = \langle x_b, x_c \rangle.$$

Relying on the gluing lemma (see Lemma 5.3.2 in [AGS05] or Section 1 in [Vil09]), there is a probability measure $\mu \in Prob[X]$ with

$$(\zeta_{ab})_\sharp(\mu) = W_{ab} \quad \text{and} \quad (\zeta_{bc})_\sharp(\mu) = W_{bc}.$$

Introduce the probability $W_{ac} \in Prob[X_a \times X_c]$ by

$$W_{ac} = (\zeta_{ac})_\sharp(\mu).$$

We claim that $W_{ac}$ is a weight function for $(\mu_a, R_{ac}, \mu_c)$. It is easy to see that $W_{ac}$ is a coupling of $(\mu_a, \mu_c)$: for all Borel sets $B_a \subseteq X_a$,

$$W_{ac}(B_a \times X_c) = \mu(B_a \times X_b \times X_c) = W_{ab}(B_a \times X_b) = \mu_a(B_a)$$

and similar, for all Borel sets $B_c \subseteq X_c$,

$$W_{ac}(X_a \times B_c) = \mu(X_a \times X_b \times B_c) = W_{bc}(X_b \times B_c) = \mu_c(B_c).$$

Define the Borel set $B \subseteq X_a \times X_b \times X_c$ by

$$B = (R'_{ab} \times X_c) \cap (X_a \times R'_{bc}).$$

Since $\mu(R'_{ab} \times X_c) = W_{ab}(R'_{ab}) = 1$ and $\mu(X_a \times R'_{bc}) = W_{bc}(R'_{bc}) = 1$, we obtain $\mu(B) = 1$. By Lemma 13, there exists a Borel set $R'_{ac} \subseteq X_a \times X_c$ with $W_{ac}(R'_{ac}) = 1$ and $R'_{ac} \subseteq \zeta_{ac}(B)$. Since

$$R'_{ac} \subseteq \zeta_{ac}(B) \subseteq R'_{ab} \diamond R'_{bc} \subseteq R_{ab},$$

it finally follows that $W_{ac}$ is a weight function for $(\mu_a, R_{ab}, \mu_c)$.

*Ad* (2). Assume that $R_{ab}$ and $R_{bc}$ are Souslin sets in $X_a \times X_b$ and $X_b \times X_c$, respectively. Let $\langle \mu_a, \mu_c \rangle \in (R_{ab})^{\mathrm{wgt}}$. Our task is to provide a probability measure $\mu_b \in Prob[X_b]$ with $\langle \mu_a, \mu_b \rangle \in (R_{ab})^{\mathrm{wgt}}$ and $\langle \mu_b, \mu_c \rangle \in (R_{bc})^{\mathrm{wgt}}$. To this end, let $W_{ac}$ be a weight function for $(\mu_a, R_{ab}, \mu_c)$ and suppose a Borel set $R'_{ac} \subseteq X_a \times X_c$ such that

$$R'_{ac} \subseteq R_{ac} \quad \text{and} \quad W_{ac}(R'_{ac}) = 1.$$

Introduce the set-valued function $F\colon X_a \times X_c \rightsquigarrow X_b$,

$$F(x_a, x_c) = \{x_b \in X_b \,;\, \langle x_a, x_b \rangle \in R_{ab} \text{ and } \langle x_b, x_c \rangle \in R_{bc}\}.$$

It is easy to see that set $Graph[F]$ is Souslin in $(X_a \times X_c) \times X_b$. In addition, since $R'_{ac} \subseteq R_{ac} = R_{ab} \diamond R_{bc}$, for every $\langle x_a, x_c \rangle \in R'_{ac}$ it holds $F(x_a, x_c) \neq \varnothing$. Since $W_{ac}(R'_{ac}) = 1$, we can apply Theorem 21. Therefore, there exists a Borel $W_{ac}$-selection of $F$, say $f$.

Define the Borel function $f' \colon X_a \times X_c \to Prob[X_b]$,

$$f'(x_a, x_c) = Dirac[f(x_a, x_c)].$$

Introduce the Borel function $\zeta \colon X_a \times X_c \times X_b \to X_a \times X_b \times X_c$,

$$\zeta(x_a, x_c, x_b) = \langle x_a, x_b, x_c \rangle.$$

Define $\mu \in Prob[X_a \times X_b \times X_c]$ by

$$\mu = \zeta_\sharp(W_{ac} \rtimes f').$$

Moreover, define $\mu_b \in Prob[X_b]$ as follows: for all Borel sets $B_b \subseteq X_b$ let

$$\mu_b(B_b) = \mu(X_a \times B_b \times X_c).$$

To justify (2), it suffices to prove $\langle \mu_a, \mu_b \rangle \in (R_{ab})^{\mathrm{wgt}}$ and $\langle \mu_b, \mu_c \rangle \in (R_{bc})^{\mathrm{wgt}}$. For reasons of symmetry it is enough to show $\langle \mu_a, \mu_b \rangle \in (R_{ab})^{\mathrm{wgt}}$.

In the remainder let $\zeta_{ab}$, $\zeta_{ac}$, and $\zeta_{bc}$ as in the first part of this proof, i.e., $\zeta_{ab}(x_a, x_b, x_c) = \langle x_a, x_b \rangle$, $\zeta_{ac}(x_a, x_b, x_c) = \langle x_a, x_c \rangle$, and $\zeta_{bc}(x_a, x_b, x_c) = \langle x_b, x_c \rangle$ for all $x_a \in X_a$, $x_b \in X_b$, and $x_c \in X_c$. Define

$$W_{ab} = (\zeta_{ab})_\sharp(\mu).$$

We claim that $W_{ab}$ is a weight function for $(\mu_a, R_{ab}, \mu_b)$.

It is easy to see that $W_{ab}$ is a coupling of $(\mu_a, \mu_b)$. Indeed, for all Borel sets $B_a \subseteq X_a$,

$$W_{ab}(B_a \times X_b) = \mu(B_a \times X_b \times X_c) = W_{ac}(B_a \times X_c) = \mu_a(B_a)$$

and moreover, for every Borel set $B_b \subseteq X_b$,

$$W_{ab}(X_a \times B_b) = \mu(X_a \times B_b \times X_c) = \mu_b(B_b).$$

To conclude that $W_{ab}$ is a weight function for $(\mu_a, R_{ab}, \mu_b)$, it suffices to show the identity $(W_{ab})^{\mathrm{out}}(R_{ab}) = 1$ applying Lemma 12. As $f$ is a Borel $W_{ac}$-selection of $F$, there exists a Borel set $R_f \subseteq X_a \times X_c$ such that

$$W_{ac}(R_f) = 1 \quad \text{and} \quad f(x_a, x_c) \in F(x_a, x_c) \text{ for all } \langle x_a, x_c \rangle \in R_f.$$

Moreover, define $M \subseteq X_a \times X_b \times X_c$ by

$$M = (R_{ab} \times X_c) \cap (X_a \times R_{bc}).$$

Observe, the set $\zeta^{-1}(M)$ constitutes a relation between the sets $X_a \times X_c$ and $X_b$. Moreover, for every $\langle x_a, x_c \rangle \in X_a \times X_b$ we have

$$\mathit{Section}[R_{ab}, x_a, \cdot] \cap \mathit{Section}[R_{bc}, \cdot, x_c] = \mathit{Section}[\zeta^{-1}(M), \langle x_a, x_c \rangle, \cdot].$$

For all $\langle x_a, x_c \rangle \in R_f$ we have $f(x_a, x_c) \in F(x_a, x_c)$ and also, $\langle x_a, f(x_a, x_c) \rangle \in R_{ab}$ and $\langle f(x_a, x_c), x_c \rangle \in R_{bc}$. Therefore, for all $\langle x_a, x_c \rangle \in R_f$ We obtain

$$f(x_a, x_c) \in \mathit{Section}[\zeta^{-1}(M), \langle x_a, x_c \rangle, \cdot]$$

and hence,

$$(f'(x_a, x_c))^{\mathrm{out}}(\mathit{Section}[\zeta^{-1}(M), \langle x_a, x_c \rangle, \cdot]) = 1.$$

Since $W_{ac}(R_f) = 1$ and as the set $M$ is Souslin in $X_a \times X_b \times X_c$, Lemma 15 yields the identity $(W_{ac} \rtimes f')^{\mathrm{out}}(M) = 1$. From this we conclude $\mu(M) = 1$. According to Lemma 13, we obtain $(W_{ab})^{\mathrm{out}}(\zeta_{ab}(M)) = 1$. As $\zeta_{ab}(M) \subseteq R_{ab}$, it finally follows $(W_{ab})^{\mathrm{out}}(R_{ab}) = 1$. This completes our argument. $\square$

We solely need the gluing lemma in our proof for Theorem 58 (1). Our argumentation for Theorem 58 (2) involves the measurable-selection principle stated in Theorem 21 for the construction of the probability measure $\mu_b$. In order to fulfil the assumptions of Theorem 21, we need the Souslin requirements for the given relations $R_{ab}$ and $R_{bc}$. It would be interesting whether Theorem 58 (2) still holds if one drops these additional Souslin requirements. However, this discussion does not affect the main result of this section as its proof relies only on the first part of Theorem 58:

*Proof of Theorem 55.* Thanks to Example 35, it is easy to see that the diagonal relation on $\mathit{Sta}$, i.e., $\mathit{Diag} = \{\langle s_a, s_b \rangle \in \mathit{Sta} \times \mathit{Sta} \; ; \; s_a = s_b\}$, forms a bisimulation. As a consequence, the relations $\preceq$ and $\simeq$ are reflexive. Moreover, it is easy to see that relations $\preceq$ is symmetric. It remains to show that the relations under consideration are transitive.

We justify the transitivity of $\preceq$. Let $s_a, s_b, s_c \in \mathit{Sta}$ be states such that $s_a \preceq s_b$ and $s_b \preceq s_c$. Our task is to justify $s_a \preceq s_c$. Let $R_{ab}$ and $R_{bc}$ be simulations satisfying $\langle s_a, s_b \rangle \in R_{ab}$ and $\langle s_b, s_c \rangle \in R_{bc}$. Denote the relation product of $R_{ab}$ and $R_{bc}$ by $R_{ac}$, i.e., $R_{ac} = R_{ab} \diamond R_{bc}$. In the remainder of this proof, we argue that $R_{ac}$ is a simulation. Note, as we have $\langle s_a, s_c \rangle \in R_{ac}$, it follows $s_a \preceq s_c$ and thus, we obtain the transitivity of $\preceq$.

Let $\langle s_a, s_c\rangle \in R_{ac}$ and $\varphi_a \in Prob[Act \times Sta]$ be such that $s_a \to \varphi_a$. Suppose $s_b \in Sta_b$ with $\langle s_a, s_b\rangle \in R_{ab}$ and $\langle s_b, s_c\rangle \in R_{bc}$. As the relations $R_{ab}$ and $R_{bc}$ are simulations, there are $\varphi_b \in Prob[Act \times Sta]$ and $\varphi_c \in Prob[Act \times Sta]$ such that the following four statements are fulfilled:

$$s_b \to \varphi_b, \quad \langle \varphi_a, \varphi_b\rangle \in ((R_{ab})^{Act})^{\mathrm{wgt}}, \quad s_c \to \varphi_c, \quad \langle \varphi_b, \varphi_c\rangle \in ((R_{bc})^{Act})^{\mathrm{wgt}}.$$

Since $(R_{ab})^{Act} \diamond (R_{bc})^{Act} \subseteq (R_{ac})^{Act}$, Theorem 58 (1) yields

$$((R_{ab})^{Act})^{\mathrm{wgt}} \diamond ((R_{bc})^{Act})^{\mathrm{wgt}} \subseteq ((R_{ab})^{Act} \diamond (R_{bc})^{Act})^{\mathrm{wgt}} \subseteq ((R_{ac})^{Act})^{\mathrm{wgt}}$$

and therefore, $\langle \varphi_a, \varphi_c\rangle \in ((R_{ac})^{Act})^{\mathrm{wgt}}$. We conclude that the relation $R_{ac}$ is a simulation. Transitivity of the relation $\simeq$ can be shown analogously. □

**The transitive closure of relations.** We conclude this section with a further application of Theorem 58 referring to transitive closures of relations. Inspecting the definitions for a simulation or bisimulation in Section 3.2 again, we require no additional assumptions on the relations under consideration compared to, e.g., [Seg95, Des99]. More precisely, given an STS $\mathcal{T}$ with state space *Sta*, it is neither required that a simulation is a preorder relation on *Sta* nor that a bisimulation is an equivalence relation on *Sta*. However, relying on the transitive closure of relations, these properties can be assumed without restricting the generality. Recall, for every relation $R \subseteq X \times X$ over some set $X$ the transitive closure is given by

$$TransClosure[R] = \bigcap\{R' \subseteq X \times X \,;\, R \subseteq R' \text{ and } R' \text{ is transitive}\}.$$

As the universal relation $X \times X$ is transitive and an arbitrary intersection of transitive relations is transitive again, the relation $TransClosure[R]$ is indeed transitive and satisfies the inclusion $R \subseteq TransClosure[R]$.

**Proposition 59.** *Let $\mathcal{T}$ be an STS with state space Sta. Define the relation $Diag \subseteq Sta \times Sta$ by $Diag = \{\langle s_a, s_b\rangle \in Sta \times Sta \,;\, s_a = s_b\}$. Then, the following two statements hold:*

(1) *For every simulation $R$ the relation $TransClosure[Diag \cup R]$ is a simulation that contains the relation $R$ and is a preorder on Sta.*

(2) *For every bisimulation $R$ the relation $TransClosure[R \cup R^{-1}]$ is a bisimulation that contains the relation $R$ and is an equivalence on Sta.*

*Proof.* *Ad* (1). Let $R$ be a simulation. Define the relation

$$R' = \bigcup_{n \in \mathbb{N}} R_n$$

where

$$R_0 = \mathit{Diag} \cup R \quad \text{and} \quad R_n = R_{n-1} \diamond R_{n-1} \text{ for every } n \in \mathbb{N} \setminus \{0\}.$$

Observe, for every $n \in \mathbb{N}$ it holds $R_n \subseteq R_{n+1}$. An easy induction over the naturals together with Theorem 58 (1) yield that $R_n$ is a simulation for all $n \in \mathbb{N}$. Here, thanks to Example 35, it directly follows that $R_0$ forms a simulation. It follows that the relation $R'$ is a simulation. This completes our proof as $\mathit{TransClosure}[\mathit{Diag} \cup R] = R'$.

*Ad* (2). The argument is analogous to the first part of this proof: simply replace the definition for the relation $R_0$ by $R_0 = R \cup R^{-1}$. Again, if $R$ is a bisimulation, then Example 35 justifies that $R_0$ yields a bisimulation. □

The insight of Proposition 59 is also of practical relevance: it may be difficult or expensive to determine a (bi)simulation that is transitive or also to check whether a given (bi)simulation is transitive for some given STS. However, the proven proposition provides a way to transform a (bi)simulation into a coarser one that is transitive in addition.

## 3.5 Combined-transition relation

This section basically revisits the mathematical theory in Section 2.4 in the context of stochastic models concerning barycentres and convex hulls of probability measures. So it is advisable to briefly recall ideas and definitions of the previously mentioned section. More precisely, the following material studies the induced combined-transition relation of an STS, which becomes a crucial ingredient of the definition of schedulers in the next section. A combined transition in an STS is obtained by a combination of a possible uncountable number of transitions relying on the concept of barycentres. Intuitively, every assignment of probabilities to enabled distributions in a picked state induces a combined transition at that state. The precise definition is as follows:

**Definition 60.** For every STS $\mathcal{T} = (\mathit{Sta}, \mathit{Act}, \rightarrow)$ the *combined-transition relation*

$$\Rightarrow \subseteq \mathit{Sta} \times \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$$

consists of exactly those pairs $\langle s, \varphi \rangle \in Sta \times Prob[Act \times Sta]$ that satisfy the following statement: there are $\beta \in Prob[Prob[Act \times Sta]]$ and a Borel set $P \subseteq Prob[Act \times Sta]$ with

$$P \subseteq Enabled[s], \quad \beta(P) = 1, \quad \text{and} \quad \varphi = Barycen(\beta).$$

⌟

The introduced notion of a combined transition is a conservative extension of the formalism presented in Section 4.2.2 in [Seg95] where one focuses on systems with countable state spaces only and Definition 8.2 in [Cat05] where one regards simple STSs only. As for the transition relation of an STS, we write $s \Rightarrow \varphi$ rather than $\langle s, \varphi \rangle \in \Rightarrow$. For every state $s \in Sta$ we moreover define the set $CombEnabled[s] \subseteq Prob[Act \times Sta]$ by

$$CombEnabled[s] = \{\varphi \in Prob[Act \times Sta] \; ; s \Rightarrow \varphi\}.$$

**Example 61.** Consider an STS $\mathcal{T} = (Sta, Act, \rightarrow)$ and a state $s \in Sta$. Let $\varphi_0, \varphi_1, \ldots \in Prob[Act \times Sta]$ be probability measures such that $Enabled[s] = \{\varphi_0, \varphi_1, \ldots\}$. Pick a distribution function $f \colon \mathbb{N} \rightarrow [0, 1]$, i.e., it holds $f(0) + f(1) + f(2) + \ldots = 1$. Intuitively, the function $f$ assigns a probability to every element in $Enabled[s]$. Define the probability measure $\varphi \in Prob[Act \times Sta]$ by

$$\varphi(B) = f(0) \cdot \varphi_0(B) + f(1) \cdot \varphi_1(B) + f(2) \cdot \varphi_2(B) + \ldots .$$

Thanks to Example 25, we obtain $s \Rightarrow \varphi$. As the set $Enabled[s]$ is countable, every combined transition enabled at $s$ is induced by a distribution function in the presented way. ⌟

The following theorem provides a generalisation of Proposition 4.2.1 in [Seg95]. Indeed, the mentioned result in [Seg95] only covers STSs where for every state $s$ the set $Enabled[s]$ of enabled distributions over action-state pairs is countable

**Theorem 62.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and $s \in Sta$ be such that the set $Enabled[s]$ is Souslin in $Act \times Sta$. Then we have the following identity*

$$Conv[Enabled[s]] = CombEnabled[s] = Conv[CombEnabled[s]].$$

*Proof.* According to Lemma 12, we have

$$CombEnabled[s] = Conv[Enabled[s]].$$

Thanks to Theorem 27, it also holds

$$Conv[Enabled[s]] = Conv[Conv[Enabled[s]]].$$

These two identities yield the claim. □

The previous result in particular shows that the set $CombEnabled[s]$ of distributions on $Act \times Sta$ is convex. The significance of this insight is illustrated below. Inspecting the definition of the combined-transition relation, for every $\varphi \in Prob[Act \times Sta]$ it holds $\varphi \in CombEnabled[s]$ provided $\varphi$ can be written as a barycentre of probability measures contained in $Enabled[s]$, i.e., $\varphi$ is a combination of measures enabled in $s$. Theorem 62 shows that $\varphi$ is also contained in $CombEnabled[s]$ in case where $\varphi$ can be represented as a barycentre of probability measures contained in $CombEnabled[s]$. In other words, the set $CombEnabled[s]$ is closed under combining probability measures.

## 3.6 Schedulers, path measures, and trace distributions

The intuitive behaviour of an STS $\mathcal{T} = (Sta, Act, \rightarrow)$ can be summarised as follows: assuming the current state of $\mathcal{T}$ is given by $s$, a probability measure $\varphi$ with $s \rightarrow \varphi$ is chosen non-deterministically in a first step. After that, an action-state pair $\langle act, s' \rangle$ is sampled according to the previously obtained $\varphi$. Intuitively, the action *act* is executed and the system enters the successor state $s'$. In other words, while the action *act* is taken, the internal system state changes from $s$ to $s'$. The sketched procedure comprising of a non-deterministic choice followed by a probabilistic choice is repeated continuously. In this way one obtains infinite sequences of action-state pairs called paths that is precisely formalised as follows. Reasoning about probabilities of sets of paths of a given STS requires the resolution of the non-determinism in terms of schedulers. In fact, every scheduler induces a path measure assigning probabilities to Borel sets of paths.

**Finite and infinite paths.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS. For every $n \in \mathbb{N}$ the set of all *finite paths* (*of length* $n$) is given by

$$Path_n = Sta \times (Act \times Sta)^n,$$

i.e., it holds $Path_0 = Sta$. The *empty path* is denoted by $\circ$. The union of the sets $\{\circ\}$ and $Path_n$, where $n$ ranges over $\mathbb{N}$, is denoted by $Path_{<\omega}$, i.e.,

$$Path_{<\omega} = \{\circ\} \cup \bigcup_{n \in \mathbb{N}} Path_n.$$

The set of all *infinite paths* is given by

$$Path = Sta \times (Act \times Sta)^{\omega}.$$

By Example 2, when equipped with the corresponding natural topologies, the set $Path_n$ constitutes a Polish space for every $n \in \mathbb{N}$ and the same holds for $Path_{<\omega}$ and $Path_{\omega}$. For every $n \in \mathbb{N}$ let $\mathrm{Last}_n \colon Path_n \to Sta$ be the function that maps every finite path of length $n$ to its last visited state, i.e., $\mathrm{Last}_n(\hat{\pi}) = s_n$ for every finite path $\hat{\pi} = s_0\, act_1\, s_1 \ldots act_n\, s_n$. Obviously, $\mathrm{Last}_n$ is a Borel function for every $n \in \mathbb{N}$.

**Schedulers and path measures.** Let $\mathcal{T} = (Sta, Act, \to)$ be an STS as before. We consider a function

$$\mathfrak{S} \colon Path_{<\omega} \to Prob[Sta] \cup Prob[Act \times Sta]$$

such that

$$\mathfrak{S}(\circ) \in Prob[Sta] \quad \text{and} \quad \mathfrak{S}(\hat{\pi}) \in Prob[Act \times Sta] \text{ for all } \hat{\pi} \in Path_{<\omega} \setminus \{\circ\}.$$

For every $n \in \mathbb{N}$ let $\mathfrak{S}_{|n} \colon Path_n \to Prob[Act \times Sta]$, $\mathfrak{S}_{|n}(\hat{\pi}) = \mathfrak{S}(\hat{\pi})$. The function $\mathfrak{S}$ is Borel iff for every $n \in \mathbb{N}$ the function $\mathfrak{S}_{|n}$ is Borel. Assuming the function $\mathfrak{S}$ is Borel, for every $n \in \mathbb{N}$ the probability measure $Pr_n[\mathfrak{S}]$ is inductively defined as follows:

$$Pr_0[\mathfrak{S}] = \mathfrak{S}(\circ) \quad \text{and} \quad Pr_{n+1}[\mathfrak{S}] = Pr_n[\mathfrak{S}] \rtimes \mathfrak{S}_{|n} \text{ for every } n \in \mathbb{N},$$

i.e., for every Borel sets $\hat{\Pi}_n \subseteq Path_n$, $A \subseteq Act$, and $S \subseteq Sta$ it holds

$$Pr_{n+1}[\mathfrak{S}](\hat{\Pi}_n \times A \times S) = \int_{\hat{\Pi}_n} \mathfrak{S}_{|n}(\hat{\pi})(A \times S)\, dPr_n[\mathfrak{S}](\hat{\pi}).$$

Intuitively, the probability measure $Pr_{n+1}[\mathfrak{S}]$ results from $Pr_n[\mathfrak{S}]$ by averaging over the respective outcomes of the function $\mathfrak{S}_{|n}$ with respect to $Pr_n[\mathfrak{S}]$. Relying on the Kolmogorov's measure extension theorem (see Corollary 7.7.2 in [Bog07]), the sequence of measures $Pr_0[\mathfrak{S}], Pr_1[\mathfrak{S}], \ldots$ uniquely determine a probability measure $Pr[\mathfrak{S}]$ on the set of all infinite paths $Path$ such that for all $n \in \mathbb{N}$ and Borel sets $\hat{\Pi}_n \subseteq Path_n$,

$$Pr[\mathfrak{S}](\hat{\Pi}_n \times (Act \times Sta)^{\omega}) = Pr_n[\mathfrak{S}](\hat{\Pi}_n).$$

We refer to $Pr[\mathfrak{S}]$ as *path measure* (*induced by* $\mathfrak{S}$). Accordingly, for every $n \in \mathbb{N}$ the probability measure $Pr_n[\mathfrak{S}]$ is called *finite-path measure* (*of length n induced by* $\mathfrak{S}$).

The function $\mathfrak{S}$ does not involve any information concerning the transition relation of the STS $\mathcal{T}$ so far. Denote the combined-transition relation of the STS $\mathcal{T}$ by $\Rightarrow$. For every probability measure $\mu \in Prob[Sta]$ we call $\mathfrak{S}$ a *$\mu$-scheduler* (*for* $\mathcal{T}$) provided the following two properties are fulfilled:

- $\mathfrak{S}$ is a Borel function.

- $\mathfrak{S}(\circ) = \mu$ and for every $n \in \mathbb{N}$ there exists a Borel set $\hat{\Pi}_n \subseteq Path_n$ with

$$Pr_n[\mathfrak{S}](\hat{\Pi}_n) = 1 \quad \text{and} \quad \mathrm{Last}_n(\hat{\pi}) \Rightarrow \mathfrak{S}(\hat{\pi}) \text{ for every } \hat{\pi} \in \hat{\Pi}_n.$$

Here, if $\mu = Dirac[s]$ for some state $s \in Sta$, we also refer to $\mathfrak{S}$ as an $s$-scheduler rather than a $Dirac[s]$-scheduler. Consequently, schedulers may randomise over enabled transitions in terms of combined transitions and may access an infinite amount of memory. The Borel requirement on a scheduler yields an important mathematical foundation concerning the definition of the respective path measure.

**Trace distributions.** Given a probability measure $\mu \in Prob[Sta]$, every $\mu$-scheduler induces a probability measure on $Act^\omega$ when projecting paths onto the consecutive sequence of taken actions. Formally, define the Borel function $Trace\colon Path \to Act^\omega$ as follows: for every infinite path $s_0\, act_1\, s_1\, act_2\, s_2\, \ldots \in Path$,

$$Trace(s_0\, act_1\, s_1\, act_2\, s_2\, \ldots) = act_1\, act_2\, \ldots\,.$$

It is easy to see that the function *Trace* is Borel. For every probability measure $\mu \in Prob[Sta]$ and every $\mu$-scheduler $\mathfrak{S}$ we can hence safely define the *trace distribution* (*induced by* $\mathfrak{S}$) as

$$PrTrace[\mathfrak{S}] = Trace_\sharp(Pr[\mathfrak{S}]),$$

i.e., for every Borel sets $A_1, A_2, \ldots \subseteq Act$ we have

$$PrTrace[\mathfrak{S}](A_1 \times A_2 \times \ldots) = Pr[\mathfrak{S}](Sta \times (A_1 \times Sta) \times (A_2 \times Sta) \times \ldots).$$

The *trace-distribution preorder* $\leq^{\mathrm{tr}} \subseteq Sta \times Sta$ is defined as follows: for every $s_a, s_b \in Sta$,

$$\begin{array}{lll} s_a \leq^{\mathrm{tr}} s_b & \text{iff} & \text{for every } s_a\text{-scheduler there is an } s_b\text{-scheduler} \\ & & \text{such that } PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b]. \end{array}$$

*Trace-distribution equivalence* is the relation $=^{\mathrm{tr}} \subseteq Sta \times Sta$ given by $=^{\mathrm{tr}} = \leq^{\mathrm{tr}} \cap (\leq^{\mathrm{tr}})^{-1}$, i.e., for every states $s_a, s_b \in Sta$ it holds

$$\begin{array}{lll} s_a =^{\mathrm{tr}} s_b & \text{iff} & \text{for every } s_a\text{-scheduler there is an } s_b\text{-scheduler} \\ & & \text{such that } PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b] \\ & & \text{and vice versa, for every } s_b\text{-scheduler there is an } s_a\text{-scheduler} \\ & & \text{such that } PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b]. \end{array}$$

It is easy to see that $\leq^{\mathrm{tr}}$ and $=^{\mathrm{tr}}$ form a preorder and an equivalence, respectively.

# 4 Simulations and trace distributions for Souslin systems

The main object of this chapter is to show that for Souslin STSs the Souslin-simulation preorder is a subset of trace-distribution preorder and the Souslin-bisimulation equivalence is finer that trace-distribution equivalence, i.e., we provide a proof of Theorem B (see Chapter 1). Our proof strategy of the introductory sketched theorem is basically the same as for the standard result for non-stochastic discrete systems [BK08]. The easy, however, crucial observation in the setting of [BK08] is that every simulation can be inductively extended to a relation between paths, i.e., given states $s_a$ and $s_b$ where $s_a$ is simulated by $s_b$ as well as a path $\pi_a$ starting from $s_a$, there exists a path $\pi_b$ starting from $s_b$ that is statewise related to $\pi_a$. The corresponding in our setting includes schedulers and their respective induced path measures:

**Lemma G.** *Let $s_a$ and $s_b$ be states of a Souslin STS and $R$ be a Souslin simulation with $\langle s_a, s_b\rangle \in R$. For every $s_a$-scheduler $\mathfrak{S}_a$ there exists an $s_b$-scheduler $\mathfrak{S}_b$ such that for all $n \in \mathbb{N}$ the finite-path measures $Pr_n[\mathfrak{S}_a]$ and $Pr_n[\mathfrak{S}_b]$ are related concerning the relation $R$, i.e.,*

$$\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b]\rangle \in (R^{\text{path},n})^{\text{wgt}}.$$

⌟

A detailed proof of the sketched key lemma can be found in Section 4.5. In what follows we elaborate on the mathematical challenge for the lemma under consideration. To obtain a proper mathematical framework for defining path measures and trace distributions, every scheduler is required to form a Borel function (see Section 3.6 and also [CSKN05, Cat05] as well as [WJ06, Wol12]). In fact, it is not straightforward how to use the measurability properties of the $s_a$-scheduler $\mathfrak{S}_a$ in a proof of the previous lemma for the construction of an $s_b$-scheduler $\mathfrak{S}_b$. The intuitive reason is that the notion of simulation involves no measurability conditions concerning enabled transitions and hence, one cannot expect that measurability properties of $\mathfrak{S}_a$ are preserved: in Section 4.6 we even see an STS involving states $s_a$ and $s_b$ such that $s_b$ simulates $s_a$, however, while there exists an $s_a$-scheduler there exists no $s_b$-scheduler at all.

Indeed, the latter discussion motivates our definition of Souslin STSs as here, assuming the STS is additionally non-blocking, there exists an $s$-scheduler for every state $s$. The high-level proof idea for this property of Souslin STSs is basically the same as for Lemma G: we introduce an appropriate set-valued function intuitively representing a set of candidates for a suitable scheduler. Relying on the given Souslin assumptions, we then show the applicability of the measurable-selection principle given by Theorem 21 that finally entails a convenient scheduler from the candidates. We remark that this approach does not yield an explicit construction for schedulers and only shows the existence of a scheduler with specific properties.

As we see in Section 4.7, it turns out that Theorem B yields non-trivial generalisation of Theorem 9.19 in [Cat05]. Moreover, in combination with the logical characterisation of the simulation preorder and the bisimulation equivalence to be investigated in the next chapter, Theorem B has various consequences that are summarised in Chapter 7. For instance, we see that labelled Markov processes [BDEP97, Des99, DEP02, Pan09] as well as image-finite non-deterministic labelled Markov processes [DTW12, Wol12] with countable action spaces and continuous controlled Markov processes [DIY79, BS96, ZEM$^+$14, TMKA16] are Souslin STSs for which the Souslin-simulation preorder and the simulation preorder are the same and accordingly, where the Souslin-bisimulation equivalence and the bisimulation equivalence collapse. Consequently, we also obtain a trace-distribution result for these prominent subclasses of STSs.

Besides the latter mentioned applications of Theorem B concerning uncountable-state-space models, the result also applies for STSs whose state space is countable. This includes probabilistic automata [Seg95], discrete Markov decision processes [Put94], continuous-time Markov chains [BHHK03, DP03], continuous-time Markov decision processes [NK07], interactive Markov chains [Her02], Markov automata [EHZ10], and stochastic automata [D'A99, BD04, DK05]. Continuous-time Markov decision processes serve, e.g., as a semantical model for stochastic Petri nets [CMBC93].

## 4.1 Souslin stochastic transition systems

Our basic stochastic model of STSs also permits non-deterministic choices over enabled distributions in a state. To reason about probabilities of Borel sets of paths, the non-determinism needs to be resolved by schedulers. Recall, schedulers are Borel functions assigning every state to an enabled distribution over action-state pairs. However, since the notion of STSs is very generic in the sense that there are no (measurability) restrictions on the transition relation, there are examples of STSs that admit no schedulers even if the STS under consideration is non-blocking. In other words, it is possible to given an STS such that there exists no function that is both Borel and compatible with the underlying transition relation at the same time. The latter fact has been already observed in [Cat05] We recall the corresponding example first:

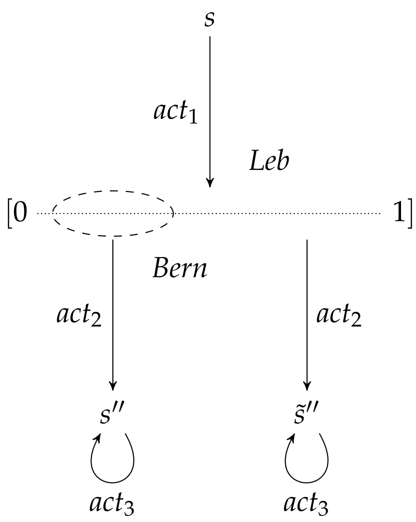


Figure 4.1: There exists no $s$-scheduler.

**Example 63** (Example 9.2 in [Cat05]). In what follows the Lebesgue measure on $[0,1]$ is denoted by *Leb*. The Lebesgue measure of an interval $[r_1, r_2]$ contained in $[0,1]$ is its length given by $r_2 - r_1$. Let *Bern* be a Bernstein set, i.e., *Bern* is a subset of $[0,1]$ such that for every Borel set $B \subseteq [0,1]$ the following property holds (see Theorem 5.4 in [Oxt71]):

$$B \subseteq \mathit{Bern} \text{ or } B \subseteq [0,1] \setminus \mathit{Bern} \quad \text{implies} \quad \mathit{Leb}(B) = 0.$$

In fact, the stated condition for the set *Bern* is a consequence of the actual definition of a Bernstein set given in [Oxt71]. However, the presented requirement is sufficient for the following purposes. Let $\mathcal{T} = (\mathit{Sta}, \mathit{Act}, \rightarrow)$ be the STS illustrated by Figure 4.1 (see also Figure 9.2 in [Cat05]). To be more precise, it holds

$$\mathit{Sta} = \{s\} \cup [0,1] \cup \{s'', \tilde{s}''\} \quad \text{and} \quad \mathit{Act} = \{act_1, act_2, act_3\}.$$

Moreover, for every $r \in [0,1]$ the following two equivalences are satisfied:

$$\begin{aligned} r &\to \langle act_2, Dirac[s''] \rangle \quad \text{iff} \quad r \in Bern, \\ r &\to \langle act_2, Dirac[\tilde{s}''] \rangle \quad \text{iff} \quad r \notin Bern. \end{aligned}$$

The remaining transitions of $\mathcal{T}$ are given as follows:

$$s \to \langle act_1, Leb \rangle, \quad s_a \to \langle act_3, Dirac[s''] \rangle, \quad s_b \to \langle act_3, Dirac[\tilde{s}''] \rangle.$$

Consequently, the STS $\mathcal{T}$ is non-blocking. However, there does not exist an $s$-scheduler. In what follows we present a detailed formal argument. Towards a contradiction assume that there is an $s$-scheduler, say $\mathfrak{S}$. Define the two Borel sets $\hat{\Pi}_a, \hat{\Pi}_b \subseteq Path_2$ by

$$\begin{aligned} \hat{\Pi}_a &= \{s\} \times \{act_1\} \times [0,1] \times \{act_2\} \times \{s''\}, \\ \hat{\Pi}_b &= \{s\} \times \{act_1\} \times [0,1] \times \{act_2\} \times \{\tilde{s}''\}. \end{aligned}$$

As the sets $\hat{\Pi}_a$ and $\hat{\Pi}_b$ form a partition of $Path_2$, we obtain $Pr_2[\mathfrak{S}](\hat{\Pi}_a) + Pr_2[\mathfrak{S}](\hat{\Pi}_b) = 1$. It follows $Pr_2[\mathfrak{S}](\hat{\Pi}_a) > 0$ or $Pr_2[\mathfrak{S}](\hat{\Pi}_b) > 0$.

Consider the case $Pr_2[\mathfrak{S}](\hat{\Pi}_a) > 0$. Then, as $\mathfrak{S}$ is an $s$-scheduler, there is a Borel set $B \subseteq [0,1]$ with the following three properties: $B \subseteq Bern$, $Pr_1[\mathfrak{S}](\{s\} \times \{act_1\} \times B) = 1$, and $r \to \mathfrak{S}(s\, act_1\, r)$ for every $r \in B$. Inspecting the STS $\mathcal{T}$, for every $r \in B$ it holds $\mathfrak{S}(s\, act_1\, r) = Dirac[\langle act_2, s'' \rangle]$. From this we conclude

$$0 < Pr_2[\mathfrak{S}](\hat{\Pi}_a) = Pr_2[\mathfrak{S}](\{\langle s, act_1, r, act_2, s_a \rangle \; ; r \in B\}) = Leb(B).$$

This yields a contradiction as we also have $Leb(B) = 0$ relying on the properties of a Bernstein set. One analogously derives a contradiction from the remaining case $Pr_2[\mathfrak{S}](\hat{\Pi}_b) > 0$. We conclude that there does not exist an $s$-scheduler. ⌟

Considering the STS $\mathcal{T}$ in Example 63 again, every path starting at $s$ admits exactly the same trace $act_1\, act_2\, act_3^\omega$. It is hence natural to expect that there exists an uniquely determined trace distribution at $s$ given by the probability measure $Dirac[act_1\, act_2\, act_3^\omega]$. However, Example 63 shows that there is no measurable resolution of the non-determinism and hence, the STS $\mathcal{T}$ with $s$ being the initial state admits no observable behaviour in terms of trace distributions. As a consequence of this discussion, the STS $\mathcal{T}$ may be seen as artificial and may be judged as a flawed model. We present a (syntactical) requirement for STSs that ensures the existence of schedulers to rule out such pathological models.

**Definition 64.** Let $\mathcal{T} = (Sta, Act, \to)$ be an STS. We say that $\mathcal{T}$ is *Souslin* if the set $\to$ of all transitions is Souslin in $Sta \times Prob[Act \times Sta]$. ⌟

The underlying motivation behind the notion of a Souslin STS is to enable the applicability of measurable-selection principles (see Section 2.3) for showing the existence of schedulers. To be more precise, the transition relation of every STS $\mathcal{T} = (Sta, Act, \to)$ can be naturally rephrased in terms of the set-valued function $F\colon Sta \rightsquigarrow Prob[Act \times Sta]$,

$$F(s) = Enabled[s].$$

Assuming an initial distribution $\mu \in Prob[Sta]$ on the states, every choice of a $\mu$-scheduler $\mathfrak{S}$ in the first step corresponds the Borel $\mu$-selection of $F$, i.e., $\mathfrak{S}_{|0}$ is a Borel $\mu$-selection of $F$. Indeed, provided $f$ is a Borel $\mu$-selection of $F$, there exists a Borel set $B \subseteq Sta$ with the following two properties:

$$\mu(B) = 1 \quad \text{and} \quad s \to f(s) \text{ for every } s \in B,$$

i.e., $f$ is compatible with the transition relation. To guarantee the existence of a Borel $\mu$-selection of $F$, we rely on the measurable-selection principle stated in Theorem 21 and thus, it suffices to assume that the set $Rel[F]$ is Souslin in $Sta \times Prob[Act \times Sta]$. However, the latter holds precisely when the STS $\mathcal{T}$ is Souslin. The idea of using measurable-selection principles in the context of Souslin STSs can be extended to a step-by-step construction of schedulers that is demonstrated by the proof of the following theorem:

**Theorem 65.** *Let $\mathcal{T} = (Sta, Act, \to)$ be a non-blocking Souslin STS. Then for every probability measure $\mu \in Prob[Sta]$ there exists a $\mu$-scheduler. In particular, for every state $s \in Sta$ and $\varphi \in Prob[Act \times Sta]$ such that $s \to \varphi$ there exists an $s$-scheduler $\mathfrak{S}$ with $\mathfrak{S}(s) = \varphi$.*

*Proof.* The theorem is an application of a measurable-selection principle. More precisely, we derive it from Theorem 21. For every $n \in \mathbb{N}$ we introduce the set-valued function $F_n\colon Path_n \rightsquigarrow Prob[Act \times Sta]$,

$$F_n(\hat{\pi}) = \{\varphi \in Prob[Act \times Sta] \, ; \, \mathrm{Last}_n(\hat{\pi}) \to \varphi\}.$$

For every $n \in \mathbb{N}$, using that the STS $\mathcal{T}$ is Souslin and as the function $\mathrm{Last}_n$ is Borel, the set $Rel[F_n]$ is Souslin in $Path_n \times Prob[Act \times Sta]$. Moreover, using that the STS under consideration is not blocking, it also follows that the set $F_n(\hat{\pi})$ is not empty for every finite path $\hat{\pi} \in Path_n$. Thus, we are in the setting of Theorem 21. Therefore, for every $n \in \mathbb{N}$ and $\hat{\chi}_n \in Prob[Path_n]$ there exists a Borel $\hat{\chi}_n$-selection of $F_n$.

Let $\mu \in Prob[Sta]$ be a probability measure. Relying on an easy inductive argument, there exist a sequence $(\hat{\chi}_n)_{n \in \mathbb{N}}$ of probability measures $\hat{\chi}_n \in Prob[Path_n]$ and a sequence $(\mathfrak{S}_n)_{n \in \mathbb{N}}$ of Borel functions $\mathfrak{S}_n\colon Path_n \to Prob[Act \times Sta]$ with the following properties:

$$\hat{\chi}_0 = \mu \quad \text{and} \quad \hat{\chi}_{n+1} = \hat{\chi}_n \rtimes \mathfrak{S}_n \text{ for every } n \in \mathbb{N}$$

and

$\mathfrak{S}_n$ is Borel $\hat{\chi}_n$-selection of $F_n$ for every $n \in \mathbb{N}$.

Let $\mathfrak{S}\colon Path_{<\omega} \to Prob[Sta] \cup Prob[Act \times Sta]$ be the function such that $\mathfrak{S}(\circ) = \mu$ and so that for every $n \in \mathbb{N}$ and $\hat{\pi} \in Path_n$ it holds $\mathfrak{S}(\hat{\pi}_n) = \mathfrak{S}_n(\hat{\pi}_n)$. It easily follows that $\mathfrak{S}$ is a Borel function and moreover, for every $n \in \mathbb{N}$ one has $Pr_n[\mathfrak{S}] = \hat{\chi}_n$. From this one can easily derive that $\mathfrak{S}$ is a $\mu$-scheduler.

Let $s \in Sta$ and $\varphi \in Prob[Act \times Sta]$ be such that $s \to \varphi$. The remaining claim of the theorem follows exactly in the same way when replacing the set-valued function $F_0$ by the function $F_0'\colon Sta \to Prob[Act \times Sta]$, $F_0'(s') = \{\varphi\}$. □

With the insight of Theorem 65, the requirement that the transition relation constitutes a Souslin set rules out pathological examples of STSs as depicted in Figure 4.1. The argumentation scheme for Theorem 65 occurs at various points throughout this thesis: in order to obtain convenient Borel functions, the basic approach is to introduce an appropriate set-valued function that fulfills the requirements of the measurable-selection principle stated in Theorem 21.

**Example 66.** We continue our discussions concerning Example 63. The STS $\mathcal{T}$ depicted in Figure 4.1 is not Souslin by Theorem 65. Let us also provide a simple direct argument for that observation without using the section's main theorem. Towards a contradiction assume that $\mathcal{T}$ is Souslin. Since

$$\to \cap\, ([0,1] \times \{Dirac[\langle act_2, s_a\rangle]\}) = Bern \times \{Dirac[\langle act_2, s_a\rangle]\},$$

the set $Bern \times \{Dirac[\langle act_2, s_a\rangle]\}$ is Souslin in $Sta \times Prob[Act \times Sta]$. By Remark 10 (1), the set $Bern$ is therefore Souslin in $[0,1]$. However, it is well-known that Bernstein sets are not Souslin [Oxt71]. This can be seen as follows. Towards a contradiction assume that $Bern$ is Souslin in $[0,1]$. Thanks to Remark 11, there are hence Borel sets $B_l, B_u \subseteq [0,1]$ with

$$B_l \subseteq Bern \subseteq B_u \quad \text{and} \quad Leb(B_l) = Leb(B_u).$$

Relying on the properties of a Bernstein set, it follows $Leb(B_l) = 0$ and hence, $Leb(B_u) = 0$. We therefore obtain $Leb([0,1] \setminus B_u) = 1$. However, as we have $[0,1] \setminus B_u \subseteq [0,1] \setminus Bern$, it also holds $Leb([0,1] \setminus B_u) = 0$ that yields a contradiction. We conclude that the set $Bern$ is not Souslin in $[0,1]$ and thus, we finally derive that the STS $\mathcal{T}$ is not Souslin. ⌟

This section is finished with the following two remarks yielding basic observations for Souslin STSs.

**Remark 67.** For the following important property of Souslin STSs it is appropriate to recall the notion of barycentres and convex hulls in the context probability theory from Section 2.4 (see also Section 3.5). Let $\mathcal{T} = (Sta, Act, \to)$ be an Souslin STS. Then for every state $s \in Sta$ the set $Enabled[s]$ is Souslin in $Prob[Act \times Sta]$, in particular, it holds

$$Conv[Enabled[s]] = CombEnabled[s] = Conv[CombEnabled[s]].$$

The argument for that claim is easy. Consider a state $s \in Sta$. It holds

$$\to \cap \left(\{s\} \times Prob[Act \times Sta]\right) = \{s\} \times Enabled[s].$$

As the STS $\mathcal{T}$ is Souslin, the set $\{s\} \times Enabled[s]$ is Souslin in $Sta \times Prob[Act \times Sta]$. Consequently, by Remark 10 (1), the set $Enabled[s]$ is Souslin in $Prob[Act \times Sta]$. Theorem 62 directly yields the remaining claim now. ⌟

**Remark 68.** Let $\mathcal{T} = (Sta, Act, \to)$ be a reactive STSs. Abbreviate

$$X = Sta \times Act \times Prob[Sta].$$

Then the following two statements are equivalent:

(1) $\mathcal{T}$ is Souslin.

(2) The set $\{\langle s, act, \mu\rangle \in X \,;\, s \to \langle act, \mu\rangle\}$ is Souslin in $X$.

The argument is easy: defining the function $\xi\colon X \to Sta \times Prob[Act \times Sta]$,

$$\xi(s, act, \mu) = \langle s, Dirac[act] \otimes \mu\rangle,$$

we obtain

$$\to = \xi(\{\langle s, act, \mu\rangle \in X \,;\, s \to \langle act, \mu\rangle\}).$$

By Example 6, the function $\xi$ is Borel and hence, the claimed equivalence is a direct consequence of Remark 10 (5). ⌟

## 4.2 Weakly Souslin stochastic transition systems

Theorem 65 states that every Souslin STS admits a scheduler. In fact, inspecting the proof again, we provide a deterministic scheduler, i.e., a scheduler that does not randomise over different enabled distributions in a given state. A priori this insight is not trivial because of the measurability requirements for schedulers and can be proven by a standard application of a measurable-selection principle. However, later on in this chapter, we are confronted with the problem of constructing schedulers that are not necessarily deterministic, i.e., schedulers that may be not compatible with the transition relation but with the induced combined-transition relation. Recalling Section 3.5 and also Section 2.4, the combined-transition relation results from the transition relation by combining the enabled distributions in a state using appropriate barycentres of probability measures. To rely again on a measurable-selection principle as in the proof of Theorem 65, we need the fact that the combined-transition relation is Souslin. This motivates the following definition:

**Definition 69.** We call an STS $\mathcal{T} = (Sta, Act, \rightarrow)$ *weakly Souslin* if the set $\Rightarrow$ is Souslin in $Sta \times Prob[Act \times Sta]$ where $\Rightarrow$ denotes the combined-transition relation of $\mathcal{T}$. ⌟

The previous definition is exact in the same spirit of Definition 64 for Souslin STSs. As the notions suggests, every Souslin STS is also weakly Souslin. However, the proof for this insight is not trivial and yields the main contribution of this section:

**Theorem 70.** *Every Souslin STS is weakly Souslin.* ⌟

Let us give the key idea for our argument before we present the precise technical details. For this purpose pick a Souslin STS $\mathcal{T} = (Sta, Act, \rightarrow)$. The basic observation is that every transition $s \rightarrow \varphi$ can be identified with the probability measure $f(s, \varphi) = Dirac[s] \otimes \varphi$ in $Prob[Sta \times (Act \times Sta)]$. As a consequence of this identification, the transition relation $\rightarrow$ can be viewed as a subset $f(\rightarrow)$ of $Prob[Sta \times (Act \times Sta)]$. The advantage is now that one can regard the convex hull $Conv[f(\rightarrow)]$ of $f(\rightarrow)$. Relying on Theorem 27, it follows that the set $Conv[f(\rightarrow)]$ is Souslin in in $Prob[Sta \times (Act \times Sta)]$. In a last proof step, roughly speaking, we see that every probability measure in $Conv[f(\rightarrow)]$ is of the form $Dirac[s] \otimes \varphi$ for some state $s \in Sta$ and probability measure $\varphi \in Prob[Act \times Sta]$ such that $s \Rightarrow \varphi$. Intuitively, the combined-transition relation can be extracted from the set $Conv[f(\rightarrow)]$. From this it finally follows that the STS $\mathcal{T}$ is weakly Souslin.

One might ask whether the reverse implication of Theorem 70 holds, i.e., whether every weakly Souslin STS is also Souslin. We believe that this is not the case but have no counterexample for our feeling. However, the latter question is neither relevant for practical purposes nor for the following material. Indeed, throughout this thesis, we solely need Theorem 70 within theoretical arguments. To show that a concretely given STS is indeed Souslin, e.g., an STS resulting from the unfolding of a modelling formalism from the literature, we directly focus on the given transition relation rather than the more complex induced combined-transition relation (see Chapter 7).

**Proof of the section's main result.** It is appropriate to recall the precise definition of barycentres and convex hulls in the context of probability measures from Section 2.4. The following sufficient criterion for showing that a probability measure constitutes a Dirac measure is routine and follows from basic results from measure theory.

**Lemma 71.** *Let $X$ be a measurable space and $\mathcal{G} \subseteq 2^X$ be a generator of the sigma algebra on $X$ that is closed under finite intersections and complements. Suppose $\mu \in Prob[X]$ and $x \in X$. If $\mu(B) = 0$ for all $B \in \mathcal{G}$ with $x \notin B$, then we have $\mu = Dirac[x]$.*

*Proof.* Applying Carathéodory extension theorem (see Section 2.1), we have $\mu = Dirac[x]$ iff $\mu(B) = Dirac[x](B)$ for all $B \in \mathcal{G}$. Assume $\mu \neq Dirac[x]$. Thus, there exists $B \in \mathcal{G}$ such that $\mu(B) \neq Dirac[x](B)$. In case where $Dirac[x](B) = 0$, it follows $x \notin B$ and $\mu(B) \neq 0$. In the case where $Dirac[x](B) = 1$, we have $X \setminus B \in \mathcal{G}$, $x \notin X \setminus B$, and $\mu(X \setminus B) \neq 0$. We are done putting things together. $\square$

*Proof of Theorem 70.* Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS and denote the corresponding combined-transition relation by $\Rightarrow$. Introduce the Polish space

$$X = Sta \times Act \times Sta$$

and define the function $f\colon Sta \times Prob[Act \times Sta] \rightarrow Prob[X]$,

$$f(s, \varphi) = Dirac[s] \otimes \varphi.$$

The function $f$ is Borel by Example 6. For every $s \in Sta$ let $f_s\colon Prob[Act \times Sta] \rightarrow Prob[X]$,

$$f_s(\varphi) = f(s, \varphi).$$

For every $s \in Sta$ it is easy to see that $f_s$ is a Borel function.

By Remark 10 (5) and since the STS $\mathcal{T}$ is Souslin, the set $f(\rightarrow)$ is Souslin in $Prob[X]$. According to Theorem 27, the set $Conv[f(\rightarrow)]$ is Souslin in $Prob[X]$. Invoking again Remark 10 (5), the set $f^{-1}(Conv[f(\rightarrow)])$ is Souslin in $Sta \times Prob[Act \times Sta]$. To conclude the theorem, it hence suffices to show the identity

$$\Rightarrow\, = f^{-1}(Conv[f(\rightarrow)]).$$

To this end let $s \in Sta$ and $\varphi \in Prob[Act \times Sta]$.

*Inclusion from left to right.* Assume $s \Rightarrow \varphi$ first. Our task is to justify $f(s, \varphi) \in Conv[f(\rightarrow)]$, i.e., $Dirac[s] \otimes \varphi = Barycen(\tilde{\beta})$ for some $\tilde{\beta} \in Prob[Prob[X]]$ with $\tilde{\beta}^{\mathrm{out}}(f(\rightarrow)) = 1$. Recall, the set $Enabled[s]$ is Souslin in $Prob[Act \times Sta]$ by Remark 67. Since we have $s \Rightarrow \varphi$, Lemma 12 thus yields the existence of some $\beta \in Prob[Prob[Act \times Sta]]$ satisfying

$$\varphi = Barycen(\beta) \quad \text{and} \quad \beta^{\text{out}}(Enabled[s]) = 1.$$

Define $\tilde{\beta} \in Prob[Prob[X]]$ by

$$\tilde{\beta} = (f_s)_\sharp(\beta).$$

As $f_s(Enabled[s]) \subseteq f(\rightarrow)$, Lemma 13 entails $\tilde{\beta}^{\text{out}}(f(\rightarrow)) = 1$. To conclude $f(s, \varphi) \in Conv[f(\rightarrow)]$, it hence suffices to show the identity $Dirac[s] \otimes \varphi = Barycen(\tilde{\beta})$. Let $S \subseteq Sta$ and $B \subseteq Act \times Sta$ be Borel sets. It holds

$$Barycen(\tilde{\beta})(S \times B) = \int f_s(\varphi')(S \times B)\, d\beta(\varphi') = \int Dirac[s](S) \cdot \varphi'(B)\, d\beta(\varphi').$$

Using the identity

$$\int Dirac[s](S) \cdot \varphi'(B)\, d\beta(\varphi') = Dirac[s](S) \cdot \int \varphi'(B)\, d\beta(\varphi'),$$

we hence obtain

$$Barycen(\tilde{\beta})(S \times B) = Dirac[s](S) \cdot Barycen(\beta)(B) = (Dirac[s] \otimes \varphi)(S \times B).$$

It follows $Barycen(\tilde{\beta}) = Dirac[s] \otimes \varphi$ applying Carathéodory extension theorem (see Section 2.1). This justifies the first inclusion $\Rightarrow \subseteq f^{-1}(Conv[f(\rightarrow)])$.

*Inclusion from right to left.* Assume $f(s, \varphi) \in Conv[f(\rightarrow)]$. Then there exists a probability measure $\tilde{\beta} \in Prob[Prob[X]]$ such that

$$\tilde{\beta}^{\text{out}}(f(\rightarrow)) = 1 \quad \text{and} \quad Dirac[s] \otimes \varphi = Barycen(\tilde{\beta}).$$

Our task is to show $s \Rightarrow \varphi$. For this purpose define the Borel function $\zeta \colon X \to Act \times Sta$,

$$\zeta(s', act, s'') = \langle act, s'' \rangle.$$

Moreover, define $\beta \in Prob[Prob[Act \times Sta]]$ by

$$\beta = (\zeta_\sharp)_\sharp(\tilde{\beta}).$$

Thanks to Lemma 12, it remains to show $Barycen(\beta) = \varphi$ and $\beta^{\text{out}}(Enabled[s]) = 1$.

The first claim $Barycen(\beta) = \varphi$ can be seen as follows. For every Borel set $B \subseteq Act \times Sta$ it holds

$$Barycen(\beta)(B) = \int \gamma(Sta \times B)\, d\tilde{\beta}(\gamma) = Barycen(\tilde{\beta})(Sta \times B) = \varphi(B).$$

It remains to show $\beta^{\text{out}}(Enabled[s]) = 1$. We prove $\tilde{\beta}^{\text{out}}(f_s(Prob[Act \times Sta])) = 1$ first. Let $\mathcal{G} \subseteq 2^{Sta}$ be a countable generator of the Borel sigma-algebra on $Sta$ that is closed under finite intersections and complements (see Remark 4). For every $B \in \mathcal{G}$ with $s \notin B$ we have

$$0 = Dirac[s] \otimes \varphi(B \times Prob[Act \times Sta])$$

and therefore,

$$0 = Barycen(\tilde{\beta})(B \times Prob[Act \times Sta]) = \int \gamma(B \times Prob[Act \times Sta])\, d\tilde{\beta}(\gamma).$$

Thus, relying on a standard result from measure theory (see, e.g., Lemma 8.2.8 in [Sch08]), for every $B \in \mathcal{G}$ with $s \notin B$ there exists a Borel set $P_B \subseteq Prob[X]$ such that

$$\tilde{\beta}(P_B) = 1 \quad \text{and} \quad \gamma(B \times Prob[Act \times Sta]) = 0 \text{ for all } \gamma \in P_B.$$

Define $P \subseteq Prob[X]$ by

$$P = \bigcap_{B \in \mathcal{G} \text{ with } s \notin B} P_B.$$

Since $\mathcal{G}$ is countable, the set $P$ is Borel in $Prob[X]$ and moreover, it holds $\tilde{\beta}(P) = 1$. According to Lemma 71, we have $\gamma(B \times Prob[Act \times Sta]) = Dirac[s](B)$ for all Borel sets $B \subseteq Sta$ and $\gamma \in P$. Thus, for every $\gamma \in P$ there exists $\varphi' \in Prob[Act \times Sta]$ such that $\gamma = Dirac[s] \otimes \varphi'$. It follows $P \subseteq f_s(Prob[Act \times Sta])$ and therefore, we finally obtain the identity $\tilde{\beta}^{\text{out}}(f_s(Prob[Act \times Sta])) = 1$.

Let $P \subseteq Prob[Act \times Sta]$ be a Borel set so that $Enabled[s] \subseteq P$. If we show $\beta(P) = 1$, then it follows $\beta^{\text{out}}(Enabled[s]) = 1$ and we are done. As $Enabled[s] \subseteq P$, we obtain

$$f(\rightarrow) \cap f_s(Prob[Act \times Sta]) \subseteq \zeta_\sharp^{-1}(P).$$

As the outer-measure function is monotonically increasing, it follows

$$\beta(P) = \tilde{\beta}(\zeta_\sharp^{-1}(P)) = \tilde{\beta}^{\text{out}}(\zeta_\sharp^{-1}(P)) \geq \tilde{\beta}^{\text{out}}(f(\rightarrow) \cap f_s(Prob[Act \times Sta])) = 1.$$

The latter equal sign follows from $\tilde{\beta}^{\text{out}}(f(\rightarrow)) = 1$ and $\tilde{\beta}^{\text{out}}(f_s(Prob[Act \times Sta])) = 1$. $\square$

## 4.3 Souslin simulation and bisimulation

A (bi)simulation is a relation over the state space of an STSs ensuring the equivalence of the branching-time structure of related states. There are no explicit measurability conditions for the relation under consideration so far, in particular, it is not required that a (bi)simulation forms a Borel or Souslin set. The following material introduces simulations and bisimulations that constitute Souslin sets in addition. Based on these new notions we develop a theory on equivalences and abstraction of Souslin STSs throughout the subsequent sections.

**Definition 72.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS. A simulation $R$ is called *Souslin* if the set $R$ is Souslin in $Sta \times Sta$. Accordingly, we refer to a bisimulation $R$ as *Souslin* provided the set $R$ is Souslin in $Sta \times Sta$. ⌟

As for the original definitions of simulation and bisimulation, the introduced notions in Definition 72 also induce relations over the state space of an STS $\mathcal{T}$. Denote the state space of $\mathcal{T}$ by $Sta$. The *Souslin-simulation preorder* and the *Souslin-bisimulation equivalence* are the binary relations $\preceq^{\mathrm{sou}}$ and $\simeq^{\mathrm{sou}}$ on $Sta$, respectively, given as follows: for every states $s_a, s_b \in Sta$,

$$
\begin{aligned}
s_a \preceq^{\mathrm{sou}} s_b \quad &\text{iff} \quad \text{there is a Souslin simulation } R \text{ with } \langle s_a, s_b\rangle \in R\\
s_a \simeq^{\mathrm{sou}} s_b \quad &\text{iff} \quad \text{there is a Souslin bisimulation } R \text{ with } \langle s_a, s_b\rangle \in R.
\end{aligned}
$$

We observe that the relation $\preceq^{\mathrm{sou}}$ is indeed a preorder on $Sta$ and that the relation $\simeq^{\mathrm{sou}}$ is an equivalence on $Sta$. The proof for this insight is analogous to our argument for Theorem 55. Indeed, Lemma 56 and Theorem 58 (1) together yield the claim.

For every states $s_a, s_b \in Sta$ we obviously have that $s_a \preceq^{\mathrm{sou}} s_b$ implies $s_a \preceq s_b$ as well as $s_a \simeq^{\mathrm{sou}} s_b$ implies $s_a \simeq s_b$. When studying different applications of the mathematical theory of this chapter (see, e.g., Chapters 5 and 7), we see that for many subclasses of STSs (from the literature) the relations $\preceq^{\mathrm{sou}}$ and $\preceq$ as well as $\simeq^{\mathrm{sou}}$ and $\simeq$ are the same.

## 4.4 Combined simulation and bisimulation

Whereas the requirements for a (Souslin) simulation and bisimulation solely focus on the transition relation, the notions of schedulers, path measures, and trace distributions also include the combined-transition relation, which properly extends the transition relation (see also Sections 3.5 and 3.6). The material of this section closes a gap that arises in our investigations throughout the next sections where connections between the (bi)simulation relation and the trace-distribution relation are investigated. To be more precise, the following question raises for states $s_a$ and $s_b$ of an STS where $s_b$ simulates $s_a$: Is it the case that every combined transition in $s_a$ can be matched by a combined transition in $s_b$? In other words, we ask whether a simulation $R$ for an STS $\mathcal{T}$ also yields a simulation for the STS $\mathcal{T}'$ where $\mathcal{T}'$ is defined as $\mathcal{T}$ expect for the transition relation that is given by the combined-transition relation of $\mathcal{T}$ instead. The main achievement of this section is that the latter question can be answered positively when considering Souslin STSs and Souslin simulation and bisimulation.

**Definition 73.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and denote its combined-transition relation by $\Rightarrow$. A relation $R \subseteq Sta \times Sta$ is called a *combined simulation* if for every states $s_a, s_b \in Sta$ and $\varphi_a \in Prob[Act \times Sta]$ such that

$$\langle s_a, s_b \rangle \in R \quad \text{and} \quad s_a \Rightarrow \varphi_a,$$

there exists $\varphi_b \in Prob[Act \times Sta]$ with

$$s_b \Rightarrow \varphi_b \quad \text{and} \quad \langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}.$$

A relation $R \subseteq Sta \times Sta$ is said to be a *combined bisimulation* if both relations $R$ and its inverse $R^{-1}$ are combined simulations. ⌟

$$\begin{array}{ccccccc} s_a & R & s_b & & s_a & R & s_b \\ \Downarrow & & & \text{can be completed to} & \Downarrow & & \Downarrow \\ \varphi_a & & & & \varphi_a & (R^{Act})^{\mathrm{wgt}} & \varphi_b \end{array}$$

Figure 4.2: Condition for a combined simulation $R$.

Figure 4.2 illustrates the condition for a combined simulation. The definition of a combined (bi)simulation is completely analogous to Definition 47, however, focusing on the combined-transition relation rather than the transition relation. Many concepts related to simulations and bisimulations can be thus easily adapted for the newly introduced notion. For instance, a combined simulation $R$ is called *Souslin* provided the set $R$ is Souslin in $Sta \times Sta$ and analogously for combined bisimulations. The *Souslin-combined-simulation preorder* and the *Souslin-combined-bisimulation equivalence* are the relations $\preceq^{\mathrm{sou,c}}$ and $\simeq^{\mathrm{sou,c}}$ over the state space *Sta*, respectively, defined as follows: for every states $s_a, s_b \in Sta$,

$$\begin{aligned} s_a \preceq^{\mathrm{sou,c}} s_b \quad &\text{iff} \quad \text{there is a Souslin combined simulation } R \text{ with } \langle s_a, s_b \rangle \in R, \\ s_a \simeq^{\mathrm{sou,c}} s_b \quad &\text{iff} \quad \text{there is a Souslin combined bisimulation } R \text{ with } \langle s_a, s_b \rangle \in R. \end{aligned}$$

An easy adaption of our proof for Theorem 55 shows that both relations $\preceq^{\mathrm{c}}$ and $\preceq^{\mathrm{sou,c}}$ are preorders on *Sta* and that the relations $\simeq^{\mathrm{c}}$ and $\simeq^{\mathrm{sou,c}}$ form equivalences on *Sta*. The main result of this section is as follows:

**Theorem 74.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS. For every states $s_a, s_b \in Sta$ the following two statements hold:*

(1) $s_a \preceq^{\text{sou}} s_b$ *implies* $s_a \preceq^{\text{sou,c}} s_b$.

(2) $s_a \simeq^{\text{sou}} s_b$ *implies* $s_a \simeq^{\text{sou,c}} s_b$.

*Moreover, it even holds that every Souslin simulation is a Souslin combined simulation and accordingly, that every Souslin bisimulation is a Souslin combined bisimulation.* ⌟

We present our proof of Theorem 74 immediately after the following discussions. One cannot expect a simple argument for Theorem 74. Roughly speaking, the mathematical difficulty arises from the fact that an uncountable number of transitions may be put together into one single combined transition relying on the notion of barycentres for probability measures. However, in contrast, the notions of a simulation or bisimulation focus only on single individual transitions. Roughly speaking, we hence need a technique to extend properties of the transition relation to the corresponding combined-transition relation.

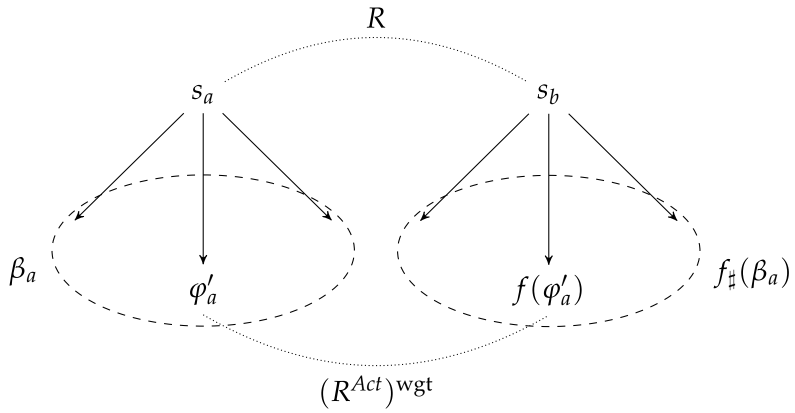


Figure 4.3: Illustration of the main idea for our proof of Theorem 74.

Using the schematic overview in Figure 4.3, we sketch the main idea for our proof of Theorem 74. This also provides a feeling for the use of the involved Souslin requirements. Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS, $s_a, s_b \in Sta$ be states, and $R$ be a Souslin simulation such that $\langle s_a, s_b \rangle \in R$. The main idea is to provide a Borel function

$$f\colon Prob[Act \times Sta] \rightarrow Prob[Act \times Sta]$$

that yields a Borel assignment of elements in $Enabled[s_a]$ to elements in $Enabled[s_b]$ while respecting the relation $R$. Formally, there exists a Borel set $P'_a \subseteq Prob[Act \times Sta]$ such that, among others, $P'_a \subseteq Enabled[s_a]$ and so that for every $\varphi'_a \in P'_a$ one has

$$f(\varphi'_a) \in Enabled[s_b] \quad \text{and} \quad \langle \varphi'_a, f(\varphi'_a) \rangle \in (R^{Act})^{\mathrm{wgt}}.$$

The existence of the function $f$ follows basically from the fact that $R$ is a simulation. However, to guarantee that $f$ can be chosen to be a Borel function, the measurable-selection principle stated in Theorem 21 is invoked. Thanks to the Souslin assumptions for the STS and the relation under consideration, it is straightforward to introduce an appropriate set-valued function for the intended purpose. The remaining argument for Theorem 74 proceeds as follows. Suppose probability measures $\varphi_a \in Prob[Act \times Sta]$ and $\beta_a \in Prob[Prob[Act \times Sta]]$ that, among others, satisfy

$$s_a \Rightarrow \varphi_a \quad \text{and} \quad \varphi_a = Barycen(\beta_a).$$

Defining $\varphi_b \in Prob[Act \times Sta]$ by

$$\varphi_b = Barycen(f_\sharp(\beta_a)),$$

it is then argued that

$$s_b \Rightarrow \varphi_b \quad \text{and} \quad \langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}.$$

This finally shows that every combined transition in $s_a$ can be matched by a combined transition in $s_b$ and therefore, the relation $R$ forms a combined simulation. The function $f$ thus formalises the intuitive idea of transferring properties of combined transitions enabled at state $s_a$ to the state $s_b$.

One might ask whether the corresponding reverse implications of Theorem 74 also hold, i.e., whether a combined simulation also forms a simulation and accordingly for bisimulations. However, the answer for this question is no in general. It is indeed easy to provide an STS with finite state space that serves as a counterexample.

**Proof of the section's main result.** In the remainder of this section we provide the details for the previously sketched proof of Theorem 74. For this we prove the following two auxiliary lemmas first whose proofs are straightforward.

**Lemma 75.** *Let $X$ be a Polish space and $R \subseteq X \times X$ be a Souslin set. Then the relation $R^{\mathrm{wgt}}$ is Souslin in $Prob[X] \times Prob[X]$.*

*Proof.* Define the Borel functions $\zeta_a\colon X \times X \to X$, $\zeta_a(x_a, x_b) = x_a$ and $\zeta_b\colon X \times X \to X$, $\zeta_b(x_a, x_b) = x_b$. Moreover, introduce the function $\zeta\colon \mathit{Prob}[X \times X] \to \mathit{Prob}[X] \times \mathit{Prob}[X]$,

$$\zeta(W) = \langle (\zeta_a)_\sharp(W), (\zeta_b)_\sharp(W) \rangle.$$

According to Remark 8, the function $\zeta$ is Borel. By Lemma 12, we moreover have

$$R^{\mathrm{wgt}} = \zeta(\{W \in \mathit{Prob}[X \times X] \;;\; W^{\mathrm{out}}(R) = 1\}).$$

This yields the claim applying Remark 10 and Lemma 14. □

**Lemma 76.** *Let $X$ be a Polish space, $R \subseteq X \times X$ be a relation, $f\colon X \to X$ be a Borel function, and $B \subseteq X$ be a Borel set such that*

$$\mathit{Graph}[f] \cap (B \times X) \subseteq R.$$

*For every $\mu \in \mathit{Prob}[X]$ such that $\mu(B) = 1$ it holds*

$$\langle \mu, f_\sharp(\mu) \rangle \in R^{\mathrm{wgt}}.$$

*Proof.* Let $\mu \in \mathit{Prob}[X]$. Introduce the Borel function $\zeta\colon X \to X \times X$,

$$\zeta(x) = \langle x, f(x) \rangle.$$

and the probability measure $W \in \mathit{Prob}[X \times X]$,

$$W = \zeta_\sharp(\mu).$$

We claim that $W$ is a weight function $(\mu, R, f_\sharp(\mu))$. It is easy to see that $W$ is a coupling of $(\mu, f_\sharp(\mu))$. Recall, $\mathit{Graph}[f]$ is Borel in $X \times X$ (see Section 2.1). It is easy to see that $W(\mathit{Graph}[f]) = \mu(X) = 1$. Assuming $\mu(B) = 1$, it moreover holds $W(B \times X) = 1$ and therefore, $W(\mathit{Graph}[f] \cap (B \times X)) = 1$. Since $\mathit{Graph}[f] \cap (B \times X) \subseteq R$, we conclude that $W$ is a weight function for $(\mu, R, f_\sharp(\mu))$. □

*Proof of Theorem 74.* It is sufficient to show statement (1) of Theorem 74 as a proof for (2) is completely analogous. Throughout this proof, the symbol $\Rightarrow$ denotes the combined-transition relation of the STS $\mathcal{T}$. Let $R$ be a Souslin simulation. Our task is to show that $R$ is even a combined simulation. For this purpose let $s_a, s_b \in \mathit{Sta}$ be states and $\varphi_a \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ be a probability measure such that

$$\langle s_a, s_b \rangle \in R \quad \text{and} \quad s_a \Rightarrow \varphi_a.$$

Moreover suppose $\beta_a \in \mathit{Prob}[\mathit{Prob}[\mathit{Sta}]]$ and a Borel set $P_a \subseteq \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ such that

$$P_a \subseteq \mathit{Enabled}[s_a], \quad \beta_a(P_a) = 1, \quad \text{and} \quad \varphi_a = \mathit{Barycen}(\beta_a).$$

In what follows we provide a probability measure $\varphi_b \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ satisfying the two properties $s_b \Rightarrow \varphi_b$ and $\langle \varphi_a, \varphi_b \rangle \in (R^{\mathit{Act}})^{\mathrm{wgt}}$. Our construction of this probability measure $\varphi_b$ relies on the measurable-selection principle in Theorem 21. Define the set-valued function $F\colon \mathit{Prob}[\mathit{Act} \times \mathit{Sta}] \rightsquigarrow \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$,

$$F(\varphi'_a) = \{\varphi'_b \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]\ ;\ s_b \to \varphi'_b \text{ and } \langle \varphi'_a, \varphi'_b \rangle \in (R^{\mathit{Act}})^{\mathrm{wgt}}\}.$$

By Lemmas 53 and 75, the set $\mathit{Rel}[F]$ is Souslin in $\mathit{Prob}[\mathit{Act} \times \mathit{Sta}] \times \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ since we have

$$\mathit{Rel}[F] = (R^{\mathit{Act}})^{\mathrm{wgt}} \cap (\mathit{Prob}[\mathit{Act} \times \mathit{Sta}] \times \mathit{Enabled}[s_b]).$$

For every $\varphi'_a \in P_a$ the set $F(\varphi'_a)$ is not empty relying on the fact that $R$ is a simulation for $\mathcal{T}$ and since we have $\langle s_a, s_b \rangle \in R$ and $s_a \to \varphi'_a$. Putting things together and as $\beta_a(P_a) = 1$, Theorem 21 yields the existence of a Borel $\beta_a$-selection of $F$.

Let $f\colon \mathit{Prob}[\mathit{Act} \times \mathit{Sta}] \to \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ a Borel function and $P'_a \subseteq \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ be a Borel set such that

$$\beta_a(P'_a) = 1 \quad \text{and} \quad f(\varphi'_a) \in F(\varphi'_a) \text{ for all } \varphi'_a \in P'_a.$$

Moreover, define the probability measures $\beta_b \in \mathit{Prob}[\mathit{Prob}[\mathit{Act} \times \mathit{Sta}]]$,

$$\beta_b = f_\sharp(\beta_a)$$

as well as $\varphi_b \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$,

$$\varphi_b = \mathit{Barycen}(\beta_b).$$

We claim that $s_b \Rightarrow \varphi_b$ and $\langle \varphi_a, \varphi_b \rangle \in (R^{\mathit{Act}})^{\mathrm{wgt}}$, which finally entails the theorem.

Since $\beta_a(P'_a) = 1$, Lemma 13 yields the existence a Borel set $P'_b \subseteq \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ such that $P'_b \subseteq f(P'_a)$ and $\beta_b(P'_b) = 1$. By the definition of the set-valued function $F$, we obtain $P'_b \subseteq f(P'_a) \subseteq \mathit{Enabled}[s_b]$ and therefore, $s_b \Rightarrow \varphi_b$.

It remains to show $\langle \varphi_a, \varphi_b \rangle \in (R^{\mathit{Act}})^{\mathrm{wgt}}$. As one has

$$\mathit{Graph}[f] \cap (P'_a \times \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]) \subseteq (R^{\mathit{Act}})^{\mathrm{wgt}},$$

Lemma 76 yields $\langle \beta_a, \beta_b \rangle \in ((R^{Act})^{\mathrm{wgt}})^{\mathrm{wgt}}$. Pick a weight function for $(\beta_a, (R^{Act})^{\mathrm{wgt}}, \beta_b)$, say $W_\beta$. According to Lemma 54, there exist a Borel function $g$ with domain $Prob[Act \times Sta] \times Prob[Act \times Sta]$ and codomain $Prob[(Act \times Sta) \times (Act \times Sta)]$ as well as a Borel set $R_\beta \subseteq Prob[Act \times Sta] \times Prob[Act \times Sta]$ so that

$$W_\beta(R_\beta) = 1$$

and

$$g(\varphi'_a, \varphi'_b) \text{ is a weight function for } (\varphi'_a, R^{Act}, \varphi'_b) \text{ for all } \langle \varphi'_a, \varphi'_b \rangle \in R_\beta.$$

Define $W'_\varphi \in Prob[Prob[(Act \times Sta) \times (Act \times Sta)]]$,

$$W'_\varphi = g_\sharp(W_\beta)$$

and $W_\varphi \in Prob[(Act \times Sta) \times (Act \times Sta)]$ by

$$W_\varphi = Barycen(W'_\varphi),$$

respectively. We claim that $W_\varphi$ is a weight function for $(\varphi_a, R^{Act}, \varphi_b)$.

For every Borel set $B \subseteq (Act \times Sta) \times (Act \times Sta)$ it holds

$$W_\varphi(B) = \int W''(B)\, dW'_\varphi(W'') = \int g(\varphi'_a, \varphi'_b)(B)\, dW_\beta(\varphi'_a, \varphi'_b).$$

From this one can easily conclude that $W_\varphi$ is a coupling of $(\varphi_a, \varphi_b)$. Indeed, using the properties of $W_\beta$ and $R_\beta$, for every Borel set $B_a \subseteq Act \times Sta$ it holds

$$\begin{aligned}
& W_\varphi(B_a \times (Act \times Sta)) \\
= \;& \int g(\varphi'_a, \varphi'_b)(B_a \times (Act \times Sta))\, dW_\beta(\varphi'_a, \varphi'_b) \\
= \;& \int \varphi'_a(B_a)\, dW_\beta(\varphi'_a, \varphi'_b) \\
= \;& \int \varphi'_a(B_a)\, d\beta_a(\varphi'_a) \\
= \;& \varphi_a(B_a).
\end{aligned}$$

One analogously shows the corresponding identity $W_\varphi((Act \times Sta) \times B_b) = \varphi_b(B_b)$ for every Borel set $B_b \subseteq Act \times Sta$.

Exploiting Lemma 12 and Lemma 53, it suffices to show $(W_\varphi)^{\mathrm{out}}(R^{Act}) = 1$ in order to conclude that $W_\varphi$ is a weight function for $(\varphi_a, R^{Act}, \varphi_b)$. Let $B \subseteq (Act \times Sta) \times (Act \times Sta)$

be a Borel set such that $R^{Act} \subseteq B$. For every $\langle \varphi'_a, \varphi'_b \rangle \in R_\beta$ we have $g(\varphi'_a, \varphi'_b)(B) = 1$. As it holds $W_\beta(R_\beta) = 1$, we obtain

$$W_\varphi(B) = \int g(\varphi'_a, \varphi'_b)(B)\, dW_\beta(\varphi'_a, \varphi'_b) = \int 1\, dW_\beta(\varphi'_a, \varphi'_b) = 1.$$

We finally conclude that $W_\varphi$ is a weight function for $(\varphi_a, R^{Act}, \varphi_b)$. This finishes our proof as already discussed before. $\square$

## 4.5 Simulations and bisimulations on finite paths

For non-stochastic systems it is easy to see that every simulation can be inductively extended to paths: indeed, given a simulation $R$ and states $s_a$ and $s_b$ with $\langle s_a, s_b \rangle \in R$, for every finite path $s_a\, act_{a1}\, s_{a1} \ldots act_{an}\, s_{an}$ there exists a path $s_b\, act_{b1}\, s_{b1} \ldots act_{bn}\, s_{bn}$ such that for every $i \in \{1, \ldots, n\}$ it holds $\langle \langle act_{ai}, s_{ai} \rangle, \langle act_{bi}, s_{bi} \rangle \rangle \in R^{Act}$, i.e., one has $act_{ai} = act_{bi}$ and $\langle s_{ai}, s_{bi} \rangle \in R$. This section presents a corresponding result for STSs where, one account of the presence of probabilities and non-determinism, schedulers and their induced path measures need to be taken into account.

The following notion is inspired by the initial sketched observation for non-stochastic systems. Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and $R \subseteq Sta \times Sta$ be a relation on its state space. For every $n \in \mathbb{N}$ we lift the relation $R$ to finite paths of length $n$. Formally, for every $n \in \mathbb{N}$ the relation $R^{\mathrm{path},n} \subseteq Path_n \times Path_n$ is given as follows: for all $\hat{\pi}_a, \hat{\pi}_b \in Path_n$,

$$\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R^{\mathrm{path},n} \quad \text{iff} \quad \begin{array}{l} \langle s_{a0}, s_{b0} \rangle \in R \text{ and} \\ \langle \langle act_{ai}, s_{ai} \rangle, \langle act_{bi}, s_{bi} \rangle \rangle \in R^{Act} \text{ for all } i \in \{1, \ldots, n\} \end{array}$$

where $\hat{\pi}_a = s_{a0}\, act_{a1}\, s_{a1} \ldots act_{an}\, s_{an}$ and $\hat{\pi}_b = s_{b0}\, act_{b1}\, s_{b1} \ldots act_{bn}\, s_{bn}$. We obviously have the identity $R^{\mathrm{path},0} = R$. The section's main result is as follows:

**Theorem 77.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a weakly Souslin STS, $R$ be a Souslin combined simulation, and $\mu_a, \mu_b \in Prob[Sta]$ be such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Then for every $\mu_a$-scheduler $\mathfrak{S}_a$ there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that for every $n \in \mathbb{N}$ it holds*

$$\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path},n})^{\mathrm{wgt}}.$$

$\lrcorner$

The formal proof is presented after the following discussions. Compared to the initial observation for non-stochastic systems, Theorem 77 requires the synthesis of a scheduler whose induced finite-path measures satisfy a certain condition. In Section 4.1 we have already illustrated mathematical difficulties caused by the measurability requirements for schedulers that motivate the notion of Souslin STSs. Invoking a measurable-selection principle, Theorem 65 shows that every non-blocking Souslin STS has a scheduler. Our proof for Theorem 77 basically relies on the same idea: we introduce a specific set-valued function that represents a set of candidates for appropriate schedulers. After that we show that the requirements of a measurable-selection principle are fulfilled that finally yields the existence of appropriate Borel functions. The previous discussions give a very high-level idea of our proof for Theorem 77. However, compared to Theorem 65, the mathematical argument below is much more intricate and technical at some points.

**Corollary 78.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS, $R$ be a Souslin simulation, as well as $\mu_a, \mu_b \in Prob[Sta]$ be such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Then for every $\mu_a$-scheduler $\mathfrak{S}_a$ there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that for all $n \in \mathbb{N}$ it holds*

$$\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path},n})^{\mathrm{wgt}}.$$

*Proof.* The claim follows from Theorem 77 together with Theorems 70 and 74. □

As a preview, Theorem 77 and its Corollary 78 are object of further discussions in subsequent chapters of this thesis. For instance, in Section 5.5 we prove the existence of a $\mu_b$-scheduler such that $\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path}})^{\mathrm{wgt}}$ where $R^{\mathrm{path}}$ denotes the lifting of the relation $R$ to infinite paths. Even though the latter sketched result follows from Theorem 77 or Corollary 78, the argument is not straightforward. Besides this, in Section 7.2 we present a much simpler argument for Theorem 77 for those Souslin STSs that are additionally purely stochastic.

**Proof of section's main results.** The remainder of this section provides the details for our proof of Theorem 77. We start with the following very simple observation.

**Lemma 79.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and $R \subseteq Sta \times Sta$ be a Souslin set. For every $n \in \mathbb{N}$ the set $R^{\mathrm{path},n}$ is Souslin in $Path_n \times Path_n$.*

*Proof.* The proof is simple and completely analogous to Lemma 53. □

For the next lemma it is convenient to recall the definition of the post operator from Section 2.1. Given measurable spaces $X$ and $Y$, a probability measure $\mu \in Prob[X]$, and a measurable function $f \colon X \rightarrow Prob[Y]$, then $Post[\mu, f]$ defines a probability measure on $Y$

that is intuitively obtained by averaging over all probability measures $f(x)$ where $x \in X$ with respect to the measure $\mu$, i.e., for every Borel set $B_Y \subseteq Y$ it holds

$$Post[\mu, f](B_Y) = \int f(x)(B_Y)\, d\mu(x).$$

**Lemma 80.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS and $s \in Sta$ be a state. Denote the combined-transition relation by $\Rightarrow$. Suppose a measurable space $X$, a measurable set $B_X \subseteq X$, a probability measure $\mu \in Prob[X]$, and a measurable function $f \colon X \to Prob[Act \times Sta]$ such that*

$$\mu(B_X) = 1 \quad \textit{and} \quad s \Rightarrow f(x) \textit{ for all } x \in B_X.$$

*Then it holds*

$$s \Rightarrow Post[\mu, f].$$

*Proof.* For convenience, let us introduce the following two abbreviations:

$$\varphi = Post[\mu, f] \quad \text{and} \quad \beta = f_\sharp(\mu).$$

First of all it holds $\varphi = Barycen(\beta)$ since for every Borel set $B \subseteq Act \times Sta$ one has

$$\varphi(B) = \int f(x)(B)\, d\mu(x) = \int \varphi'(B)\, df_\sharp(\mu)(\varphi') = Barycen(\beta)(B).$$

The set $Enabled[s]$ is Souslin in $Prob[Act \times Sta]$ by Remark 67 and therefore, the set $Conv[Enabled[s]]$ is Souslin in $Prob[Act \times Sta]$ by Theorem 27. The combined-transition relation of $\mathcal{T}$ is closed under combining inspecting again Remark 67. To conclude the lemma, it hence suffices to show $\beta^{\text{out}}(Conv[Enabled[s]]) = 1$ thanks to Lemma 12.

Applying Lemma 13, it holds

$$\beta^{\text{out}}(f(B)) = 1.$$

We moreover have the following inclusion:

$$f(B) \subseteq Conv[Enabled[s]].$$

Indeed, given $\varphi' \in f(B)$, then there exists $x \in B$ such that $f(x) = \varphi'$, which directly entails $s \Rightarrow \varphi'$. From this we conclude $\beta^{\text{out}}(Conv[Enabled[s]]) = 1$. $\square$

Intuitively, assuming the setting as in Lemma 80, the function $f$ assigns every element of the measurable set $B_X$ to a combined transition enabled at the state $s$. The lemma then shows that averaging over these annotated combined transitions with respect to the given probability measure $\mu$ also yields a combined transition enabled at the state $s$.

*Proof of Theorem 77.* As usual the symbol $\Rightarrow$ denotes the combined-transition relation of the given STS $\mathcal{T}$. Pick a $\mu_a$-scheduler, say $\mathfrak{S}_a$. Our task is to provide a $\mu_b$-scheduler $\mathfrak{S}_b$ such that for every $n \in \mathbb{N}$ the induced finite-path measures $Pr_n[\mathfrak{S}_a]$ and $Pr_n[\mathfrak{S}_b]$ are related concerning the weight lifting of the relation $R^{\mathrm{path},n}$.

For every $n \in \mathbb{N}$ define the function $\hat{\mathfrak{S}}_{a,n} \colon Path_n \to Prob[Act \times Sta]$,

$$\hat{\mathfrak{S}}_{a,n}(\hat{\pi}) = \mathfrak{S}_a(\hat{\pi})$$

as well as the probability measure $\hat{\chi}_{a,n} \in Prob[Path_n]$,

$$\hat{\chi}_{a,n} = Pr_n[\mathfrak{S}_a].$$

We obviously have $\hat{\chi}_{a,0} = \mu_a$. For every $n \in \mathbb{N}$ the function $\hat{\mathfrak{S}}_{a,n}$ is Borel, it holds

$$Pr_{n+1}[\mathfrak{S}_a] = \hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}$$

and moreover, there exists a Borel set $\hat{\Pi}_{a,n} \subseteq Path_n$ such that

$$\hat{\chi}_{a,n}(\hat{\Pi}_{a,n}) = 1 \quad \text{and} \quad \mathrm{Last}_n(\hat{\pi}_a) \Rightarrow \hat{\mathfrak{S}}_{a,n}(\hat{\pi}_b) \text{ for all } \hat{\pi}_a \in \hat{\Pi}_{a,n}.$$

*Main claim.* The key observation of this proof is summarised by the following statement. For every $n \in \mathbb{N}$ there exist a probability measure $\hat{\chi}_{b,n} \in Prob[Path_n]$ as well as a Borel function $\hat{\mathfrak{S}}_{b,n} \colon Path_n \to Prob[Act \times Sta]$ such that the following three properties hold:

(a) It holds $\hat{\chi}_{b,0} = \mu_b$ and for every $n \in \mathbb{N} \setminus \{0\}$ one has $\hat{\chi}_{b,n+1} = \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n}$.

(b) For every $n \in \mathbb{N}$ there is a Borel set $\hat{\Pi}_{b,n} \subseteq Path_n$ with the properties

$$\hat{\chi}_{b,n}(\hat{\Pi}_{b,n}) = 1 \quad \text{and} \quad \mathrm{Last}_n(\hat{\pi}_b) \Rightarrow \hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b) \text{ for all } \hat{\pi}_b \in \hat{\Pi}_{b,n}.$$

(c) For every $n \in \mathbb{N}$ one has $\langle \hat{\chi}_{a,n}, \hat{\chi}_{b,n} \rangle \in (R^{\mathrm{path},n})^{\mathrm{wgt}}$.

Let us derive the theorem assuming for a moment that the previous claim is already proven. For every $n \in \mathbb{N}$ let $\hat{\chi}_{b,n}$ and $\hat{\mathfrak{S}}_{b,n}$ as in the claim, in particular, satisfying the three statements (1), (2), and (3). Moreover, let $\mathfrak{S}_b \colon Path_{<\omega} \to Prob[Sta] \cup Prob[Act \times Sta]$ be the function such that

$$\mathfrak{S}_b(\circ) = \mu_b \quad \text{and} \quad \mathfrak{S}_{b|n} = \hat{\mathfrak{S}}_{b,n} \text{ for all } n \in \mathbb{N},$$

In particular, for every $n \in \mathbb{N}$ and $\hat{\pi}_b \in Path_n$ it holds $\mathfrak{S}_b(\hat{\pi}_b) = \hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b)$.

As the function $\hat{\mathfrak{S}}_{b,n}$ is Borel for every $n \in \mathbb{N}$, it directly follows that $\mathfrak{S}_b$ is Borel. An easy induction over the natural numbers together with statement (1) show that for all $n \in \mathbb{N}$,

$$Pr_n[\mathfrak{S}_b] = \hat{\chi}_{b,n} \quad \text{and} \quad Pr_{n+1}[\mathfrak{S}_b] = \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n}.$$

By statement (2), it hence follows that $\mathfrak{S}_b$ forms a $\mu_b$-scheduler. Moreover, statement (3) yields the existence of a weight function for $(Pr_n[\mathfrak{S}_a], R^{\text{path},n}, Pr_n[\mathfrak{S}_b])$ for every $n \in \mathbb{N}$. Putting things together, we obtain a proof of Theorem 77.

*Proof of main claim.* Let $n \in \mathbb{N}$ be a natural number. Suppose a probability measure $\hat{\chi}_{b,n} \in Prob[Path_n]$ satisfying $\langle \hat{\chi}_{a,n}, \hat{\chi}_{b,n} \rangle \in (R^{\text{path},n})^{\text{wgt}}$. Let $W_n$ be a weight function for $(\hat{\chi}_{a,n}, R^{\text{path},n}, \hat{\chi}_{b,n})$. In the remainder of this proof we provide a Borel function $\hat{\mathfrak{S}}_{b,n} \colon Path_n \to Prob[Act \times Sta]$ such that there exists a Borel set $\hat{\Pi}_{b,n} \subseteq Path_n$ with

$$\hat{\chi}_{b,n}(\hat{\Pi}_{b,n}) = 1 \quad \text{and} \quad \text{Last}_n(\hat{\pi}_b) \Rightarrow \hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b) \text{ for all } \hat{\pi}_b \in \hat{\Pi}_{b,n}$$

and moreover, such that it holds

$$\langle \hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}, \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n} \rangle \in (R^{\text{path},n+1})^{\text{wgt}}.$$

Relying on an inductive argument, it is then easy to deduce the previously stated main claim of this proof.

The following proof is comparatively long and in some of its parts also rather technical. In a first step, we introduce an appropriate set-valued function and justify the existence of a Borel selection. Exploiting this Borel selection, we then provide a definition of the function $\hat{\mathfrak{S}}_{b,n}$. After that, it remains to show that the desired properties mentioned before are fulfilled.

*Set-valued function $F$ and Borel $W_n$-selection.* We introduce a set-valued function $F$ and justify the existence of a Borel $W_n$-selection. Define $F\colon Path_n \times Path_n \rightsquigarrow Prob[Act \times Sta]$,

$$F(\hat{\pi}_a, \hat{\pi}_b) = \{\varphi_b \in Prob[Act \times Sta] \; ; \; \text{Last}_n(\hat{\pi}_b) \Rightarrow \varphi_b \text{ and } \langle \hat{\mathfrak{S}}_{a,n}(\hat{\pi}_a), \varphi_b \rangle \in (R^{Act})^{\text{wgt}}\}.$$

As the probability measure $W_n$ is a weight function for $(\hat{\chi}_{a,n}, R^{\text{path},n}, \hat{\chi}_{b,n})$, there exists a Borel set $R_{W,n} \subseteq Path_n \times Path_n$ with $R_{W,n} \subseteq R^{\text{path},n}$ and $W_n(R_{W,n}) = 1$. Define the Borel set $R'_{W,n} \subseteq Path_n \times Path_n$ by $R'_{W,n} = R_{W,n} \cap (\hat{\Pi}_{a,n} \times Path_n)$. As $\hat{\chi}_{a,n}(\hat{\Pi}_{a,n}) = 1$, we obtain $W_n(R'_{W,n}) = 1$. Moreover, using that $R$ is a combined-transition simulation, the set $F(\hat{\pi}_a, \hat{\pi}_b)$ is not empty for every $\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R'_{W,n}$.

The set $Rel[F]$ is Souslin in $Path_n \times Path_n \times Prob[Act \times Sta]$ that can be seen as follows. Introduce the Borel function $\zeta$ with domain $Path_n \times Path_n \times Prob[Act \times Sta]$ and codomain $Sta \times Prob[Act \times Sta] \times Prob[Act \times Sta]$ as follows: for all $\hat{\pi}_a \in Path_n$, $\hat{\pi}_b \in Path_n$, and $\varphi_b \in Prob[Act \times Sta]$ let

$$\zeta(\hat{\pi}_a, \hat{\pi}_b, \varphi_b) = \langle \mathrm{Last}_n(\hat{\pi}_b), \varphi_b, \hat{\mathfrak{S}}_{a,n}(\hat{\pi}_a) \rangle.$$

We immediately obtain the following identity

$$Rel[F] = \zeta^{-1}((\Rightarrow \times Prob[Act \times Sta]) \cap (Sta \times (R^{Act})^{\mathrm{wgt}})).$$

The set $(R^{Act})^{\mathrm{wgt}}$ is Souslin in $Prob[Act \times Sta] \times Prob[Act \times Sta]$ exploiting Lemmas 53 and 75. As the STS $\mathcal{T}$ is required to be weakly Souslin, it follows that the set $Rel[F]$ is Souslin in $Path_n \times Path_n \times Prob[Act \times Sta]$ by Remark 10 (5).

*Definition of the Borel function* $\hat{\mathfrak{S}}_{b,n}$. We saw that there exists a Borel $W_n$-selection of $F$, say $f$. For every $\hat{\pi}_b \in Path_n$ define the function $f_{\hat{\pi}_b}\colon Path_n \to Prob[Act \times Sta]$,

$$f_{\hat{\pi}_b}(\hat{\pi}_a) = f(\hat{\pi}_a, \hat{\pi}_b).$$

A standard argument shows that the function $f_{\hat{\pi}_b}$ is Borel for all $\hat{\pi}_b \in Path_n$.

By the disintegration theorem (see Section 2.1), there moreover exists a Borel function $g\colon Path_n \to Prob[Path_n]$ with

$$W_n = g \ltimes \hat{\chi}_{b,n}.$$

Using the functions $f$ and $g$, the function $\hat{\mathfrak{S}}_{b,n}\colon Path_n \to Prob[Act \times Sta]$ is introduced as follows: for all $\hat{\pi}_b \in Path_n$ define

$$\hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b) = Post[g(\hat{\pi}_b), f_{\hat{\pi}_b}],$$

i.e., for every $\hat{\pi}_b \in Path_n$ and Borel set $B_b \subseteq Act \times Sta$ it holds

$$\hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b)(B_b) = \int f_{\hat{\pi}_b}(\hat{\pi}_a)(B_b)\, dg(\hat{\pi}_b)(\hat{\pi}_a).$$

Fubini's theorem (see Section 2.1) justifies that $\hat{\mathfrak{S}}_{b,n}$ is a Borel function.

*Compatibility with the combined-transition relation.* Since $f$ is a Borel $W_n$-selection of $F$, there exists a Borel set $R_f \subseteq Path_n \times Path_n$ such that

$$W_n(R_f) = 1 \quad \text{and} \quad f(\hat{\pi}_a, \hat{\pi}_b) \in F(\hat{\pi}_a, \hat{\pi}_b) \text{ for all } \langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R_f.$$

The set $Section[R_f, \cdot, \hat{\pi}_b]$ is Borel in $Path_n$ for all $\hat{\pi}_b \in Path_n$ (see Section 2.1). Thanks to Lemma 15, there hence exists a Borel set $\hat{\Pi}_{b,n} \subseteq Path_n$ such that the following properties are fulfilled:

$$\hat{\chi}_{b,n}(\hat{\Pi}_{b,n}) = 1 \quad \text{and} \quad g(\hat{\pi}_b)(Section[R_f, \cdot, \hat{\pi}_b]) = 1 \text{ for all } \hat{\pi}_b \in \hat{\Pi}_{b,n}.$$

Let $\hat{\pi}_b \in \hat{\Pi}_{b,n}$ It remains to show $\text{Last}(\hat{\pi}_b) \Rightarrow \hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b)$. For all $\hat{\pi}_a \in Section[R_f, \cdot, \hat{\pi}_b]$ one has $f(\hat{\pi}_a, \hat{\pi}_b) \in F(\hat{\pi}_a, \hat{\pi}_b)$ and therefore, it holds

$$\text{Last}_n(\hat{\pi}_b) \Rightarrow f_{\hat{\pi}_b}(\hat{\pi}_a).$$

Since $g(\hat{\pi}_b)(Section[R_f, \cdot, \hat{\pi}_b]) = 1$, we can apply Lemma 80 and conclude,

$$\text{Last}_n(\hat{\pi}_b) \Rightarrow Post[g(\hat{\pi}_b), f_{\hat{\pi}_b}].$$

By the definition of $\hat{\mathfrak{S}}_{b,n}$, this finally yields $\text{Last}(\hat{\pi}_b) \Rightarrow \hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b)$.

*Existence of a weight function.* As $f$ is a Borel $W_n$-selection of the set-valued function $F$ and since $W_n$ is a weight function for $(\hat{\chi}_{a,n}, R^{\text{path},n}, \hat{\chi}_{b,n})$, Lemma 54 yields the existence of a Borel function $h\colon Path_n \times Path_n \to Prob[(Act \times Sta) \times (Act \times Sta)]$ and of a Borel set $R_h \subseteq Path_n \times Path_n$ with the following properties:

$$W_n(R_h) = 1 \quad \text{and} \quad R_h \subseteq R^{\text{path},n}$$

and for all $\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R_h$,

$$h(\hat{\pi}_a, \hat{\pi}_b) \text{ is a weight function for } (\hat{\mathfrak{S}}_{a,n}(\hat{\pi}_a), R^{Act}, f(\hat{\pi}_a, \hat{\pi}_b)).$$

Let $\zeta$ be the function with domain $Path_n \times Path_n \times (Act \times Sta) \times (Act \times Sta)$ and codomain $Path_{n+1} \times Path_{n+1}$ such that for all $\hat{\pi}_a \in Path_n$, $\hat{\pi}_b \in Path_n$, $\langle act_a, s_a \rangle \in Act \times Sta$, and $\langle act_b, s_b \rangle \in Act \times Sta$ it holds

$$\zeta(\hat{\pi}_a, \hat{\pi}_b, \langle act_a, s_a \rangle, \langle act_b, s_b \rangle) = \langle \hat{\pi}_a\, act_a\, s_a, \hat{\pi}_b\, act_b\, s_b \rangle.$$

Moreover, define $W_{n+1} \in Prob[Path_{n+1} \times Path_{n+1}]$ by

$$W_{n+1} = \zeta_\sharp(W_n \rtimes h).$$

It remains to show that $W_{n+1}$ is a weight function for $(\hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}, R^{\text{path},n+1}, \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n})$.

We justify that $W_{n+1}$ is a coupling of $(\hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}, \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n})$ first. Let $\hat{\Pi}'_{b,n} \subseteq Path_n$ and $B'_b \subseteq Act \times Sta$ be Borel sets. It holds,

$$W_{n+1}(Path_{n+1} \times (\hat{\Pi}'_{b,n} \times B'_b))$$

$$
\begin{aligned}
&= \int_{Path_n \times \hat{\Pi}'_{b,n}} h(\hat{\pi}_a, \hat{\pi}_b)((Act \times Sta) \times B'_b)\, dW_n(\hat{\pi}_a, \hat{\pi}_b) \\
&= \int_{Path_n \times \hat{\Pi}'_{b,n}} f(\hat{\pi}_a, \hat{\pi}_b)(B'_b)\, dW_n(\hat{\pi}_a, \hat{\pi}_b) \\
&= \int_{\hat{\Pi}'_{b,n}} \left( \int f_{\hat{\pi}_b}(\hat{\pi}_a)(B'_b)\, dg(\hat{\pi}_b)(\hat{\pi}_a) \right) d\hat{\chi}_{b,n}(\hat{\pi}_b) \\
&= \int_{\hat{\Pi}'_{b,n}} \hat{\mathfrak{S}}_{b,n}(\hat{\pi}_b)(B'_b)\, d\hat{\chi}_{b,n}(\hat{\pi}_b) \\
&= \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n}(\hat{\Pi}'_{b,n} \times B'_b)
\end{aligned}
$$

From this, we immediately obtain $W_{n+1}(Path_{n+1} \times \hat{\Pi}'_{b,n+1}) = \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n}(\hat{\Pi}'_{b,n+1})$ for all Borel sets $\hat{\Pi}'_{b,n+1} \subseteq Path_{n+1}$ by Carathéodory extension theorem (see Section 2.1).

Let $\hat{\Pi}'_{a,n} \subseteq Path_n$ and $B'_a \subseteq Act \times Sta$ be Borel sets. We have

$$
\begin{aligned}
& W_{n+1}((\hat{\Pi}'_{a,n} \times B'_a) \times Path_{n+1}) \\
&= \int_{\hat{\Pi}_{a,n} \times Path_n} h(\hat{\pi}_a, \hat{\pi}_b)(B'_a \times (Act \times Sta))\, dW_n(\hat{\pi}_a, \hat{\pi}_b) \\
&= \int_{\hat{\Pi}'_{a,n} \times Path_n} \hat{\mathfrak{S}}_{a,n}(\hat{\pi}_a)(B'_a)\, dW_n(\hat{\pi}_a, \hat{\pi}_b) \\
&= \int_{\hat{\Pi}'_{a,n}} \hat{\mathfrak{S}}_{a,n}(\hat{\pi}_a)(B'_a)\, d\hat{\chi}_{a,n}(\hat{\pi}_a) \\
&= \hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}(\hat{\Pi}'_{a,n} \times B'_a)
\end{aligned}
$$

Again, thanks to Carathéodory extension, for all Borel sets $\hat{\Pi}'_{a,n+1} \subseteq Path_{n+1}$ it holds $W_{n+1}(\hat{\Pi}'_{a,n+1} \times Path_{n+1}) = \hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}(\hat{\Pi}'_{a,n+1})$. Putting things together, $W_{n+1}$ is a coupling of $(\hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}, \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n})$.

By Lemma 79, the set $R^{\text{path},n+1}$ is Souslin in $Path_{n+1} \times Path_{n+1}$. Thanks to Lemma 12, it suffices to show $(W_{n+1})^{\text{out}}(R^{\text{path},n+1}) = 1$ in order to finish the argument for (3). Let $R_W \subseteq Path_{n+1} \times Path_{n+1}$ be a Borel set such that $R^{\text{path},n+1} \subseteq R_W$. For every $\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R^{\text{path},n}$ we obtain

$$
R^{Act} \subseteq Section[\zeta^{-1}(R_W), \langle \hat{\pi}_a, \hat{\pi}_b \rangle, \cdot].
$$

For all $\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R_h$ it holds $(h(\hat{\pi}_a, \hat{\pi}_b))^{\text{out}}(R^{Act}) = 1$ and therefore,

$$
h(\hat{\pi}_a, \hat{\pi}_b)(Section[\zeta^{-1}(R_W), \langle \hat{\pi}_a, \hat{\pi}_b \rangle, \cdot]) = 1.
$$

Since $W_n(R_h) = 1$, we thus obtain

$$
W_{n+1}(R_W) = \int h(\hat{\pi}_a, \hat{\pi}_b)(Section[\zeta^{-1}(R_W), \langle \hat{\pi}_a, \hat{\pi}_b \rangle, \cdot])\, dW_n(\hat{\pi}_a, \hat{\pi}_b) = 1.
$$

This proves $(W_{n+1})^{\text{out}}(R^{\text{path},n+1}) = 1$. As a consequence, the probability measure $W_{n+1}$ forms a weight function for $(\hat{\chi}_{a,n} \rtimes \hat{\mathfrak{S}}_{a,n}, R^{\text{path},n+1}, \hat{\chi}_{b,n} \rtimes \hat{\mathfrak{S}}_{b,n})$. □

## 4.6 Preservation of trace distributions

We show that in Souslin STSs the simulation preorder is subsumed by the trace-distribution preorder and accordingly, that the Souslin-bisimulation equivalence is finer than the trace-distribution equivalence. At the end of this section, we provide two slight generalisations of this result. Given the preparatory work carried out by the previous sections, it remains to show the following simple observation:

**Lemma 81.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS, $R \subseteq Sta \times Sta$ be a relation, as well as $\mu_a, \mu_b \in Prob[Sta]$ be probability measures. Moreover, consider a $\mu_a$-scheduler $\mathfrak{S}_a$ and a $\mu_b$-scheduler $\mathfrak{S}_b$ such that for all $n \in \mathbb{N}$,*

$$\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b]\rangle \in (R^{\text{path},n})^{\text{wgt}}.$$

*Then the following identity holds:*

$$PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b].$$

*Proof.* Let $n \in \mathbb{N} \setminus \{0\}$ and $A_1, \ldots, A_n$ be Borel subsets of $Act$. Define

$$A = A_1 \times \ldots \times A_n \times Act^{\omega}.$$

as well as

$$\hat{\Pi} = Sta \times (A_1 \times Sta) \times \ldots \times (A_n \times Sta).$$

Then the sets $A$ and $\hat{\Pi}$ are Borel in $Act^{\omega}$ and $Path_n$, respectively. Moreover, it is easy to see that the set $\hat{\Pi}$ is $R^{\text{path},n}$-stable. Since $\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b]\rangle \in (R^{\text{path},n})^{\text{wgt}}$, Remark 34 yields

$$PrTrace[\mathfrak{S}_a](A) = Pr_n[\mathfrak{S}_a](\hat{\Pi}) = Pr_n[\mathfrak{S}_b](\hat{\Pi}) = PrTrace[\mathfrak{S}_b](A).$$

Thanks to Carathéodory's uniqueness theorem (see Section 2.1), we finally obtain the identity $PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b]$. □

The lemma gives a simple sufficient criterion for proving that two schedulers induce the same trace distribution. This together with the achievements of the previous sections in this chapter yield the first main result of this thesis:

**Theorem 82.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a weakly Souslin STS. For every states $s_a, s_b \in Sta$ the following two statements hold:*

(1) $s_a \preceq^{\text{sou,c}} s_b$ *implies* $s_a \leq^{\text{tr}} s_b$.

(2) $s_a \simeq^{\text{sou,c}} s_b$ *implies* $s_a =^{\text{tr}} s_b$.

*Proof.* Let $s_a, s_b \in Sta$ be states with $s_a \preceq^{\text{sou,c}} s_b$. Moreover let $R$ be a Souslin combined simulation such that $\langle s_a, s_b \rangle \in R$. Suppose an $s_a$-scheduler $\mathfrak{S}_a$. According to Theorem 77, there exists an $s_b$-scheduler $\mathfrak{S}_b$ such that for all $n \in \mathbb{N}$ it holds

$$\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b] \rangle \in (R^{\text{path},n})^{\text{wgt}}.$$

Thanks to Lemma 81, we obtain

$$PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b].$$

This entails $s_a \leq^{\text{tr}} s_b$ that justifies statement (1). Obviously, (2) follows directly from (1). □

**Corollary 83.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS. For every states $s_a, s_b \in Sta$ the following two statements hold:*

(1) $s_a \preceq^{\text{sou}} s_b$ *implies* $s_a \leq^{\text{tr}} s_b$.

(2) $s_a \simeq^{\text{sou}} s_b$ *implies* $s_a =^{\text{tr}} s_b$.

*Proof.* The claim follows from Theorem 82 together with Theorems 70 and 74. □

As a consequence, establishing a Souslin-simulation and a Souslin-bisimulation are sound proof techniques for proving trace-distribution preorder and trace-distribution equivalence in Souslin STSs, respectively. Standard discussions concerning linear-time and branching-time behaviour of operational systems (see, e.g., [BK08]) show that the respective reverse implications of Theorem 82 and its Corollary 83 do not hold in general. It can be regarded as folklore that there are examples of finite non-stochastic systems where $s_a =^{\text{tr}} s_b$ while neither $s_a \preceq s_b$ nor $s_a \simeq s_b$ for some states $s_a$ and $s_b$. However, as we see in Section 4.8, for the subclass of Souslin STSs consisting of those systems that are deterministic and purely stochastic the relations induced by (Souslin) simulation, (Souslin) bisimulation, and trace distributions are the same.

It turns out that Theorem 82 and its Corollary 83 do not hold if one drops the Souslin requirement for the STS under consideration. The following example is borrowed from Example 9.2 in [Cat05] and basically continues our discussions from Section 4.1.

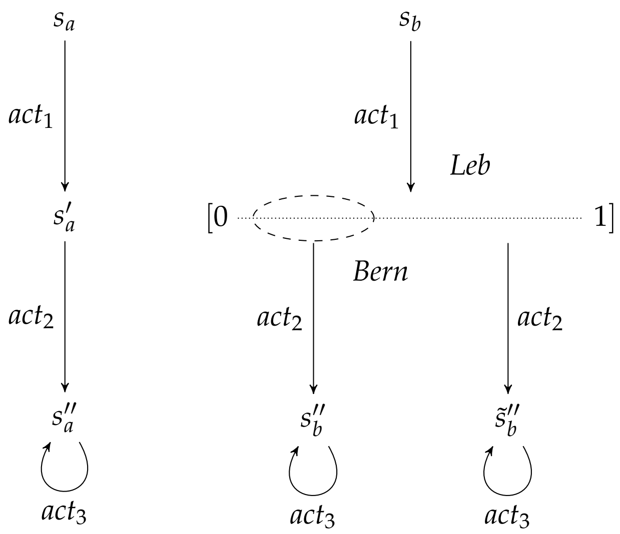


Figure 4.4: It holds $s_a \simeq^{\mathrm{sou}} s_b$ but neither $s_a \leq^{\mathrm{tr}} s_b$ nor $s_a =^{\mathrm{tr}} s_b$.

**Example 84** (Example 9.2 in [Cat05])**.** Suppose the STS $\mathcal{T}$ involving the states $s_a$ and $s_b$ is given as in Figure 4.4. Here, the reachable part from state $s_b$ is defined as in Example 63 involving a Bernstein set *Bern*. Thanks to Examples 63 and 66, the STS $\mathcal{T}$ is not Souslin and moreover, there exists no $s_b$-scheduler. In contrast, it is easy to see that there is a $s_a$-scheduler and thus, the assertion $s_a \leq^{\mathrm{tr}} s_b$ does not hold. However, we obviously have $s_a \simeq^{\mathrm{sou,c}} s_b$ as well as $s_a \simeq s_b$. As a result of this example, the thesis [Cat05] proposes a global bisimulation notion. This approach is detailed discussed in the next section. ⌟

Theorem 82 as well as its Corollary 83 consider only trace distributions where the initial conditions of corresponding schedulers are given by a single state. However, the following proposition illustrates that the main results of this section can be easily adapted for trace distributions whose generating schedulers involve arbitrary initial distributions over the state space:

**Proposition 85.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS, $R$ be an Souslin simulation, and $\mu_a, \mu_b \in Prob[Sta]$ be such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Then for every $\mu_a$-scheduler $\mathfrak{S}_a$ there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that*

$$PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b].$$

*Proof.* Inspecting Lemmas 81 and 81, one can apply the same argument as for Theorem 82 and its Corollary 83. □

## 4.7 Discussion on Cattani's result

The measurability requirements on schedulers cause the main mathematical challenges for a proof of Theorem 82 and its Corollary 83 stating that, besides others, trace-distribution equivalence is finer than bisimulation equivalence under Souslin side constraints. Illustrated by Example 84 (see also Example 9.2 in [Cat05]), the problem is that the used standard notion for bisimulation does not distinguish between states with, roughly speaking, different measurable structure. Indeed, there might be states $s_a$ and $s_b$ of an STS with $s_a \simeq^{\mathrm{sou}} s_b$ such that there exists an $s_a$-scheduler while there is no $s_b$-scheduler. For that reason, reflecting our idea introduced in Section 4.1, we develop a theory based on Souslin STSs (see also Theorem 65). In contrast, Cattani proposes another approach in his thesis [Cat05] to treat the measurability problems caused by schedulers: instead of focusing on the classical notion for bisimulation, Cattani regards a global variant of bisimulation. Given some side constraints, Theorem 9.19 in [Cat05] then shows that global-bisimulation equivalence is finer than trace-distribution equivalence. This section discusses this approach and relates our result with that in [Cat05].

The starting point of [Cat05] is to view STSs as transformers of probabilities [KVAK10, DMS14, AAGT15]. In the standard semantics for stochastic systems, stochastic choices over successor states are resolved in the sense that each step continues from a single state. However, viewing stochastic systems as transformers of probabilities, the stochastic choice is not resolved and one continuous from a distribution over states in every step instead. This informal description leads to the notion of the global-transition relation and corresponding notions of (bi)simulations [Cat05, DHR08, EHZ10, Hen12, HKK14, FZ14, YJZ17].

For the following material it is appropriate to recall the definition of the post operator from Section 2.1.

**Definition 86.** For every STS $\mathcal{T} = (Sta, Act, \rightarrow)$ the *global-transition relation*

$$\rightrightarrows \subseteq Prob[Sta] \times Prob[Act \times Sta]$$

consists of exactly those pairs $\langle \mu, \varphi \rangle \in Prob[Sta] \times Prob[Act \times Sta]$ that satisfy the following statement: there exist a Borel function $\tilde{\mathfrak{S}} \colon Sta \rightarrow Prob[Act \times Sta]$ and a Borel set $S \subseteq Sta$ such that

$$\mu(S) = 1, \quad s \Rightarrow \tilde{\mathfrak{S}}(s) \text{ for all } s \in S, \quad \text{and} \quad \varphi = Post[\mu, \tilde{\mathfrak{S}}].$$

In this context, we refer to the pair $(\tilde{\mathfrak{S}}, S)$ as a *witness for* $\mu \rightrightarrows \varphi$. ⌟

Our previous definition corresponds to Definition 9.7 in [Cat05] where, however, one refers to global transitions as hyper transitions. Whereas each transition of an STS starts in a single state, a global transition continues from a distribution over states. Intuitively, a global transition includes a measurable bundle of transitions that are put together by averaging with respect to a probability measure.

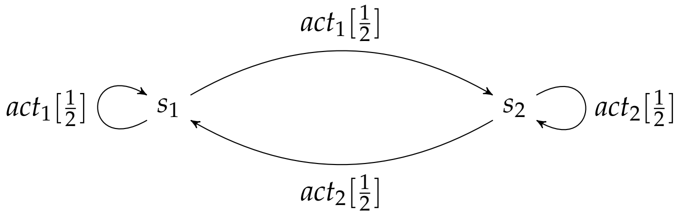


Figure 4.5: Simple STS for the illustration of the global transition relation.

**Example 87.** Consider the simple STS $\mathcal{T} = (Sta, Act, \rightarrow)$ depicted in Figure 4.5, i.e., $Sta = \{s_1, s_2\}$ and $Act = \{act_1, act_2\}$. Denote the corresponding global transition by $\rightrightarrows$. Let $\mu \in Prob[Sta]$ be the probability measure such that $\mu(\{s_1\}) = 1/2$ and $\mu(\{s_2\}) = 1/2$. For every $\varphi \in Prob[Act \times Sta]$ one has $\mu \rightrightarrows \varphi$ iff the following four statements hold:

$$\begin{aligned}
&\varphi(\{act_1\} \times \{s_1\}) = 1/4, \quad \varphi(\{act_2\} \times \{s_1\}) = 1/4,\\
&\varphi(\{act_1\} \times \{s_2\}) = 1/4, \quad \varphi(\{act_2\} \times \{s_2\}) = 1/4.
\end{aligned}$$

Consequently, for every probability measure $\varphi \in Prob[Act \times Sta]$ such that $\mu \rightrightarrows \varphi$ one obtains $\varphi(Act \times \{s_1\}) = 1/2 = \mu(\{s_1\})$ and $\varphi(Act \times \{s_2\}) = 1/2 = \mu(\{s_2\})$. Thus, for every execution continuing from $\mu$ with respect to the global-transition relation the distribution over the state space remains invariant and is given by $\mu$. ⌟

Relying on the global-transition relation, we introduce the following alternative notions for the comparison of the branching-time structure of states in an STS:

$$\begin{array}{ccc c ccc}
\mu_a & R^{\mathrm{wgt}} & \mu_b & & \mu_a & R^{\mathrm{wgt}} & \mu_b \\
\downdownarrows & & & \text{can be completed to} & \downdownarrows & & \downdownarrows \\
\varphi_a & & & & \varphi_a & (R^{Act})^{\mathrm{wgt}} & \varphi_b
\end{array}$$

Figure 4.6: Condition for a global simulation $R$.

**Definition 88.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and denote the corresponding global-transition relation by $\rightrightarrows$. A relation $R \subseteq Sta \times Sta$ is called a *global simulation* if for every probability measures $\mu_a, \mu_b \in Prob[Sta]$ and $\varphi_a \in Prob[Act \times Sta]$ such that

$$\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}} \quad \text{and} \quad \mu_a \rightrightarrows \varphi_a,$$

there exists $\varphi_b \in Prob[Act \times Sta]$ with

$$\mu_b \rightrightarrows \varphi_b \quad \text{and} \quad \langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}.$$

A relation $R \subseteq Sta \times Sta$ is said to be a *global bisimulation* if both relations $R$ and its inverse relation $R^{-1}$ are combined simulations. ⌟

Figure 4.6 illustrates the condition for a global simulation. Conceptually, Definition 9.8 in [Cat05] and Definition 88 are the same, however, while we rely on the weight lifting of relations, [Cat05] focuses on an alternative lifting that is precised in the second part of this section. The classical notions of simulation and bisimulation introduced in Section 3.2 are local in the sense that only individual states and their outgoing transitions are considered. Compared with this, global simulation and bisimulation provide a global view on the branching-time structure of an STS such that measurability aspects are taken into account (see also the discussions in Chapter 9 in [Cat05]).

The following notions for an STS $\mathcal{T} = (Sta, Act, \rightarrow)$ are routine. A global simulation $R$ is called *Souslin* if the set $R$ is Souslin in $Sta \times Sta$. Analogously, we refer to a global bisimulation $R$ as *Souslin* provided the set $R$ is Souslin in $Sta \times Sta$. The *Souslin-global-simulation preorder* $\preceq^{\mathrm{sou,g}}$ and the *Souslin-global-bisimulation equivalence* $\simeq^{\mathrm{sou,g}}$ are defined in a natural fashion, e.g., for every states $s_a, s_b \in Sta$,

$$\begin{aligned} s_a \preceq^{\mathrm{sou,g}} s_b \quad &\text{iff} \quad \text{there is a Souslin global simulation } R \text{ with } \langle s_a, s_b \rangle \in R, \\ s_a \simeq^{\mathrm{sou,g}} s_b \quad &\text{iff} \quad \text{there is a Souslin glob bisimulation } R \text{ with } \langle s_a, s_b \rangle \in R. \end{aligned}$$

Relying on Lemma 56 as well as Theorem 58 (2), it is easy to see that the relations $\preceq^{\mathrm{sou,g}}$ and $\simeq^{\mathrm{sou,g}}$ form a preorder and an equivalence on the state space *Sta*, respectively (see also the proof of Theorem 55). In what follows we relate the introduced global notions for simulation and bisimulation with combined simulation and bisimulation investigated in Section 4.3 (see Definition 73).

**Example 89.** Considering states $s_a$ and $s_b$ of an STS, it is easy to see that the following two statements hold:

(1) $s_a \preceq^{\text{sou,g}} s_b \quad \text{implies} \quad s_a \preceq^{\text{sou,c}} s_b$.

(2) $s_a \simeq^{\text{sou,g}} s_b \quad \text{implies} \quad s_a \simeq^{\text{sou,c}} s_b$.

This follows directly from the fact that for every transition $s \to \varphi$ in the STS it holds $Dirac[s] \rightrightarrows \varphi$ where $\rightrightarrows$ denotes the corresponding global-transition relation.

However, the reverse implications of (1) and (2) do not hold in general. To see this, consider the STS $\mathcal{T}$ from Example 84 (see also Figure 4.4). Denote the corresponding state and action space by *Sta* and *Act*, respectively. Moreover, let $\rightrightarrows$ be the associated global-transition relation. Relying on the same argument as in Example 63, there exists no $\varphi \in Prob[Act \times Sta]$ such that $Leb \rightrightarrows \varphi$. This implies that neither $s_a \preceq^{\text{sou,g}} s_b$ nor $s_a \simeq^{\text{sou,g}} s_b$ hold. However, it is easy to see that $s_a \simeq^{\text{sou,c}} s_b$. ⌟

According to Example 66, the STS for the counterexample in the previous example in not Souslin. In fact, if one focuses on Souslin STSs, the notions of Souslin global (bi)simulation and Souslin combined (bi)simulation are the same:

**Theorem 90.** *Let $\mathcal{T} = (Sta, Act, \to)$ be a Souslin STS. For every states $s_a, s_b \in Sta$ the two statements below hold:*

(1) $s_a \preceq^{\text{sou,g}} s_b \quad \textit{iff} \quad s_a \preceq^{\text{sou,c}} s_b$.

(2) $s_a \simeq^{\text{sou,g}} s_b \quad \textit{iff} \quad s_a \simeq^{\text{sou,c}} s_b$.

*Moreover, it even holds that every Souslin combined simulation is a Souslin global simulation and accordingly, that every Souslin combined bisimulation is a Souslin global bisimulation.* ⌟

If one takes the setting of this chapter into account, the involved Souslin requirements of the previous result are as expected. Relying on exactly the same argument as for Corollary 78 (or more precisely, Theorem 77), one obtains a proof for Theorem 90 that, in particular, uses a measurable-selection principle. Nevertheless, at the end of this section, we provide some more details on an argument. According to Example 89, the previous theorem does not hold if one drops the Souslin assumption for the involved STS. Theorem 90 also serves as an answer to concluding questions in Section 9.5.4 in [Cat05] where one asks, besides others, under which conditions the combined-bisimulation relation and the global-bisimulation relation are the same.

**Corollary 91.** *Let $\mathcal{T} = (Sta, Act, \to)$ be a Souslin STS. For every states $s_a, s_b \in Sta$ the two statements below hold:*

(1) $s_a \preceq^{\mathrm{sou,g}} s_b$ *implies* $s_a \leq^{\mathrm{tr}} s_b$.

(2) $s_a \simeq^{\mathrm{sou,g}} s_b$ *implies* $s_a =^{\mathrm{tr}} s_b$.

*Proof.* The claim follows directly from Corollary 83 and Theorem 90. □

**Discussion on Cattani's result.** In what follows we provide a detailed comparison of Theorem 82 as well as Corollaries 83 and 91 with a main result of Cattani's thesis [Cat05]. For this let us first remark that in [Cat05] two probability measures $\mu_a, \mu_b \in Prob[Sta]$ are related in terms of the relation $R$ if for every $R$-stable Borel $S \subseteq Sta$ it holds $\mu_a(S) = \mu_b(S)$. According to Theorem 40, the this notion and the weight lifting are the same for smooth relations. This insight is important in regards to the following discussion.

**Proposition 92** (Theorem 9.19 in [Cat05])**.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a simple STS. Then for every states $s_a, s_b \in Sta$ the following implication holds: if there is a global bisimulation $R$ with $\langle s_a, s_b \rangle \in R$ and such that the relation $R$ is smooth, then it holds $s_a =^{\mathrm{tr}} s_b$.* ⌟

To be more precise, the given proof of Theorem 9.19 in [Cat05] also works if one assumes that certain regular conditional probabilities with respect to the quotient functions of a given (not necessarily smooth) bisimulation exist. However, in applications and for concrete examples it is difficult to determine whether this requirement is indeed fulfilled. That is why one typically considers smooth relations where this condition is automatically fulfilled (see Theorem 2.12 in [Cat05] and Theorem 7.11 in [Pan09] as well as the discussions on this topic in Section 7.5 in [Pan09]). Besides this, Theorem 9.19 in [Cat05] in fact considers a weak notion of global equivalence. Weak equivalences abstract from internal steps supposed to be not observable from an external agent [SL94, BH97, PLS00, DGJP10, CSKN05, DH13b]. However, this does not affect the discussions on measurability aspects for schedulers assuming every action is observable.

Conceptually, Corollary 91 and Proposition 92 are in the same spirit. Indeed, both results focus on the generic model of STSs and on a global notion for the comparison of the branching-time structure of states. However, the results differ concerning their side constraints. While Proposition 92 holds for arbitrary STSs, Corollary 91 requires a Souslin STS. The smoothness requirement on the global bisimulation in Proposition 92 is stronger than the corresponding Souslin assumption in Corollary 91 (see also Remark 38). Furthermore, Proposition 92 provides no corresponding result for global simulations and is restricted to simple STSs.

Besides the latter mentioned facts, the main ideas of the proofs in this thesis and in [Cat05] completely differ. Whereas [Cat05] provides a construction of schedulers based on quotient spaces and associated quotient functions, we conservatively extend the classical approach for non-stochastic discrete systems using measurable-selection principles. It is moreover not clear whether the technique in [Cat05] can be extended for the simulation preorder as for preorders there is no natural notion concerning quotient spaces and corresponding quotient functions.

Taking Cattani's results into account, the precise new insights of this chapter for the class of all Souslin STS can be summarised as follows. First of all, in Theorem 90 we establish a connection between global simulation and combined simulation as well as for the corresponding notions for bisimulation. Taking Theorem 90 and the result in Cattani's thesis, i.e., Proposition 92, we derive that for every states $s_a$ and $s_b$ of a simple Souslin STS the following implication holds: if there is a combined bisimulation $R$ with $\langle s_a, s_b \rangle \in R$ and such that the relation $R$ is smooth, then it holds $s_a =^{\mathrm{tr}} s_b$. Thus, in Theorem 82 we moreover show that the premise of the latter implication can be even relaxed by the requirement $s_a \simeq^{\mathrm{sou,c}} s_b$ (see also Theorem 70). Viewing our results from this perspective, Corollary 83 finally states that the condition $s_a \simeq^{\mathrm{sou,c}} s_b$ can be replaced by the statement $s_a \simeq^{\mathrm{sou}} s_b$ considering the ordinary notion of simulation.

Besides the latter mentioned facts, compared to [Cat05], Theorem 82 and Corollary 83 also provide corresponding result for simulations.

**Proof of section's main result.** Our argument for Theorem 90 does not need any new insight. In fact, the proof constitutes a simplification of the argument for Theorem 77. It basically remains to show the following lemma:

**Lemma 93.** *Let $X$ and $Y$ be Polish spaces, $R \subseteq X \times X$ be a relation, and $f \colon X \to Y$ be a Borel function. Define the relation $R_f \subseteq Y \times Y$ by*

$$R_f = \{\langle f(x_a), f(x_b)\rangle \ ; \ \langle x_a, x_b \rangle \in R\}.$$

*For every probability measures $\mu_a, \mu_b \in \mathit{Prob}[X]$ the following implication holds:*

$$\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}} \quad \textit{implies} \quad \langle f_\sharp(\mu_a), f_\sharp(\mu_b)\rangle \in (R_f)^{\mathrm{wgt}}.$$

*Proof.* Define the Borel function $g \colon X \times X \to Y \times Y$, $g(x_a, x_b) = \langle f(x_a), f(x_b)\rangle$. Considering a weight function $W$ for $(\mu_a, R, \mu_b)$, we introduce $W' \in \mathit{Prob}[Y \times Y]$ by $W' = g_\sharp(W)$. Then $W'$ is a weight function for $(f_\sharp(\mu_a), R_f, f_\sharp(\mu_b))$. Let us see why. First of all, it is easy to see that $W'$ is a coupling for $(f_\sharp(\mu_a), f_\sharp(\mu_b))$. Let $R' \subseteq X \times X$ be a Borel set such that $W(R') = 1$ and $R' \subseteq R$. According to Lemma 13, there is a Borel set $R'_f \subseteq Y \times Y$ with $W'(R'_f) = 1$ and $R'_f \subseteq g(R')$. Since we have $R'_f \subseteq g(R') \subseteq R_f$, we can finally conclude that the probability measure $W'$ is a weight function for $(f_\sharp(\mu_a), R_f, f_\sharp(\mu_b))$. $\square$

*Proof of Theorem 90.* We provide a proof for claim (1). as the argument for (2) is completely analogous. First of all, the implication stating that $s_a \preceq^{\text{sou,g}} s_b$ implies $s_a \preceq^{\text{sou,c}} s_b$ for every states $s_a, s_b \in Sta$ follows directly from the definitions. We consider the reverse direction now. For this let $R$ be a Souslin combined simulation. In the remainder of this proof, we show that $R$ is a global simulation that finally yields a proof for claim (1). Let $\mu_a, \mu_b \in Prob[Sta]$ and $\varphi_a \in Prob[Act \times Sta]$ be such that

$$\langle \mu_a, \mu_b \rangle \in R^{\text{wgt}} \quad \text{and} \quad \mu_a \rightrightarrows \varphi_a.$$

Suppose that $(\tilde{\mathfrak{S}}_a, S_a)$ is a witness for $\mu_a \rightrightarrows \varphi_a$. The argument for Theorem 77 can be easily simplified to obtain a probability measure $\varphi_b \in Prob[Act \times Sta]$ such that $\mu_b \rightrightarrows \varphi_b$ and $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\text{wgt}}$. More precisely, inspecting this proof again, there exists a Borel function $\tilde{\mathfrak{S}}_b \colon Sta \to Prob[Act \times Sta]$ and a Borel set $S_b \subseteq Sta$ such that

$$\mu_b(S_b) = 1, \quad \text{and} \quad s_b \rightrightarrows \tilde{\mathfrak{S}}_b(s_b) \text{ for all } s_b \in S_b$$

as well as

$$\langle \mu_a \rtimes \tilde{\mathfrak{S}}_a, \mu_b \rtimes \tilde{\mathfrak{S}}_b \rangle \in (R^{\text{path},1})^{\text{wgt}}.$$

Defining the probability measure $\varphi_b \in Prob[Act \times Sta]$,

$$\varphi_b = Post[\mu_b, \tilde{\mathfrak{S}}_b],$$

one easily derives that $\mu_b \rightrightarrows \mathfrak{S}_b$. Here, $(\tilde{\mathfrak{S}}_b, S_b)$ is a witness for $\mu_b \rightrightarrows \mathfrak{S}_b$. It remains to show $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\text{wgt}}$. To this end define the Borel function $\xi \colon Sta \times Act \times Sta \to Act \times Sta$,

$$\xi(s, act, s') = \langle act, s' \rangle.$$

Then one easily justifies that

$$\xi_\sharp(\mu_a \rtimes \tilde{\mathfrak{S}}_a) = \varphi_a \quad \text{and} \quad \xi_\sharp(\mu_b \rtimes \tilde{\mathfrak{S}}_b) = \varphi_b$$

as well as

$$R^{Act} = \{\langle \xi(\hat{\pi}_a), \xi(\hat{\pi}_b) \rangle \ ; \ \langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R^{\text{path},1}\}.$$

Thanks to Lemma 93, we finally obtain $\langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\text{wgt}}$. $\square$

## 4.8 Deterministic purely stochastic systems

It can be regarded as folklore that for non-stochastic systems whose observable branching-time structure is deterministic the trace relation and the bisimulation relation are the same, i.e., for every states $s_a$ and $s_b$ one has $s_a \simeq s_b$ iff $s_a =^{\mathrm{tr}} s_b$. The motivation of this section is to provide a corresponding result for the setting of this thesis. For this purpose we introduce purely stochastic STSs first. After that, we define the property of being deterministic in the context of purely stochastic STSs. It is then shown that for every deterministic purely stochastic Souslin STS the relations induced by bisimulation, Souslin bisimulation, and trace distributions are the same.

**Definition 94.** An STS $\mathcal{T} = (Sta, Act, \rightarrow)$ is called *purely stochastic* provided there exists a function (*control law*)

$$\mathfrak{K}\colon Sta \rightarrow Prob[Act \times Sta]$$

such that for every state $s \in Sta$ and probability measure $\varphi \in Prob[Act \times Sta]$,

$$s \rightarrow \varphi \quad \text{iff} \quad \mathfrak{K}(s) = \varphi,$$

i.e., the transition relation $\rightarrow$ and the graph of the function $\mathfrak{K}$ are the same. ⌟

In other words, an STS is purely stochastic precisely when the set $Enabled[s]$ is a singleton for every state $s$. If an STS is purely stochastic, the corresponding control law is uniquely determined. Vice versa, the control law of a purely stochastic STS uniquely determines the associated transition relation. To simplify notions, we can hence safely refer to a tuple $(Sta, Act, \mathfrak{K})$ including the control law $\mathfrak{K}$ rather than the transition relation $\rightarrow$ as a purely stochastic STS where the symbols are given as before.

**Remark 95.** Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a purely stochastic STS. Then the following two statements are equivalent:

(1) The control law $\mathfrak{K}$ is a Borel function.

(2) The STS $\mathcal{T}$ is Souslin.

The claimed equivalence follows directly from Remark 10 (6). ⌟

The definition of a purely stochastic STS involves no measurability requirement concerning the control law. The previous remark shows that the class of all Souslin STSs covers precisely those purely stochastic STSs whose control law constitutes a Borel function. In the context of stochastic processes, the Borel assumption on the control law is standard. Thus, the Souslin property for purely stochastic STSs has a natural characterisation that also provides confidence for the applicability of the key concepts of this chapter. Modelling formalisms from the literature that are covered by purely stochastic Souslin STSs are presented in Chapter 7.2.

**Remark 96.** Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a purely stochastic Souslin STS and $\mu \in Prob[Sta]$ be a probability measure intuitively serving as an initial distribution. According to Theorem 65, there exists a $\mu$-scheduler for $\mathcal{T}$. Since the set $Enabled[s]$ is a singleton for every state $s \in Sta$, there exists exactly one $\mu$-scheduler, say $\mathfrak{S}_\mu$. Let $n \in \mathbb{N}$ be a natural number. For every finite path $\hat{\pi} \in Path_n$ it obviously holds

$$\mathfrak{S}_\mu(\hat{\pi}) = \mathfrak{K}(\mathrm{Last}_n(\hat{\pi})).$$

Consequently, for every Borel sets $\hat{\Pi}_n \subseteq Path_n$ and $B \subseteq Act \times Sta$ we obtain

$$Pr_{n+1}[\mathfrak{S}_\mu](\hat{\Pi}_n \times B) = \int_{\hat{\Pi}_n} \mathfrak{K}(\mathrm{Last}_n(\hat{\pi}))(B)\, dPr_n[\mathfrak{S}_\mu](\hat{\pi}),$$

i.e., the probability measure $Pr_{n+1}[\mathfrak{S}_\mu]$ is uniquely determined by the probability measure $Pr_n[\mathfrak{S}_\mu]$ as well as the the control law $\mathfrak{K}$. ⌟

According to Remark 96, the concept of schedulers is not needed for the definition of path measures and trace distributions of purely stochastic Souslin STS. For that reason we simplify notions as follows. Consider a purely stochastic Souslin STS with state space $Sta$. For every $\mu \in Prob[Sta]$ we write $Pr[\mu]$ rather than $Pr[\mathfrak{S}_\mu]$. For every $n \in \mathbb{N}$ we use $Pr_n[\mu]$ instead of $Pr_n[\mathfrak{S}_\mu]$. The notion $PrTrace[\mu]$ referring to the respective trace distribution is used analogously. For every state $s \in Sta$ we moreover use $Pr[s]$ rather than $Pr[Dirac[s]]$ and accordingly for finite-path measures and trace distributions.

**Remark 97.** Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a purely stochastic STS and $R \subseteq Sta \times Sta$ be a relation. Then the following three statements are equivalent:

(1) $R$ is a bisimulation.

(2) $R$ is a simulation.

(3) For every states $s_a, s_b \in Sta$,

$$\langle s_a, s_b \rangle \in R \quad \text{implies} \quad \langle \mathfrak{K}(s_a), \mathfrak{K}(s_b) \rangle \in (R^{Act})^{\mathrm{wgt}}.$$

This equivalence follows directly from the definitions. ⌟

The notions of simulation and bisimulation are the same for the class of purely stochastic STSs. In this chapter all the results are written from the bisimulation perspective, however, every result can be also formulated from the simulation point of view by Remark 97. Besides this observation, the combined-transition relation of a purely stochastic STS collapses with the transition relation and hence, the notion of bisimulation and combined bisimulation are also the same.

We first thought that the bisimulation relation and the trace-distribution relation are the same for purely stochastic Souslin STS. However, our initial idea is false as the following (standard) example illustrates:

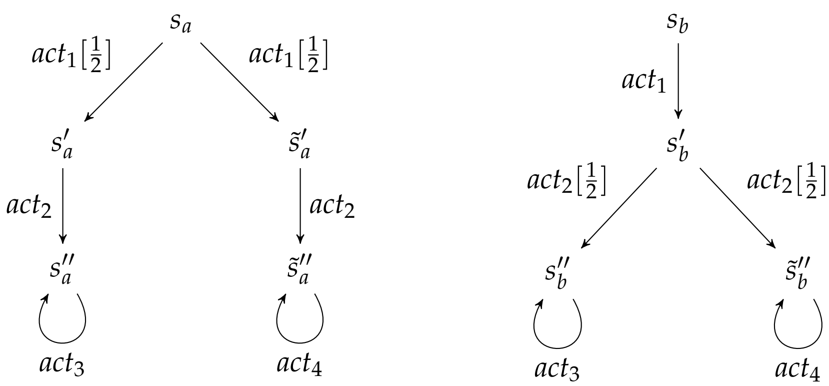


Figure 4.7: It holds $s_a =^{\mathrm{tr}} s_b$ but not $s_a \simeq s_b$.

**Example 98.** Consider the purely stochastic Souslin STS $\mathcal{T}$ illustrated by Figure 4.7. Inspecting the states $s_a$ and $s_b$, it is easy to see that $s_a =^{\mathrm{tr}} s_b$. Indeed, the uniquely determined trace distribution in these states is given by the probability measure on $Act^{\omega}$ that assigns probability $1/2$ to the traces $act_1\, act_2\, act_3^{\omega}$ and $act_1\, act_2\, act_4^{\omega}$, respectively. However, we have that $s_a$ and $s_b$ are not bisimilar to each other. ⌟

The purely stochastic STS depicted in Figure 4.7 (see also Example 98) is not deterministic from the action-based point of view on stochastic systems. Indeed, both states $s'_a$ and $\tilde{s}'_a$ can be reached from state $s_a$ by taking the same action $act_1$ and accordingly, the execution of the action $act_2$ in state $s'_b$ does not lead to an uniquely determined successor state. This observation motivates to the following definition:

**Definition 99.** A purely stochastic STS $\mathcal{T} = (Sta, Act, \mathfrak{K})$ is called *deterministic* provided for every state $s \in Sta$ there exists a Borel function $f_s \colon Act \to Sta$ such that

$$\mathfrak{K}(s)(Graph[f_s]) = 1.$$

In this context, we refer to the function $f_s$ as a *witness for the determinism in* $s$. ⌟

By Remark 10 (6), the graph of the Borel function $f_s$ is indeed Borel in $Act \times Sta$ and thus, we can safely consider its probability mass concerning the probability measure $\mathfrak{K}(s)$ in the previous definition. The intuitive role of the function $f_s$ is as follows: when the action $act$ is executed in the state $s$, the successor state is almost surely determined by the state $f_s(act)$. Assuming the sets $Sta$ and $Act$ are countable, it is easy to see that the purely stochastic STS $\mathcal{T}$ is deterministic precisely when for every states $s, s', \tilde{s}' \in Sta$ and action $act \in Act$ the following implication holds:

$$\mathfrak{K}(s)(\{act\} \times \{s'\}) > 0 \text{ and } \mathfrak{K}(s)(\{act\} \times \{\tilde{s}'\}) > 0 \quad \text{implies} \quad s' = \tilde{s}'.$$

This observation immediately justifies that the STS in Figure 4.7 is not deterministic.

**Theorem 100.** *Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a deterministic purely stochastic Souslin STS. For every states $s_a, s_b \in Sta$ the following equivalence holds:*

$$s_a \simeq^{\text{sou}} s_b \quad \textit{iff} \quad s_a =^{\text{tr}} s_b.$$

*In particular, the relation $=^{\text{tr}}$ is a Souslin bisimulation.* ⌟

As a consequence, establishing Souslin-bisimulation equivalence is sound and complete for proving trace-distribution equivalence of states for the subclass of STSs under consideration. Relying on Corollary 83, it suffices to show that the relation $=^{\text{tr}}$ is a Souslin bisimulation. Our proof for this is technical in some points, however, relies only on basic concepts of measure theory. Interestingly, the following argument uses the characterisation of the weight lifting of an action lifting from Section 3.3. Indeed, this is the only place in this thesis where we exploit the two-step view on distributions over action-state pairs discussed in Section 3.3 in a mathematical argument.

**Proof of section's main result.** Our argumentation for Theorem 100 uses the following additional notations. Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a purely stochastic Souslin STS. The function $\mathfrak{K}_{Act} \colon Sta \to Prob[Act]$ results from the control law $\mathfrak{K}$ as follows: for every state $s \in Sta$ and Borel set $A \subseteq Act$ let

$$\mathfrak{K}_{Act}(s)(A) = \mathfrak{K}(s)(A \times Sta),$$

i.e., $\mathfrak{K}_{Act}(s)$ is obtained by projecting the probability measure $\mathfrak{K}(s)$ onto the action space *Act*. Obviously, the function $\mathfrak{K}_{Act}$ is Borel since the control law $\mathfrak{K}$ forms a Borel function by Remark 95. We additionally define the function $\mathit{PrTraceFunc}\colon \mathit{Sta} \to \mathit{Prob}[\mathit{Act}^\omega]$,

$$\mathit{PrTraceFunc}(s) = \mathit{PrTrace}[s].$$

Hence, *PrTraceFunc* maps every state to its uniquely determined trace distribution (see also Remark 96). This function is needed for a careful treatment of measurability issues. For instance, to ensure the well-definedness of integrals over specific trace distributions, we are required to show that the function *PrTraceFunc* is Borel. In this context, the following lemma can be seen as folklore:

**Lemma 101.** *Let $\mathcal{T} = (\mathit{Sta}, \mathit{Act}, \mathfrak{K})$ be a purely stochastic Souslin STS. The function PrTraceFunc is Borel and moreover, for every state $s \in \mathit{Sta}$ and Borel sets $A \subseteq \mathit{Act}$ and $B \subseteq \mathit{Act}^\omega$ it holds*

$$\mathit{PrTraceFunc}(s)(A \times B) = \int_{A \times \mathit{Sta}} \mathit{PrTraceFunc}(s')(B)\, d\mathfrak{K}(s)(\mathit{act}, s')$$

*Proof.* We rely on a standard argumentation scheme from measure theory. Pick a Borel sets $A \subseteq \mathit{Act}$. For every Borel set $B \subseteq \mathit{Act}^\omega$ define the function $f_B\colon \mathit{Sta} \to [0, 1]$,

$$f_B(s) = \mathit{PrTraceFunc}(s)(B).$$

Let $\mathcal{M}$ be the family consisting of all Borel subsets $B \subseteq \mathit{Act}^\omega$ with the following property: the function $f_B$ is Borel and for every state $s \in \mathit{Sta}$ one has the identity

$$\mathit{PrTraceFunc}(s)(A \times B) = \int_{A \times \mathit{Sta}} f_B(s')\, d\mathfrak{K}(s)(\mathit{act}, s').$$

We show that $\mathcal{M}$ and the Borel sigma algebra on the Polish space $\mathit{Act}^\omega$ are the same. By Remark 5, it then follows that the function *PrTraceFunc* is Borel and by Carathéodory uniqueness theorem (see Section 2.1), the claimed identity then follows.

The family of sets $\mathcal{M}$ is a monotone class, i.e., the family $\mathcal{M}$ is closed under unions of increasing chains and under intersections of decreasing chains. This follows from the following three facts: every measure is continuous from below and above, the point-wise limit of a convergence sequence of Borel functions is Borel, and the monotone convergence theorem (also known as Beppo Levi's theorem, see Satz 8.2.1 in [Sch08]).

Define $\mathcal{G}_0 = \{\mathit{Act}^\omega\}$ and moreover, for every $n \in \mathbb{N} \setminus \{0\}$ let

$$\mathcal{G}_n = \{A_1 \times \ldots \times A_n \times \mathit{Act}^\omega \ ; \ A_1, \ldots, A_n \subseteq \mathit{Act} \text{ are Borel sets}\}.$$

Relying on an induction over the natural numbers, we show that for every $n \in \mathbb{N}$ it holds $\mathcal{G}_n \subseteq \mathcal{M}$. For every state $s \in Sta$ it holds $f_{Act^\omega}(s) = 1$ as well as

$$PrTraceFunc(s)(A \times Act^\omega) = \mathfrak{K}_{Act}(s)(A) = \int_{A \times Sta} 1 \, d\mathfrak{K}(s)(act, s')$$

We hence obtain $\mathcal{G}_0 \subseteq \mathcal{M}$. For every $n \in \mathbb{N}$, if $\mathcal{G}_n \subseteq \mathcal{M}$, then Fubini's theorem (see Section 2.1) yields the inclusion $\mathcal{G}_{n+1} \subseteq \mathcal{M}$.

The union $\mathcal{G}_1 \cup \mathcal{G}_2 \cup \ldots$ generates the Borel sigma algebra on $Act^\omega$ (see Remark 3 (5) as well as Section 3.5 in [Bog07]). Thus, putting things together, we can apply the monotone class theorem (see Theorem 136B in [Fre01]) and therefore, the family $\mathcal{M}$ and the Borel sigma algebra on $Act^\omega$ are the same. □

Whereas Lemma 101 holds for arbitrary purely stochastic Souslin STSs, the following representation of the function *PrTraceFunc* is adapted for deterministic systems:

**Lemma 102.** *Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a deterministic purely stochastic Souslin STS, $s \in Sta$ be a state, and $f$ be a witness of the determinism in $s$. For every Borel set $B \subseteq Act^\omega$ define the function $g_B \colon Act \to [0,1]$,*

$$g_B(act) = PrTraceFunc(f(act))(B).$$

*For every Borel set $B \subseteq Act^\omega$ the function $g_B$ is Borel and moreover, for all Borel sets $A \subseteq Act$,*

$$PrTraceFunc(s)(A \times B) = \int_A g_B(act) \, d\mathfrak{K}_{Act}(s)(act).$$

*Proof.* Let $B \subseteq Act^\omega$ be a Borel set. Remark 5 immediately proves the first claim stating that the function $g_B$ is Borel. It remains to show the claimed identity. For this purpose define the function $h_B \colon Sta \to [0,1]$,

$$h_B(s') = PrTraceFunc(s')(B).$$

Using the same argument as for the function $g_B$, the function $h_B$ is also Borel. For every pair $\langle act, s' \rangle \in Graph[f]$ it moreover holds $f(act) = s'$ and therefore, $g_B(act) = h_B(s')$. Relying on the identity $\varphi(Graph[f]) = 1$, we hence obtain

$$\begin{aligned}
& \int_A g_B(act) \, d\mathfrak{K}_{Act}(s)(act) \\
= & \int_{A \times Sta} g_B(act) \, d\mathfrak{K}(s)(act, s') \\
= & \int_{A \times Sta} h_B(s') \, d\mathfrak{K}(s)(act, s')
\end{aligned}$$

$$= \quad PrTraceFunc(s)(A \times B)$$

where the latter equality follows from Lemma 101. □

**Lemma 103.** *Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a deterministic purely stochastic Souslin STS, $s_a, s_b \in Sta$ be states, as well as $f_a$ and $f_b$ be witnesses of the determinism in $s_a$ and $s_b$, respectively. Assume $s_a =^{\mathrm{tr}} s_b$. Then the following two statements hold:*

(1) $\mathfrak{K}_{Act}(s_a) = \mathfrak{K}_{Act}(s_b)$.

(2) *There exists a Borel set $A \subseteq Act$ such that $\mathfrak{K}_{Act}(s_a)(A) = 1$ and for all $act \in A$ it holds*

$$PrTraceFunc(f_a(act)) = PrTraceFunc(f_a(act)).$$

*Proof.* For every state $s \in Sta$ and Borel set $A \subseteq Act^{\omega}$ it holds

$$\mathfrak{K}_{Act}(s)(A) = PrTraceFunc(s)(A \times Act^{\omega}).$$

This directly yields claim (1) since the identity $PrTraceFunc(s_a) = PrTraceFunc(s_b)$ is implied by the assumption $s_a =^{\mathrm{tr}} s_b$ (see also Remark 96). The remainder of this proof is devoted to claim (2). Define the probability measure $\alpha \in Prob[Act]$,

$$\alpha = \mathfrak{K}_{Act}(s_a).$$

Clearly, it also holds $\alpha = \mathfrak{K}_{Act}(s_b)$. Let $\mathcal{C}$ be a countable family of Borel sets in $Act^{\omega}$ that generates the Borel sigma-algebra on $Act^{\omega}$, and moreover, that is closed under finite intersections (see Remark 4). Suppose Borel sets $B_0, B_1, B_2, \ldots \subseteq Act^{\omega}$ such that $\mathcal{C} = \{B_0, B_1, B_2, \ldots\}$. For every $n \in \mathbb{N}$ introduce the functions $g_{a,n} \colon Act \to [0,1]$ and $g_{b,n} \colon Act \to [0,1]$ as follows: for every $act \in Act$ let

$$g_{a,n}(act) = PrTraceFunc(f_a(act))(B_n)$$

as well as

$$g_{b,n}(act) = PrTraceFunc(f_b(act))(B_n).$$

For every $n \in \mathbb{N}$, thanks to Lemma 102, both functions $g_{a,n}$ and $g_{b,n}$ are Borel and moreover, for all Borel sets $A \subseteq Act$ it holds

$$PrTraceFunc(s_a)(A \times B_n) = \int_A g_{a,n}(act)\, d\alpha(act)$$

and analogously,

$$PrTraceFunc(s_b)(A \times B_n) = \int_A g_{b,n}(act)\, d\alpha(act).$$

Using again the identity $PrTraceFunc(s_a) = PrTraceFunc(s_b)$, for every $n \in \mathbb{N}$ and every Borel set $A \subseteq Act$ it follows

$$\int_A g_{a,n}(act)\, d\alpha(act) = \int_A g_{b,n}(act)\, d\alpha(act).$$

According to a standard result from measure theory (see, e.g., Folgerung 9.2.8 in [Sch08]), for every $n \in \mathbb{N}$ there exists a Borel set $A_n \subseteq Act$ such that

$$\alpha(A_n) = 1 \quad \text{and} \quad g_{a,n}(act) = g_{b,n}(act) \text{ for all } act \in A_n.$$

Define the subset $A_{\mathcal{C}}$ of $Act$ by

$$A_{\mathcal{C}} = A_0 \cap A_1 \cap A_2 \cap \dots .$$

Since the family $\mathcal{C}$ is countable, the set $A_{\mathcal{C}}$ is Borel in $Act$ and it holds $\alpha(A_{\mathcal{C}}) = 1$. For every $act \in A_{\mathcal{C}}$ the probability measures $PrTraceFunc(f_a(act))$ and $PrTraceFunc(f_b(act))$ agree on $\mathcal{C}$, i.e., for every $n \in \mathbb{N}$ one has

$$PrTraceFunc(f_a(act))(B_n) = PrTraceFunc(f_b(act))(B_n).$$

Since the family $\mathcal{C}$ generates the Borel sigma-algebra on $Act^\omega$ and is also closed under finite intersection, for every $act \in A_{\mathcal{C}}$ it follows

$$PrTraceFunc(f_a(act)) = PrTraceFunc(f_b(act))$$

by Carathéodory extension theorem (see Section 2.1). This completes our proof. □

**Lemma 104.** *Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a deterministic purely stochastic STS, $s \in Sta$ be a state, and $f_s$ be a witness of the determinism in $s$. Define the function $f'_s \colon Act \to Prob[Sta]$,*

$$f'_s(act) = Dirac[f_s(act)].$$

*Then the following identity holds*

$$\mathfrak{K}(s) = \mathfrak{K}_{Act}(s) \rtimes f'_s.$$

*Proof.* Note, by Example 6, the function $f'_s$ is indeed Borel. Let $A \subseteq Act$ and $S \subseteq Sta$ be Borel sets. It holds

$$\mathfrak{K}_{Act}(s) \rtimes f'_s(A \times S) = \int_A Dirac[f_s(act)](S)\, d\mathfrak{K}_{Act}(s)(act) = \mathfrak{K}_{Act}(s)(A \cap f_s^{-1}(S)).$$

Using the identity

$$((A \cap f_s^{-1}(S)) \times Sta) \cap Graph[f_s] = (A \times S) \cap Graph[f_s],$$

together with $\mathfrak{K}(s)(Graph[f_s]) = 1$, we obtain

$$\mathfrak{K}_{Act}(s)(A \cap f_s^{-1}(S)) = \mathfrak{K}(s)((A \cap f_s^{-1}(S)) \times Sta) = \mathfrak{K}(s)(A \times S).$$

Carathéodory extension theorem (see Section 2.1) finally yields $\mathfrak{K}_{Act}(s) \rtimes f'_s = \mathfrak{K}(s)$. $\square$

*Proof of Theorem 100.* Using Remarks 37 and 96, it is easy to see that the relation $=^{\mathrm{tr}}$ is smooth. In fact, for every states $s_a, s_b \in Sta$ it holds

$$s_a =^{\mathrm{tr}} s_b \quad \text{iff} \quad PrTraceFunc(s_a) = PrTraceFunc(s_b).$$

Recall, the function *PrTraceFunc* is Borel by Lemma 101. Thanks to Remark 38, the set $=^{\mathrm{tr}}$ is in particular Souslin in $Sta \times Sta$. Hence, thanks to Corollary 41, for every probability measures $\mu_a, \mu_b \in Prob[Sta]$ we have

$$\langle \mu_a, \mu_b \rangle \in (=^{\mathrm{tr}})^{\mathrm{wgt}} \quad \text{iff} \quad PrTraceFunc_\sharp(\mu_a) = PrTraceFunc_\sharp(\mu_b).$$

Let $s_a, s_b \in Sta$ be states with $s_a =^{\mathrm{tr}} s_b$. To obtain a proof of the theorem, remembering Remark 97, it suffices to show $\langle \mathfrak{K}(s_a), \mathfrak{K}(s_b) \rangle \in ((=^{\mathrm{tr}})^{Act})^{\mathrm{wgt}}$. We rely on Theorem 52 and hence, it suffices to show $\langle \mathfrak{K}(s_a), \mathfrak{K}(s_b) \rangle \in ((=^{\mathrm{tr}})^{Act})^{\mathrm{two}}$.

Suppose witnesses $f_a$ and $f_b$ of the determinism in $s_a$ and $s_b$, respectively. Introduce the two functions $f'_a \colon Act \to Prob[Sta]$ and $f'_b \colon Act \to Prob[Sta]$ as follows: for every action $act \in Act$ let

$$f'_a(act) = Dirac[f_a(act)] \quad \text{and} \quad f'_b(act) = Dirac[f_a(act)].$$

Lemma 104 yields the identities

$$\mathfrak{K}(s_a) = \mathfrak{K}_{Act}(s_a) \rtimes f'_a \quad \text{and} \quad \mathfrak{K}(s_b) = \mathfrak{K}_{Act}(s_b) \rtimes f'_b.$$

Invoking Lemma 103, we obtain $\mathfrak{K}_{Act}(s_a) = \mathfrak{K}_{Act}(s_b)$ and there exists a Borel set $A \subseteq Act$ such that $\mathfrak{K}_{Act}(s_a)(A) = 1$ and so that for every $act \in A$,

$$PrTraceFunc(f_a(act)) = PrTraceFunc(f_a(act)).$$

From this one can easily derive that for every $act \in A$,

$$PrTraceFunc_{\sharp}(f'_a(act)) = PrTraceFunc_{\sharp}(f'_a(act)).$$

For every action $act \in A$, exploiting the initial observation in this proof, we therefore obtain $\langle f'_a(act), f'_a(act)\rangle \in (=^{\mathrm{tr}})^{\mathrm{wgt}}$. This implies $\langle \mathfrak{K}(s_a), \mathfrak{K}(s_b)\rangle \in ((=^{\mathrm{tr}})^{Act})^{\mathrm{two}}$. As already discussed before, the latter insight completes the proof. $\square$

# 5 Action-based probabilistic temporal logics

This chapter is devoted to the proof of Theorem D (see Chapter 1) where simulation preorder and bisimulation equivalence is related to relations induced by temporal logics. For the specification of probabilities for complex path properties the temporal logics APCTL$^*$ and its existential fragment $\exists$APCTL$^*$ yield comparable expressive temporal logics that. As already mentioned in Chapter 1, they combine features of several logics that have been introduced in the literature. The temporal logics APCTL$_\circ$ and APCTL$_\bullet$ form Hennessy-Milner-like sublogics of APCTL$^*$ and $\exists$APCTL$^*$, respectively. These sublogics are only allowed to specify properties on direct successor distributions over action-state pairs.

To the best of our knowledge, Theorem D is the first theorem providing logical characterisations of the simulation preorder and the bisimulation equivalence for a general stochastic model possibly having an uncountable state and action space in terms of a weak and an expressive temporal logic. In view of the developed theory in Chapter 4, we emphasise that Theorem D also identifies a subclass of (Souslin) STSs where Souslin simulation and simulation are the same and accordingly, where the concepts of Souslin bisimulation and bisimulation collapse.

To cover STSs whose action spaces are uncountable, the novelty of our syntax for APCTL$^*$ is that the basic atomic observables are certain subsets of the action space that are specified by an action event family. Recall, an action space *Act* can be viewed as the set of all relevant atomic observables. In particular, every action contained in *Act* stands for a process activity. However, since the set *Act* is not restricted to be countable, it is not plausible to assume that the occurrence of every individual action can be indeed observed, e.g., from an external agent. Intuitively, an action event family $\mathcal{A}$ is a countable family of subsets of *Act* specifying those subsets of the action space for which one can determine whether a given action is contained or not: taking $A \in \mathcal{A}$ and $act \in Act$, one can provide the answer yes if $act \in A$ and no otherwise.

Initially, we interpret APCTL$^*$ over arbitrary STSs augmented with an action event family and a reward function. While the given semantics is mostly standard, the probability modality needs some care due to measurability issues caused by uncountable state and action spaces. In fact, it turns out that specific satisfactions of APCTL$^*$ formulas may not be

Borel. To this end, we rely on the concept of outer-measure functions within the definition of our semantics. Recall that an outer-measure function extends a probability measure defined on some sigma algebra to the whole powerset (see Section 2.1). This approach causes no trouble for theoretical arguments and allows us to extend classical results from the discrete setting to the setting of this thesis.

We present a measurability condition on STSs called: Borel concerning the hit sigma algebra. This notion in particular ensures that the outer-measure semantics and the classical semantics for the two weak logics $\mathrm{APCTL}_{\circ}$ and $\mathrm{APCTL}_{\bullet}$ are the same since corresponding satisfaction sets can be proven to yield Borel sets. The latter fact is crucial for the application of the techniques in [FKP17] concerning our proof of one part of Theorem D. The mentioned measurability requirement on an STS is closely connected to the definition of non-deterministic labelled Markov processes [DTW12, Wol12]. However, there is an seemingly unimportant difference in the corresponding definitions that has crucial consequences for our compositional framework. The motivation of the concepts also differ.

In this chapter we also provide a variant of Theorem D for simple STSs whose action space is countable. Consequently, on the one hand, this chapter's main result covers the established logical characterisations for the simulation preorder and the bisimulation equivalence for labelled Markov processes (LMPs) [FKP17] in terms of corresponding weak modal logics (see also [BDEP97, DEP02, DGJP03, Des99, Pan09]). On the other hand, this chapter extends the theory on LMPs by providing corresponding preservation results for the expressive temporal logics $\mathrm{APCTL}^*$ and $\exists\mathrm{APCTL}^*$. The same discussion applies for non-deterministic labelled Markov processes (NLMPs) [DTW12, Wol12] where, to the best of our knowledge, the existing literature is restricted to bisimulations. Thus, as a byproduct, Theorem D provides a first result referring to the simulation relations of NLMPs.

Theorem D also extends [NK07] by a complete characterisation of both the simulation preorder and the bisimulation equivalence. Note, [NK07] solely provides a preservation result for bisimulations of continuous-time Markov decision processes concerning an expressive temporal logic. The main result of this chapter also covers results in [DP03] and [DGJP10] providing a logical characterisation of the bisimulation equivalence for continuous-time Markov processes and labelled concurrent Markov chains, respectively. We moreover emphasise that the logical characterisations given by Theorem D refer to a generative model in the classification of [GSS95]. In this context, to the best of our knowledge, a generic logical characterisation of the simulation preorder and the bisimulation equivalence has been not proven before. Chapter 7 provides more detailed discussions regarding related models.

## 5.1 Syntax using action event spaces

This section presents the syntax of the action-based probabilistic logic APCTL$^*$ and of the corresponding sublogics $\exists$APCTL$^*$, APCTL$_\circ$, and APCTL$_\bullet$. To this end, we introduce the notion of an action event family:

**Definition 105.** For every Polish space *Act* an *action event family* (*on Act*) is a countable generator $\mathcal{A}$ of the Borel sigma algebra on *Act* that satisfies the following additional property: the family of sets $\mathcal{A}$ separates the points in *Act*, i.e., for every $act_1, act_2 \in Act$,

$$act_1 = act_2 \quad \text{iff} \quad \text{for every } A \in \mathcal{A} \text{ one has } act_1 \in A \text{ iff } act_2 \in A.$$

An *action event space* is a pair $(Act, \mathcal{A})$ such that *Act* is a Polish space and $\mathcal{A}$ is an action event family on *Act*. ⌟

Let $(Act, \mathcal{A})$ be an action event space. The elements of the action event family $\mathcal{A}$ can be viewed as test sets to distinguish two actions $act_1$ and $act_2$ in *Act*. Indeed, $act_1$ is not equal to $act_2$ precisely when there exists a test set $A \in \mathcal{A}$ such that $act_1 \in A$ and $act_2 \notin A$ or the other way round, i.e., $act_1 \notin A$ and $act_2 \in A$. The smallest family of subsets of *Act* that contains every set in $\mathcal{A}$ and that is additionally closed under complementation and finite intersections is denoted by $Bool[\mathcal{A}]$. The family $Bool[\mathcal{A}]$ is hence also closed under finite unions. Moreover, it is easy to see that $Bool[\mathcal{A}]$ also yields an action event family, in particular, constitutes a countable family of subsets of *Act*. For every $A, A_1, A_2 \in Bool[\mathcal{A}]$ we write $\neg A$ rather than $Act \setminus A$, $A_1 \wedge A_2$ instead of $A_1 \cap A_2$. and accordingly for the other boolean connectives (disjunction $\vee$, implication $\rightarrow, \ldots$).

**Example 106.** Based on Examples 2 and 3, we provide examples for action event spaces:

(1) Let $Act = Q$ be a countable set. Define $\mathcal{A}_Q$ as the family of sets consisting of all singletons $\{q\}$ with $q \in Q$. Then it is easy to see that the pair $(Q, \mathcal{A}_Q)$ yields an action event space.

(2) We regard the action space $Act = \mathbb{R}_{\geq 0}$ representing continuous time. Suppose $\mathcal{A}_\mathcal{I}$ is the family consisting of all intervals $I$ in $\mathbb{R}_{\geq 0}$ with rational endpoints, i.e., there are $q_1, q_1 \in \mathbb{Q}_{\geq 0}$ with $I \in \{(q_1, q_2), (q_1, q_2], [q_1, q_2), [q_1, q_2]\}$. Then the pair $(\mathbb{R}_{\geq 0}, \mathcal{A}_\mathcal{I})$ also forms an action event space.

(3) Consider the action space $Act = Eval[Var]$ of all variable evaluations over some finite set of variables *Var*. Let $\mathcal{A}_{Var}$ be the family consisting of all conditions over *Var*, i.e., $\mathcal{A}_{Var} = Cond[Var]$. Again the pair $(Eval[Var], \mathcal{A}_{Var})$ constitutes an action event space (see also Section 3.1).

Clearly, one can also combine the presented action event spaces. Consider for instance a countable set $Q$ that is disjoint from $\mathbb{R}_{\geq 0}$. Then the union $\mathcal{A}_Q \cup \mathcal{A}_\mathcal{I}$ yields an action event family on the action space $Q \cup \mathbb{R}_{\geq 0}$. Moreover, $\{A_Q \times A_\mathcal{I} \,;\, A_Q \in \mathcal{A}_Q \text{ and } A_\mathcal{I} \in \mathcal{A}_\mathcal{I}\}$ is an action event family on the action space $Q \times \mathbb{R}_{\geq 0}$. ⌟

It is important to realise that while an action space may be uncountable an action event family is required to be countable. From a theoretical point of view, according to Remark 4, every Polish space gives rise to an action event space. However, the choice of a concrete action event family can be seen as an engineering task performed while modelling the system under consideration. The previous example illustrates that action event families appear naturally in the modelling of stochastic systems.

The following definition provides the syntax of the action-based probabilistic modal logic APCTL*. The basic atomic building blocks are given by the sets of actions contained in a picked action event family:

**Definition 107.** Let $(Act, \mathcal{A})$ be an action event space. An *APCTL* state formula* (*over* $(Act, \mathcal{A})$) is formed by the following grammar:

$$\sigma \quad ::= \quad \sigma \wedge \sigma \;\mid\; \neg\sigma \;\mid\; \exists\tau$$

where $\tau$ is an APCTL* path-measure formula. Here, an *APCTL* path-measure formula* (*over* $(Act, \mathcal{A})$) is generated by the following grammar:

$$\tau \quad ::= \quad \tau \wedge \tau \;\mid\; \neg\tau \;\mid\; \mathbb{P}_{\bowtie q}[v]$$

where $\bowtie \in \{<, \leq, =, \geq, >\}$, $q \in [0,1] \cap \mathbb{Q}$, and $v$ is a APCTL* path formula. An *APCTL* path formula* (*over* $(Act, \mathcal{A})$) is defined by the following grammar:

$$v \quad ::= \quad \sigma \;\mid\; A \;\mid\; v \wedge v \;\mid\; \neg v \;\mid\; \bigcirc v \;\mid\; v \,\mathrm{U}\, v \;\mid\; Acc[\bowtie r]$$

where $\sigma$ is an APCTL* state formula, $A \in Bool[\mathcal{A}]$, $\bowtie \in \{<, \leq, =, \geq, >\}$, and $r \in \mathbb{Q}$. ⌟

As the names in the previous definition suggest, APCTL* state formulas are interpreted over states of an STS, APCTL* path-measure formulas constrain path measures induced by schedulers, and APCTL* path formulas refer to paths. The precise semantics is presented in the next section, however, let us give a feeling for the meaning of the introduced formulas. Considering a state $s$ of an STS, the exists modality $\exists\tau$ asserts the existence of an $s$-scheduler $\mathfrak{S}$ whose induced path measure $Pr[\mathfrak{S}]$ satisfies the formula $\tau$. For instance, assuming $\tau$ is equal to the probability modality $\mathbb{P}_{>q}[v]$, the path measure $Pr[\mathfrak{S}]$ fulfils $\tau$ if the probability of all paths satisfying the formula $v$ is greater than $q$ (this needs some care due to measurability issues). The next modality $\bigcirc\, v$ and the until modality $v_1 \,\mathrm{U}\, v_2$ have the same meaning as in standard linear temporal logic (see also, e.g., [BK08]). The accumulation modality $Acc[{>}r]$ asserts that the accumulated reward of a path up to the current position is greater than the reward bound $r$.

For each of the three levels of APCTL* the operators from propositional logic (true $tt$, disjunction $\vee$, implication $\rightarrow, \ldots$) are derived as usual. For every APCTL* path-measure formula $\tau$ we moreover define the forall modality

$$\forall\tau = \neg\exists(\neg\tau).$$

The forall modality constitutes an APCTL* state formula constraining every $s$-scheduler of a picked state $s$. The eventually and the always modality are defined as in standard linear temporal logics, i.e., for every APCTL* path formula $v$ we have

$$\Diamond\, v = tt \,\mathrm{U}\, v \quad \text{and} \quad \Box\, v = \neg(\Diamond\, \neg v).$$

Considering a path, the formula $\Diamond\, v$ intuitively states that $v$ holds eventually in the future and $\Box\, v$ specifies that $v$ holds from now on forever.

**Example 108.** We present properties that can be specified by APCTL*. In every case the action spaces is equipped with the corresponding action event family from Example 106.

(1) Let $Act = Eval[\{\mathrm{temp}\}]$ where the variable temp intuitively stands for the temperature of room with a temperature controller . The APCTL* state formula

$$\sigma_1 = \exists\mathbb{P}_{\geq 0.7}[\Box(\mathrm{temp}{\geq}18 \wedge \mathrm{temp}{\leq}22)]$$

asserts the existence of a scheduler such that the temperature forever stays between 18 and 22 degree Celsius with a probability greater than or equal to 0.7.

(2) Consider an action space $Act$ with goal $\in Act$. The APCTL* state formula

$$\sigma_2 = \forall\mathbb{P}_{>0.9}[Acc[{\geq}0] \,\mathrm{U}\, \{\mathrm{goal}\}]$$

expresses that for all resolutions of the non-determinism the accumulated reward referring, e.g., to the available energy or the allocated resources, until reaching the goal action never drops below zero with a probability greater than 0.9.

(3) The following formula nests the exists modality. Regard an action space *Act* such that $\{\text{error}, \text{operating}\} \subseteq Act$. The APCTL$^*$ state formula

$$\sigma_3 = \forall \mathbb{P}_{=1}[\Box(\{\text{error}\} \rightarrow \exists \mathbb{P}_{>0.8}[\Diamond(Acc[\leq 17] \wedge \{\text{operating}\})])]$$

specifies the following resilient condition on an adaptive control system: almost surely, after an error has been occurred, the system fosters towards an operating mode with probability greater than 0.8 accumulating at most 17 cost. The cost may refer to the time that has been elapsed since the occurrence of the indicated error.

⌟

The remainder of this section introduces three fragments of APCTL$^*$ that become important in our discussions on the logical characterisation of the simulation preorder and the bisimulation equivalence later on in this chapter. For instance, having corresponding results from the discrete setting in mind (see, e.g., Section 7.5 in [BK08]), one cannot expect that the full temporal logic APCTL$^*$ is preserved by the simulation preorder. This leads to the following definition of the existential fragment of APCTL$^*$.

**Definition 109.** Let $(Act, \mathcal{A})$ be an action event space. An *$\exists$APCTL$^*$ state formula* (*over* $(Act, \mathcal{A})$) is formed by the following grammar:

$$\sigma \quad ::= \quad tt \mid \sigma \wedge \sigma \mid \sigma \vee \sigma \mid \exists \tau$$

where $\tau$ is an $\exists$APCTL$^*$ path-measure formula. Here, an *$\exists$APCTL$^*$ path-measure formula* (*over* $(Act, \mathcal{A})$) is generated by the following grammar:

$$\tau \quad ::= \quad \tau \wedge \tau \mid \tau \vee \tau \mid \mathbb{P}_{\trianglerighteq q}[v]$$

where $\trianglerighteq \in \{=, \geq, >\}$, $q \in [0,1] \cap \mathbb{Q}$, and $v$ is a $\exists$APCTL$^*$ path formula. An *$\exists$APCTL$^*$ path formula* (*over* $(Act, \mathcal{A})$) is defined by the following grammar:

$$v \quad ::= \quad \sigma \mid A \mid v \wedge v \mid v \vee v \mid \bigcirc v \mid v \,\mathrm{U}\, v \mid Acc[\bowtie r]$$

where $\sigma$ is an APCTL$^*$ state formula, $A \in Bool[\mathcal{A}]$, $\bowtie \in \{<, \leq, =, \geq, >\}$, and $r \in \mathbb{Q}$. ⌟

It is easy to see that the logic $\exists$APCTL$^*$ yields a fragment of APCTL$^*$. While for every APCTL$^*$ state formula $\sigma$ the negation $\neg\sigma$ is an APCTL$^*$ state formula, the corresponding statement does not hold for $\exists$APCTL$^*$. In particular, in $\exists$APCTL$^*$ the forall quantification over schedulers is missing that actually motivates the previously introduced sublogic. Besides this, probabilities for path properties can only be bounded from below by the probability modality.

**Definition 110.** Let $(Act, \mathcal{A})$ be an action event space. An *APCTL$_\circ$ state formula* (*over* $(Act, \mathcal{A})$) is formed by the following grammar:

$$\sigma \quad ::= \quad tt \mid \sigma \wedge \sigma \mid \exists \tau$$

where $\tau$ is an APCTL$_\circ$ path-measure formula. Here, an *APCTL$_\circ$ path-measure formula* (*over* $(Act, \mathcal{A})$) is generated by the following grammar:

$$\tau \quad ::= \quad \tau \wedge \tau \mid \mathbb{P}_{>q}[v] \mid \mathbb{P}_{<q}[v]$$

where $q \in [0,1] \cap \mathbb{Q}$ and $v$ is a APCTL$_\circ$ path formula. An *APCTL$_\circ$ path formula* (*over* $(Act, \mathcal{A})$) is defined by the following grammar:

$$v \quad ::= \quad \bigcirc\, (A \wedge \sigma)$$

where $A \in Bool[\mathcal{A}]$ and $\sigma$ is an APCTL$_\circ$ state formula. ⌟

Although the boolean fragment of APCTL$_\circ$ is restricted to conjunctions and include no negations, the expressiveness of APCTL$_\circ$ and $\exists$APCTL$^*$ are incomparable as in APCTL$_\circ$ probabilities for path properties can be also bounded from above by arbitrary probability thresholds. Most importantly, the logic APCTL$_\circ$ basically allows only for those APCTL$^*$ path formulas that include solely the next modality, in particular, the until modality is missing. Consequently, there is no chance to express the eventually and the always modality in APCTL$_\circ$. Intuitively, formulas of APCTL$_\circ$ are only capable of specifying constraints for direct action-state successors in a specific state. Besides this, the logic APCTL$_\circ$ also includes no accumulation modality.

**Definition 111.** Let $(Act, \mathcal{A})$ be an action event space. An *APCTL$_\bullet$ state formula* (*over* $(Act, \mathcal{A})$) is formed by the following grammar:

$$\sigma \quad ::= \quad tt \mid \sigma \wedge \sigma \mid \exists \tau$$

where $\tau$ is an APCTL$_\bullet$ path-measure formula. Here, an *APCTL$_\bullet$ path-measure formula* (*over* $(Act, \mathcal{A})$) is generated by the following grammar:

$$\tau \quad ::= \quad \tau \wedge \tau \mid \mathbb{P}_{>q}[v]$$

where $q \in [0,1] \cap \mathbb{Q}$ and $v$ is a APCTL$_\bullet$ path formula. An *APCTL$_\bullet$ path formula* (*over* $(Act, \mathcal{A})$) is defined by the following grammar:

$$v \quad ::= \quad v \vee v \mid \bigcirc\, (A \wedge \sigma)$$

where $A \in Bool[\mathcal{A}]$ and $\sigma$ is an APCTL$_\bullet$ state formula. ⌟

The previously introduced sublogic APCTL$_\bullet$ of APCTL$^*$ is in the same spirit as the fragment APCTL$_\circ$. However, compared to APCTL$_\circ$, the latter defined fragment involves no form of negation, in particular, probabilities of path properties can only be bounded from below by the probability modality and hence, APCTL$_\circ$ yields a fragment of $\exists$APCTL$^*$. In comparison to APCTL$_\circ$, there is an additional disjunction operator for APCTL$_\bullet$ path formulas. As we see in Section 5.4, this disjunction operator is indispensable.

## 5.2 Outer-measure semantics

We interpret the temporal logic APCTL$^*$ over STSs that are additionally equipped with an action event family and a reward function. Action event families have been introduced in the previous section (see Definition 105). Reward functions are defined as expected:

**Definition 112.** For every Polish space *Act* a *reward function* (*on Act*) is Borel function $rew\colon Act \to \mathbb{R}$ mapping every action in *Act* to a real number (*reward*). ⌟

The reward function may, e.g., specify the consumed energy or the produced utility accompanied with the execution of an action. Let $\mathcal{T} = (Sta, Act, \to)$ be an STS and *rew* be a reward function. Intuitively, if the action *act* is executed, the reward $rew(act)$ is gained. It is therefore convenient to extend the reward function *rew* to the set of all finite paths (also denoted by *rew*) as follows: for every state $s \in Sta$ let

$$rew(s) = 0$$

and for every finite path $\hat{\pi} = s_0\, act_1\, s_1 \ldots act_n\, s_n$ where $n \in \mathbb{N}$ with $n > 0$ let

$$rew(\hat{\pi}) = rew(act_1) + \ldots + rew(act_n),$$

i.e., $rew(\hat{\pi})$ records the accumulated reward along the given finite path.

**Definition 113.** Let $\mathcal{T} = (Sta, Act, \to)$ be an STS, $\mathcal{A}$ be an action event family, and *rew* be a reward function. The *satisfaction relation* $\models$ for APCTL$^*$ state formulas is defined as follows where $s \in Sta$ is a state:

$$\begin{array}{lll} s \models \sigma_1 \wedge \sigma_2 & \text{iff} & s \models \sigma_1 \text{ and } s \models \sigma_2, \\ s \models \neg\sigma & \text{iff} & s \not\models \sigma, \\ s \models \exists\tau & \text{iff} & \text{there exists an } s\text{-scheduler } \mathfrak{S} \text{ with } Pr[\mathfrak{S}] \models \tau. \end{array}$$

The *satisfaction relation* $\models$ for APCTL* path-measure formulas is given below where $\chi$ denotes a probability measure on *Path*:

$$\begin{aligned}
&\chi \models \tau_1 \wedge \tau_2 && \text{iff} \quad \chi \models \tau_1 \text{ and } \chi \models \tau_2,\\
&\chi \models \neg\tau && \text{iff} \quad \chi \not\models \tau,\\
&\chi \models \mathbb{P}_{\bowtie q}[v] && \text{iff} \quad \chi^{\text{out}}(\{\pi \in \mathit{Path}\ ;\ \pi \models v\}) \bowtie q.
\end{aligned}$$

The *satisfaction relation* $\models$ for APCTL* path formulas is defined as follows where $n \in \mathbb{N}$ is a natural number and $\pi = s_0\, act_1\, s_1 \dots$ is an infinite path in *Path*:

$$\begin{aligned}
&\langle \pi, n\rangle \models \sigma && \text{iff} \quad s_n \models \sigma,\\
&\langle \pi, n\rangle \models A && \text{iff} \quad act_{n+1} \in A,\\
&\langle \pi, n\rangle \models v_1 \wedge v_2 && \text{iff} \quad \langle \pi, n\rangle \models v_1 \text{ and } \langle \pi, n\rangle \models v_2,\\
&\langle \pi, n\rangle \models \neg v && \text{iff} \quad \langle \pi, n\rangle \not\models v,\\
&\langle \pi, n\rangle \models \bigcirc v && \text{iff} \quad \langle \pi, n+1\rangle \models v,\\
&\langle \pi, n\rangle \models v_1 \,\mathrm{U}\, v_2 && \text{iff} \quad \text{there exists } i \in \mathbb{N} \text{ with } n \leq i \text{ so that } \langle \pi, i\rangle \models v_2\\
& && \qquad\quad \text{and } \langle \pi, j\rangle \models v_1 \text{ for all } j \in \mathbb{N} \text{ with } n \leq j < i,\\
&\langle \pi, n\rangle \models \mathit{Acc}[\bowtie r] && \text{iff} \quad \mathit{rew}(s_0\, act_1\, s_1 \dots act_n\, s_n) \bowtie r.
\end{aligned}$$

For every path $\pi \in \mathit{Path}$ and APCTL* path formula $v$ we also use $\pi \models v$ as a shorthand notation for $\langle \pi, 0\rangle \models v$. ⌟

As before, let $\mathcal{T} = (\mathit{Sta}, \mathit{Act}, \to)$ be an STS, $\mathcal{A}$ be an action event family, and *rew* be a reward function. For every APCTL* state formula $\sigma$, path-measure formula $\tau$, and path formula $v$ we introduce the respective satisfaction sets by

$$\begin{aligned}
&[\![\sigma]\!] = \{s \in \mathit{Sta}\ ;\ s \models \sigma\},\\
&[\![\tau]\!] = \{\chi \in \mathit{Prob}[\mathit{Path}]\ ;\ \chi \models \tau\},\\
&[\![v]\!] = \{\pi \in \mathit{Path}\ ;\ \pi \models v\}.
\end{aligned}$$

Relying on the presented semantics, the temporal logics introduced in Section 5.1 induce the following relations over the state space. The *APCTL* equivalence* and the *APCTL$_\circ$ equivalence* are the binary relations $\simeq^*$ and $\simeq_\circ$ over the state space *Sta*, respectively, given as follows: for every states $s_a, s_b \in \mathit{Sta}$,

$$s_a \simeq^* s_b \quad \text{iff} \quad \text{for every APCTL* state formula one has } s_a \models \sigma \text{ iff } s_b \models \sigma,$$

$$s_a \simeq_\circ s_b \quad \text{iff} \quad \text{for every APCTL}_\circ \text{ state formula one has } s_a \models \sigma \text{ iff } s_b \models \sigma.$$

Accordingly, the *$\exists$APCTL$^*$ preorder* and the *APCTL$_\bullet$ preorder* are the binary relations $\preceq^\exists$ and $\preceq_\bullet$ over *Sta*, respectively, defined as follows: for every states $s_a, s_b \in Sta$,

$$s_a \preceq^\exists s_b \quad \text{iff} \quad \text{for every } \exists\text{APCTL}^* \text{ state formula one has } s_a \models \sigma \text{ implies } s_b \models \sigma,$$

$$s_a \preceq_\bullet s_b \quad \text{iff} \quad \text{for every APCTL}_\bullet \text{ state formula one has } s_a \models \sigma \text{ implies } s_b \models \sigma.$$

In Definition 113 we rely on the outer-measure function of $\chi$ for the definition of the semantics of the probability modality $\mathbb{P}_{\bowtie q}[v]$. The reason is that we believe that the set $[\![v]\!]$ may not be Borel in *Path* for arbitrary STSs. As we have no counterexample for this yet, let us briefly sketch the difficulty. If one wants to show that $[\![v]\!]$ is Borel in *Path*, the idea is to rely on an induction over the construction of formulas. Hence, one also has to show that the sets $[\![\sigma]\!]$ and $[\![\tau]\!]$ are Borel in *Sta* and *Prob*[*Path*], respectively. Clearly, considering for instance the APCTL$^*$ state formula $\sigma = \sigma_1 \wedge \sigma_2$ for some APCTL$^*$ path formulas $\sigma_1$ and $\sigma_2$, it easily follows that $[\![\sigma]\!]$ is Borel in *Sta* assuming the sets $[\![\sigma_1]\!]$ and $[\![\sigma_2]\!]$ are Borel in *Sta* since $[\![\sigma]\!] = [\![\sigma_1]\!] \cap [\![\sigma_2]\!]$. However, the case $\sigma = \exists\tau$ where $\tau$ is an APCTL$^*$ path-measure formula with $[\![\tau]\!]$ being a Borel set in *Prob*[*Path*] is more intricate: since the semantics of the exists modality involves an existential quantification ranging over all schedulers, we think that one cannot expect that the set $[\![\sigma]\!]$ is Borel in *Sta*. Recall, sets defined in terms of an existential quantification are typically not Borel (see also Section 2.2). At this point, one may ask whether the set $[\![\sigma]\!]$ is Souslin in *Sta*. Indeed, we see a chance to answer the latter question positively for every Souslin STS (see also [MS00]). Such a result needs a deeper analysis of structural properties of the set of all path measures induced by a scheduler.

Interestingly, the presented approach to circumvent the latter mentioned problems using outer measures causes no mathematical difficulties with regard to the main object of this chapter concerning the logical characterisations of simulation preorder and bisimulation equivalence. Inspecting the previous discussion carefully, it turns out that the concept of outer measure can be omitted for PCTL$^*$ formulas where the exists modality is not nested as then all the satisfaction sets of the corresponding subformulas are Borel. This is for instance the case for the formulas $\sigma_1$ and $\sigma_2$ in Example 108. However, compared to that, the formula $\sigma_3$ in Example 108 needs some care as here, the exists modality occurs nested. Obviously, if the state space of the STS under consideration is countable, then the outer-measure semantics and the standard semantics are the same.

**Example 114.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a non-blocking Souslin STS, $\mathcal{A}$ be an action event family, and *rew* be a reward function. Consider the APCTL$^*$ state formula

$$\sigma = \exists\mathbb{P}_{\bowtie q}[\bigcirc (A \wedge \sigma')]$$

with $\bowtie \in \{<, \leq, =, \geq, >\}$, $q \in [0,1] \cap \mathbb{Q}$, $A \in \mathit{Bool}[\mathcal{A}]$, and $\sigma'$ being an APCTL$^*$ state formula. Then for every state $s \in \mathit{Sta}$ we have the following equivalence:

$$s \models \sigma \quad \text{iff} \quad \text{there is } \varphi \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}] \text{ with } s \to \varphi \text{ and } \varphi^{\text{out}}(A \times [\![\sigma']\!]) \bowtie q.$$

The formal argument for this claim is as follows. Let $s \in \mathit{Sta}$ be a state. Abbreviate the APCTL$^*$ path formula $\bigcirc(A \wedge \sigma')$ by $v$. One has $[\![v]\!] = \mathit{Sta} \times (A \times [\![\sigma']\!]) \times (\mathit{Act} \times \mathit{Sta})^{\omega}$. For every $s$-scheduler $\mathfrak{S}$ it is hence easy to see that $\mathit{Pr}[\mathfrak{S}]^{\text{out}}([\![v]\!]) = \mathfrak{S}(s)^{\text{out}}(A \times [\![\sigma']\!])$. Consequently, the implication from the left to the right of the claimed equivalence follows directly from the fact that $s \to \mathfrak{S}(s)$ for every $s$-scheduler $\mathfrak{S}$. The reverse implication is a consequence of Theorem 65. Indeed, thanks to the assumptions on the STS under consideration, for every $\varphi \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ with $s \to \varphi$ there exists an $s$-scheduler $\mathfrak{S}$ with $\mathfrak{S}(s) = \varphi$. ⌟

Clearly, the semantics for APCTL$^*$ directly transfers to the fragments $\exists$APCTL$^*$, APCTL$_\circ$, and APCTL$_\bullet$. In this context, the previous example illustrates that the concepts of schedulers and induced path measures can be avoided for a direct definition of the corresponding semantics of APCTL$_\circ$ and APCTL$_\bullet$. This also confirms corresponding observations in Section 5.1 stating that APCTL$_\circ$ and APCTL$_\bullet$ yield inexpressive sublogics of APCTL$^*$.

## 5.3 Borel concerning the hit sigma algebra

We present a sublcass of (Souslin) STSs such that all the satisfaction sets of APCTL$_\circ$ formulas and APCTL$_\bullet$ formulas are Borel. Recalling discussions in Section 3.1, the Souslin property in the context of STSs is basically motivated by a requirement of the measurable-selection principle given by Theorem 21. The following condition on an STS is in the spirit of an assumption of the classical measurable-selection principle given by Theorem 18:

**Definition 115.** An STS $\mathcal{T} = (\mathit{Sta}, \mathit{Act}, \to)$ is called *Borel concerning the hit sigma algebra* provided for every open set $O \subseteq \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ the set

$$\{s \in \mathit{Sta} \,;\, \mathit{Enabled}[s] \cap O \neq \varnothing\}$$

is Borel in *Sta*. ⌟

In the discussions of Section 2.3 concerning Theorem 18 we have introduced the notion of hit sigma algebras that also justifies the choice for the name of the new subclass of STSs. Intuitively, the set $\{s \in \mathit{Sta} \,;\, \mathit{Enabled}[s] \cap O \neq \varnothing\}$ consists of those states $s$ such that there exists an enabled distribution over action-state pairs that hit the given test set $O$. These test sets are given by all the open subsets of $Prob[Act \times Sta]$. Definition 115 is also in the spirit of non-deterministic labelled Markov processes [DTW12, Wol12] that is precisely discussed in Section 7.4. The following remark establishes a connection between Souslin STSs and those STSs satisfying the condition in Definition 115:

**Remark 116.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS that is non-blocking and such that the set $Enabled[s]$ is closed in $Prob[Act \times Sta]$. The latter assumption in particular includes the case where $\mathcal{T}$ is image finite, i.e., where the set $Enabled[s]$ is finite for every state $s \in Sta$. By Remark 23, if $\mathcal{T}$ is Borel concerning the hit sigma algebra, then the STS $\mathcal{T}$ is Souslin. The cited remark also shows that the reverse implication does not hold in general, i.e., the STS $\mathcal{T}$ may not be Borel concerning the hit sigma algebra even though it is Souslin. ⌟

**Theorem 117.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and $\mathcal{A}$ be an action event family. Assume that the STS $\mathcal{T}$ is non-blocking, image-finite, and Borel concerning the hit sigma algebra. Then the following two statements hold:*

(1) *The set $[\![\sigma]\!]$ is Borel in Sta for every APCTL$_\bullet$ state formula $\sigma$.*

(2) *The set $[\![\sigma]\!]$ is Borel in Sta for every APCTL$_\circ$ state formula $\sigma$.*

*Proof.* We rely on Remark 19 providing a refinement of the measurable-selection principle given by Theorem 18: there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of Borel functions $f_n\colon Sta \rightarrow Prob[Act \times Sta]$ such that for every state $s \in Sta$ it holds

$$Enabled[s] = \{f_n(s)\ ;\ n \in \mathbb{N}\}.$$

Let $k \in \mathbb{N}$ and define the APCTL$_\bullet$ state formula $\sigma$ by

$$\sigma \quad = \quad \exists(\mathbb{P}_{>q_0}[v_0] \ \wedge \ \ldots \ \wedge \ \mathbb{P}_{>q_k}[v_k])$$

where for every $i \in \{0, \ldots, k\}$, $q_i$ is a rational number in $[0, 1]$ and $v_i$ is the APCTL$_\bullet$ path-measure formula given by

$$v_i \quad = \quad \bigcirc(A_{i0} \wedge \sigma_{i0}) \ \vee \ \ldots \ \vee \ \bigcirc(A_{ik_i} \wedge \sigma_{ik_i})$$

for some natural number $k_i \in \mathbb{N}$, $A_{i0}, \ldots, A_{ik_i} \in Bool[\mathcal{A}]$ and APCTL$_\bullet$ state formula $\sigma_{i0}, \ldots, \sigma_{ik_i}$. Assume that for all $i \in \{0, \ldots, k\}$ the sets $[\![\sigma_{i0}]\!], \ldots, [\![\sigma_{ik_i}]\!]$ are Borel in $Sta$. Inspecting the syntax of APCTL$_\bullet$ state formulas and relying on an induction over the construction of $\mathrm{APCTL}_\bullet$ state formulas, it suffices to show that the set $[\![\sigma]\!]$ is Borel in *Sta* in order to conclude the lemma.

For every $n \in \mathbb{N}$ and $i \in \{0, \ldots, k\}$ define the function $g_{ni} \colon Sta \to [0, 1]$,

$$g_{ni}(s) = f_n(s)((A_{i0} \times [\![\sigma_{i0}]\!]) \cup \ldots \cup (A_{ik_i} \times [\![\sigma_{ik_i}]\!])).$$

Inspecting Example 114, it is easy to see that

$$[\![\sigma]\!] = \bigcup_{n \in \mathbb{N}} \bigcap_{i \in \{0,\ldots,k\}} g_{ni}^{-1}((q_i, 1]).$$

According to Remark 5, for every $n \in \mathbb{N}$ and $i \in \{0, \ldots, k\}$ the function $g_{ni}$ is Borel and therefore, it directly follows that the set $[\![\sigma]\!]$ is Borel in *Sta*. This shows statement (1). Claim (2) follows analogously. □

As a consequence, the outer-measure semantics and the standard semantics for $\mathrm{APCTL}_\circ$ and $\mathrm{APCTL}_\bullet$ are the same concerning the subclass of STSs investigated in the previous theorem. In particular, it is not necessary to consider outer-measure functions of probability measures for the interpretation of properties specified by $\mathrm{APCTL}_\circ$ or $\mathrm{APCTL}_\bullet$. According to Remark 116, every STS satisfying the conditions in Theorem 117 is also Souslin.

## 5.4 Logical characterisation

The main object of this section is to show that the $\mathrm{APCTL}_\bullet$ preorder $\preceq_\bullet$ and the simulation preorder $\preceq$ are the same and accordingly, that the $\mathrm{APCTL}_\circ$ equivalence $\simeq_\circ$ and the bisimulation equivalence $\simeq$ coincide for a subclass of STSs. The proof below adapts the argumentation scheme of the recent contribution [FKP17] for labelled Markov processes. More precisely, the key ingredients for the following argumentation are given by Theorems 39 and 40 whose proofs heavily rely on the techniques in [FKP17]. Moreover, as in [FKP17], the facts that the two temporal logics $\mathrm{APCTL}_\bullet$ and $\mathrm{APCTL}_\circ$ are countable turn out to be crucial.

**Theorem 118.** *Let $\mathcal{T} = (Sta, Act, \to)$ be an STS and $\mathcal{A}$ be an action event family. Assume that the STS $\mathcal{T}$ is non-blocking, image-finite, and Borel concerning the hit sigma algebra. For every states $s_a, s_b \in Sta$ the following two equivalences hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\mathrm{sou}} s_b$ *iff* $s_a \preceq_\bullet s_b$.

(2) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$ *iff* $s_a \simeq_\circ s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth.* ⌟

The proof of the theorem can be found below. As a byproduct of the presented logical characterisation, we derive a subclass of Souslin STSs (see also Remark 116), where the simulation preorder and the Souslin-simulation preorder are the same and accordingly, where the bisimulation equivalence and the Souslin-bisimulation equivalence collapse. Applying our results from Chapter 4, we obtain the following corollary:

**Corollary 119.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS. Assume that $\mathcal{T}$ is non-blocking, image-finite, and Borel with respect to open hit sets. For every $s_a, s_b \in Sta$ the two statements below hold:*

(1) $s_a \preceq s_b$ *implies* $s_a \leq^{\mathrm{tr}} s_b$.

(2) $s_a \simeq s_b$ *implies* $s_a =^{\mathrm{tr}} s_b$.

*Proof.* By Remark 4, we can safely assume some action event family on $Act$. Remark 116 shows that the STS $\mathcal{T}$ is Souslin. Consequently, the claim immediately follows Corollary 83 and Theorem 118. □

It remains to show the section's main result:

*Proof of Theorem 118.* The proof schema for both claimed chain of equivalences is the same. In what follows we first show that formulas that can be formulated in APCTL$_\circ$ and APCTL$_\bullet$ are preserved by the simulation preorder and the bisimulation equivalence, respectively. In a second step, we show that the relations $\preceq_\bullet$ and $\simeq_\circ$ induced by the temporal logics are a simulation and bisimulation, respectively. Statements (1) and (2) then easily follow as the relations $\preceq_\bullet$ and $\simeq_\circ$ are weakly smooth and smooth, respectively. To be more precise, relying on Theorem 117, the facts the temporal logics APCTL$_\bullet$ and APCTL$_\circ$ are countable, and Remark 38, the family

$$\mathcal{C}_\preceq = \{[\![\sigma]\!] \; ; \sigma \text{ is an APCTL}_\bullet \text{ state formula}\}$$

is a witnesses of the weakly smoothness of $\preceq_\bullet$ and accordingly,

$$\mathcal{C}_\simeq = \{[\![\sigma]\!] \; ; \sigma \text{ is an APCTL}_\circ \text{ state formula}\}$$

yields a witnesses of the smoothness of $\simeq_\circ$.

*Preservation of temporal properties.* Let $s_a, s_b \in Sta$ be states. We only show that $s_a \preceq_\bullet s_b$ follows from $s_a \preceq s_b$. Indeed, one analogously derives that $s_a \simeq s_b$ implies $s_a \simeq_\circ s_b$. Consider a simulation $R$ with $\langle s_a, s_b\rangle \in R$. Let $k \in \mathbb{N}$ and define the APCTL$_\bullet$ state formula $\sigma$ by

$$\sigma = \exists(\mathbb{P}_{>q_0}[v_0] \;\wedge\; \ldots \;\wedge\; \mathbb{P}_{>q_k}[v_k])$$

where $q_0, \ldots, q_k \in [0,1] \cap \mathbb{Q}$ are rational numbers between zero and one and $v_0, \ldots, v_k$ are APCTL$_\bullet$ path-measure formulas defined as follows: for every $i \in \{0, \ldots, k\}$,

$$v_i = \bigcirc (A_{i0} \wedge \sigma_{i0}) \;\vee\; \ldots \;\vee\; \bigcirc (A_{ik_i} \wedge \sigma_{ik_i})$$

where $A_{i0}, \ldots, A_{ik_i} \in \mathit{Bool}[\mathcal{A}]$ and $\sigma_{i0}, \ldots, \sigma_{ik_i}$ are APCTL$_\bullet$ state formulas. We assume that for every $i \in \{0, \ldots, k\}$ and $j \in \{0, \ldots, k_i\}$ the implication below holds:

$$s_a \models \sigma_{ij} \quad \text{implies} \quad s_b \models \sigma_{ij}.$$

Assume $s_a \models \sigma$. According the syntax of a APCTL$_\bullet$ state formula and relying on an induction over the construction of APCTL$_\bullet$ formulas, it suffices to show $s_b \models \sigma$ in order to conclude $s_a \preceq_\bullet s_b$.

For every $i \in \{0, \ldots, k\}$ and $j \in \{0, \ldots, k_i\}$ it is easy to see that the set $A_{ij} \times [\![\sigma_{ij}]\!]$ is upper $R^{Act}$-stable. Therefore, for every $i \in \{0, \ldots, k\}$ the set $B_i$ is upper $R^{Act}$-stable where

$$B_i = (A_{i0} \times [\![\sigma_{i0}]\!]) \cup \ldots \cup (A_{ik_i} \times [\![\sigma_{ik_i}]\!]).$$

Applying Lemma 117, for every $i \in \{0, \ldots, k\}$ the set $B_i$ is Borel in $\mathit{Act} \times \mathit{Sta}$. Relying on the insights in Example 114 and using the assumption $s_a \models \sigma$, there exists a probability measure $\varphi_a \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ such that

$$s_a \to \varphi_a \quad \text{and} \quad \varphi_a(B_i) > q_i \text{ for every } i \in \{0, \ldots, k\}.$$

Using that the relation $R$ is a simulation with $\langle s_a, s_b \rangle \in R$, there hence exists a probability measure $\varphi_b \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ such that

$$s_b \to \varphi_b \quad \text{and} \quad \langle \varphi_a, \varphi_b \rangle \in (R^{Act})^{\mathrm{wgt}}.$$

Thanks to Remark 34, for every $i \in \{0, \ldots, k\}$ it hence holds

$$\varphi_b(B_i) \geq \varphi_a(B_i) > q_i.$$

Applying the same argument as in Example 114, we can conclude $s_b \models \sigma$.

*Characterisation of simulation preorder and bisimulation equivalence.* To obtain a proof for the theorem under consideration, we show that the relation $\preceq_\bullet$ is a simulation and that the relation $\simeq_\circ$ is a bisimulation. this proof uses the following additional notions For presentation purposes the remainder of this proof uses the following additional notions. for every family $\mathcal{C}$ of subsets of $\mathit{Sta}$ the relation $\mathit{Rel}_\preceq[\mathcal{C}]$ over $\mathit{Sta}$ is given by

$$\mathit{Rel}_\preceq[\mathcal{C}] = \{\langle s_a, s_b \rangle \in \mathit{Sta} \times \mathit{Sta} \;;\; s_a \in S \text{ implies } s_b \in S \text{ for every } S \in \mathcal{C}\}.$$

and accordingly, the relation $Rel_{\simeq}[\mathcal{C}]$ over $Sta$ is defined as

$$Rel_{\simeq}[\mathcal{C}] = \{\langle s_a, s_b\rangle \in Sta \times Sta\ ;\ s_a \in S \text{ iff } s_b \in S \text{ for every } S \in \mathcal{C}\}.$$

*Ad* (1). Let $\mathcal{C}'_{\preceq}$ be the family of subsets of $Act \times Sta$ given by

$$\mathcal{C}'_{\preceq} = \{A \times [\![\sigma]\!]\ ;\ A \in Bool[\mathcal{A}] \text{ and } \sigma \text{ is a APCTL}_{\bullet} \text{ state formula}\}.$$

Relying on the requirements on action event families, it is easy to see that the family $\mathcal{C}'_{\preceq}$ is countable. Moreover, according to Theorem 117, every element in $\mathcal{C}'_{\preceq}$ is a Borel set in $Act \times Sta$. It is easy to see that $\mathcal{C}'_{\preceq}$ is closed under finite intersections. Indeed, for every $A_1, A_2 \in Bool[\mathcal{A}]$ and every APCTL$_{\bullet}$ state formulas $\sigma_1$ and $\sigma_2$ one has

$$(A_1 \times [\![\sigma_1]\!]) \cap (A_2 \times [\![\sigma_2]\!]) = (A_1 \cap A_2) \times ([\![\sigma_1]\!] \cap [\![\sigma_2]\!]) = (A_1 \wedge A_2) \times [\![\sigma_1 \wedge \sigma_2]\!].$$

On top of that, it holds $Rel_{\preceq}[\mathcal{C}'_{\preceq}] = (\preceq_{\bullet})^{Act}$. Here, the inclusion $Rel_{\preceq}[\mathcal{C}'_{\preceq}] \supseteq (\preceq_{\bullet})^{Act}$ is trivial and the reverse inclusion $Rel_{\preceq}[\mathcal{C}'_{\preceq}] \subseteq (\preceq_{\bullet})^{Act}$ follows immediately from the fact that the action event family $\mathcal{A}$ separates the points in $Act$.

The family $\mathcal{C}_{\preceq}$ is defined as the smallest family of subsets of $Act \times Sta$ such that $\mathcal{C}_{\preceq}$ is closed under finite unions and so that $\mathcal{C}'_{\preceq} \subseteq \mathcal{C}_{\preceq}$. It is easy to see that every set contained in $\mathcal{C}_{\preceq}$ is Borel in $Act \times Sta$. Moreover, since the family of sets $\mathcal{C}'_{\preceq}$ is countable, the newly introduced family of sets $\mathcal{C}_{\preceq}$ is also countable. Using that $\mathcal{C}'_{\preceq}$ is closed under finite intersections, the family of sets $\mathcal{C}_{\preceq}$ is also closed under finite intersections. Additionally, it is easy to see that $Rel_{\preceq}[\mathcal{C}_{\preceq}] = Rel_{\preceq}[\mathcal{C}'_{\preceq}] = (\preceq_{\bullet})^{Act}$.

It follows that $(\preceq_{\bullet})^{Act}$ is weakly smooth where $\mathcal{C}_{\preceq}$ is a witness that is closed under finite intersections. We are hence in the situation of Theorem 39 that yields the following statement. For every $\varphi'_a, \varphi'_b \in Prob[Act \times Sta]$ one has the equivalence below:

$$\langle \varphi'_a, \varphi'_b\rangle \in ((\preceq_{\bullet})^{Act})^{\mathrm{wgt}} \quad \text{iff} \quad \text{for every } B \in \mathcal{C}_{\preceq} \text{ it holds } \varphi'_a(B) \leq \varphi'_b(B).$$

Statement (1) can be derived as follows now. Towards a contradiction assume that the relation $\preceq_{\bullet}$ is no simulation. Hence, there are states $s_a, s_b \in Sta$ and a probability measure $\varphi_a \in Prob[Act \times Sta]$ satisfying the following three statements:

$$s_a \preceq_{\bullet} s_b, \quad s_a \to \varphi_a, \quad \text{and} \quad \langle \varphi_a, \varphi_b\rangle \notin ((\preceq_{\bullet})^{Act})^{\mathrm{wgt}} \text{ for all } \varphi_b \in Enabled[s_b].$$

As the set $Enabled[s_b]$ is required to be finite and not empty, there exists $n \in \mathbb{N} \setminus \{0\}$ and $\varphi_{b1}, \ldots, \varphi_{bn} \in Prob[Act \times Sta]$ with

$$Enabled[s_b] = \{\varphi_{b1}, \ldots, \varphi_{bn}\}.$$

For every $i \in \{1, \ldots, n\}$ it holds $\langle \varphi_a, \varphi_{bi} \rangle \notin ((\preceq_\bullet)^{Act})^{\mathrm{wgt}}$ and thus, relying on the first part of this proof, there exists a set $B_i \in \mathcal{C}_\cup$ and a rational number $q_i \in [0,1] \cap \mathbb{Q}$ with

$$\varphi_a(B_i) > q_i > \varphi_{bi}(B_i).$$

Inspecting the definition of $\mathcal{C}_\preceq$, for every $i \in \{1, \ldots, n\}$ there are a natural number $k_i \in \mathbb{N} \setminus \{0\}$, elements $A_{i1}, \ldots, A_{ik_i}$ of the action event family $\mathcal{A}$, and APCTL$_\bullet$ path formulas $\sigma_{i1}, \ldots, \sigma_{ik_i}$ such that

$$B_i = (A_{i1} \times [\![\sigma_{i1}]\!]) \cup \ldots \cup (A_{ik_i} \times [\![\sigma_{ik_i}]\!]).$$

For every $i \in \{1, \ldots, n\}$ define the APCTL$_\bullet$ path formula $v_i$ by

$$v_i = \bigcirc (A_{i1} \wedge \sigma_{i1}) \vee \ldots \vee \bigcirc (A_{ik_i} \wedge \sigma_{ik_i}).$$

The APCTL$_\bullet$ state formula $\sigma$ is introduced by

$$\sigma = \exists(\mathbb{P}_{>q_1}[v_1] \wedge \ldots \wedge \mathbb{P}_{>q_n}[v_n]).$$

Thanks to Example 114 and Theorem 117, it follows $s_a \models \sigma$ while $s_b \not\models \sigma$. The latter insight shows $\langle s_a, s_b \rangle \in (Sta \times Sta) \setminus \preceq_\bullet$ that yields a contradiction. We finally conclude that the relation $\preceq_\bullet$ is a simulation.

*Ad* (2). Our argument for (2) is similar to (1), however, relies on Theorem 40 rather than Theorem 39. As Theorem 40 requires a family of sets that is closed under finite intersections only (and not necessarily under finite unions), the following proof for (2) is less technical than (1). However, the core idea is the same.

Define the family of sets $\mathcal{C}_\simeq$ by

$$\mathcal{C}_\simeq = \{A \times [\![\sigma]\!] \,;\, A \in Bool[\mathcal{A}] \text{ and } \sigma \text{ is a APCTL}_\circ \text{ state formula}\}.$$

As for $\mathcal{C}'_\preceq$ in the first part of this proof, one can show that $\mathcal{C}_\simeq$ is countable, closed under finite intersections, and consists solely of Borel subsets of $Act \times Sta$. Therefore, applying Theorem 40, for every probability measures $\varphi'_a, \varphi'_b \in Prob[Act \times Sta]$ the equivalence below holds:

$$\langle \varphi'_a, \varphi'_b \rangle \in ((\simeq_\circ)^{Act})^{\mathrm{wgt}} \quad \text{iff} \quad \varphi'_a(B) = \varphi'_b(B) \text{ for all } B \in \mathcal{C}_\circ.$$

Towards a contradiction assume that $\simeq_\circ$ is no bisimulation. Thus, there are states $s_a, s_b \in Sta$ and $\varphi_a \in Prob[Act \times Sta]$ with $s_a \simeq_\circ s_b$, $s_a \to \varphi_a$, and $\langle \varphi_a, \varphi_b \rangle \notin ((\simeq_\circ)^{Act})^{\mathrm{wgt}}$

for all $\varphi_b \in \mathit{Enabled}[s_b]$ or vice versa, there are states $s_a, s_b \in \mathit{Sta}$ and $\varphi_b \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ with $s_a \simeq_\circ s_b$, $s_b \to \varphi_b$, and $\langle \varphi_a, \varphi_b \rangle \notin ((\simeq_\circ)^{\mathit{Act}})^{\mathrm{wgt}}$ for all $\varphi_a \in \mathit{Enabled}[s_a]$. For reasons of symmetry it suffices to discuss the first case. Let $s_a, s_b \in \mathit{Sta}$ and $\varphi_a \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ be such that

$$s_a \simeq_\circ s_b, \quad s_a \to \varphi_a, \quad \text{and} \quad \langle \varphi_a, \varphi_b \rangle \notin ((\simeq_\circ)^{\mathit{Act}})^{\mathrm{wgt}} \text{ for all } \varphi_b \in \mathit{Enabled}[s_b].$$

Let $n \in \mathbb{N}$ and $\varphi_{b0}, \ldots, \varphi_{bn} \in \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ be such that

$$\mathit{Enabled}[s_b] = \{\varphi_{b0}, \ldots, \varphi_{bn}\}.$$

For every $i \in \{0, \ldots, n\}$ there exists $A_i \in \mathcal{A}$, a APCTL$_\circ$ state formula $\sigma_i$, a comparison operator $\bowtie_i \in \{<, >\}$, and a rational number $q_i \in [0,1] \cap \mathbb{Q}$ such that

$$\varphi_a(A_i \times [\![\sigma_i]\!]) \ \bowtie_i \ q_i \ \bowtie_i \ \varphi_{bi}(A_i \times [\![\sigma_i]\!]).$$

Introduce the APCTL$_\circ$ state formula $\sigma$ by

$$\sigma = \exists(\mathbb{P}_{\bowtie q_0}[\bigcirc (A_0 \wedge \sigma_0)] \ \wedge \ \ldots \ \wedge \ \mathbb{P}_{\bowtie q_n}[\bigcirc (A_n \wedge \sigma_n)])$$

According to Example 114 and Theorem 117, we have $s_a \models \sigma$ while $s_b \not\models \sigma$. This yields a contradiction to $s_a \simeq_\circ s_b$ and hence, we are done. $\square$

Whereas the logic APCTL$_\bullet$ allows for disjunction of path formulas, the logic APCTL$_\circ$ does not. The proof of statement (1) of the previous theorem indeed requires disjunction of APCTL$_\bullet$ path formulas. It is natural to ask whether the argument can be adapted such that these disjunction are no longer needed in view of obtaining a simpler logic inducing a simulation. However, Example 4.3.4 in [Des99] shows that statement (1) of Theorem 118 does not longer hold when dropping the disjunction of APCTL$_\bullet$ path formulas. This example provides an STS with finitely many states where disjunction of APCTL$_\bullet$ path formulas is needed to distinguish states $s_a$ and $s_b$ where $s_b$ does not simulate $s_a$.

## 5.5 Simulation and bisimulation on infinite paths

This section revisits and extends the results presented in Section 4.5. To recall these contribution, consider a Souslin simulation $R$ for some Souslin STS $\mathcal{T} = (\mathit{Sta}, \mathit{Act}, \to)$ as well as probability measures $\mu_a, \mu_b \in \mathit{Prob}[\mathit{Sta}]$ such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. We have proved that for every $\mu_a$-scheduler $\mathfrak{S}_a$ there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that the corresponding induced finite-path measures of $\mathfrak{S}_a$ and $\mathfrak{S}_b$ are related by appropriate weight functions. The following material shows that even the (infinite-)path measures induced by the schedulers $\mathfrak{S}_a$ and $\mathfrak{S}_b$ are related by a certain weight function. However, compared to the setting in Section 4.5, our argument needs the stronger assumption that the simulation $R$ is even weakly smooth (and not merely Souslin, see also Remark 38).

Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS. Recalling Section 4.5, for every relation $R \subseteq Sta \times Sta$ and $n \in \mathbb{N}$ the lifting of $R$ to the set $Path_n$ of all finite paths of length $n$ is denoted by $R^{\mathrm{path},n}$. Accordingly, every relation $R \subseteq Sta \times Sta$ extends naturally to the relation $R^{\mathrm{path}}$ over the set of all infinite paths as follows: for every infinite paths $\pi_a = s_{a0}\, act_{a1}\, s_{a1} \ldots$ and $\pi_b = s_{b0}\, act_{b1}\, s_{b1} \ldots$,

$$\langle \pi_a, \pi_b \rangle \in R^{\mathrm{path}} \quad \text{iff} \quad \langle \pi_{a|n}, \pi_{b|n} \rangle \in R^{\mathrm{path},n} \text{ for every } n \in \mathbb{N}$$

where the two finite paths $\pi_{a|n}, \pi_{b|n} \in Path_n$ are given by $\pi_{a|n} = s_{a0}\, act_{a1}\, s_{a1} \ldots act_{an}\, s_{an}$ and $\pi_{b|n} = s_{b0}\, act_{b1}\, s_{b1} \ldots act_{bn}\, s_{bn}$, respectively, for every $n \in \mathbb{N}$. It obviously holds $\langle \pi_a, \pi_b \rangle \in R^{\mathrm{path}}$ precisely when both $\langle s_{an}, s_{bn} \rangle \in R$ for every $n \in \mathbb{N}$ and $act_{an} = act_{bn}$ for every $n \in \mathbb{N} \setminus \{0\}$. The following result is in the same spirit as Theorem 77.

**Theorem 120.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a , $R$ be a combined simulation, and $\mu_a, \mu_b \in Prob[Sta]$ be such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Assume that the relation $R$ is weakly smooth. Then for every $\mu_a$-scheduler $\mathfrak{S}_a$ there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that*

$$\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path}})^{\mathrm{wgt}}.$$

⌟

The formal proof of the stated result is presented below. Compared to Theorem 77, the assumption on the relation $R$ is stronger as every weakly smooth relation is in particular Souslin (see Remark 38). In exchange, the conclusion of Theorem 120 is stronger. Indeed, relying on Lemma 93, it is easy to see that $\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path},n})^{\mathrm{wgt}}$ for every $n \in \mathbb{N}$ is implied by $\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path}})^{\mathrm{wgt}}$.

Let us give an naive proof idea for Theorem 120, which yields an intuition for the mathematical difficulties associated to the stated result. According to Theorem 77, for every $\mu_a$-scheduler $\mathfrak{S}_a$ there is a $\mu_b$-scheduler $\mathfrak{S}_b$ such that for every $n \in \mathbb{N}$ there exists a weight function for $(Pr_n[\mathfrak{S}_a], R^{\mathrm{path},n}, Pr_n[\mathfrak{S}_b])$, say $W_n$. Applying, e.g., Kolmogorov's extension theorem (see Corollary 7.7.2 in [Bog07]), a first naive idea is to extend these weight functions to a single weight function for $(Pr[\mathfrak{S}_a], R^{\mathrm{path}}, Pr[\mathfrak{S}_b])$. However, there is no reason why the weight functions $W_n$ with $n \in \mathbb{N}$ are compatible in some sense as they are picked completely independent of each other. At this point, the assumption concerning the weakly smoothness of $R$ comes into play. Roughly speaking, this assumption guarantees the existence of compatible weight functions for $(Pr_n[\mathfrak{S}_a], R^{\mathrm{path},n}, Pr_n[\mathfrak{S}_b])$ for every $n \in \mathbb{N}$ that give rise to a weight function for $(Pr[\mathfrak{S}_a], R^{\mathrm{path}}, Pr[\mathfrak{S}_b])$.

**Corollary 121.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS, $R$ be a simulation, as well as $\mu_a, \mu_b \in Prob[Sta]$ be such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Assume that the relation $R$ is weakly smooth. Then for every $\mu_a$-scheduler $\mathfrak{S}_a$ there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that*

$$\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path}})^{\mathrm{wgt}}.$$

*Proof.* By Remark 38, the relation $R$ is Souslin in $Sta \times Sta$. Thus, similar to Corollary 78, the claim is a consequence of Theorem 120 together with Theorems 70 and 74. □

**Proof of main results.** The remainder of this section is devoted to a proof of Theorem 120. We need the following two auxiliary lemmas. Whereas the first lemma is routine, the second lemma formalises the given intuition for our proof of the main result.

**Lemma 122.** *Let $X$ be a Polish spaces and $(R_n)_{n \in \mathbb{N}}$ be a family of relations $R_n \subseteq X \times X$. Define the relation $R \subseteq X \times X$ by*

$$R = \bigcap_{n \in \mathbb{N}} R_n.$$

*Then the following two statements hold:*

(1) *If $R_n$ is weakly smooth for every $n \in \mathbb{N}$, then $R$ is weakly smooth.*

(2) *If $R_n$ is smooth for every $n \in \mathbb{N}$, then $R$ is smooth.*

*Proof.* We only provide a proof of claim (1) as the argument for (2) is completely analogous. For every $n \in \mathbb{N}$ assume that the relation $R_n$ is weakly smooth. For every $n \in \mathbb{N}$ let $\mathcal{C}_n$ be a witness for the weakly smoothness of $R_n$. Define the family of sets $\mathcal{C}$ by

$$\mathcal{C} = \bigcup_{n \in \mathbb{N}} \mathcal{C}_n.$$

For every $x_a, x_b \in X$ it easy to see that

$$\langle x_a, x_b \rangle \in R \quad \text{iff} \quad \text{for every } B \in \mathcal{C} \text{ it holds } x_a \in B \text{ implies } x_b \in B.$$

It follows that the relation $R$ is weakly smooth where $\mathcal{C}$ is a witness. □

**Lemma 123.** *Let $X$ be a Polish space. For every $n \in \mathbb{N}$ let $R_n \subseteq X \times X$ be a weakly smooth relation. Assume that for every $n \in \mathbb{N}$ it holds $R_n \supseteq R_{n+1}$. Define the relation $R \subseteq X \times X$ by*

$$R = \bigcap_{n \in \mathbb{N}} R_n.$$

*Then for every probability measures $\mu_a, \mu_b \in Prob[X]$ the implication below holds:*

$$\langle \mu_a, \mu_b \rangle \in (R_n)^{\mathrm{wgt}} \textit{ for every } n \in \mathbb{N} \quad \textit{implies} \quad \langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}.$$

*Proof.* For every $n \in \mathbb{N}$ let $\mathcal{C}'_n$ be a witness of the weakly smoothness of $R_n$. For every $n \in \mathbb{N}$ define the family $\mathcal{C}_n$ of Borel sets in $X$ as the smallest family of subsets of $X$ that is closed under both finite intersections and finite unions and such that $\mathcal{C}'_0 \cup \ldots \cup \mathcal{C}'_n \subseteq \mathcal{C}_n$. For every $n \in \mathbb{N}$ the family $\mathcal{C}'_n$ is countable and consequently, for every $n \in \mathbb{N}$ the family $\mathcal{C}_n$ is countable. For every $n \in \mathbb{N}$ and $x_a, x_b \in X$, exploiting the assumed chain of inclusions $R_0 \supseteq R_1 \supseteq \ldots \supseteq R_n$, it holds

$$\langle x_a, x_b \rangle \in R_n \quad \text{iff} \quad \text{for every } B \in \mathcal{C}_n \text{ one has } x_a \in B \text{ implies } x_b \in B.$$

Hence, for every $n \in \mathbb{N}$ the family of sets $\mathcal{C}_n$ constitutes a witness of the weakly smoothness of $R_n$. For every $n \in \mathbb{N}$ we additionally have the inclusion $\mathcal{C}_n \subseteq \mathcal{C}_{n+1}$.

Define the family $\mathcal{C}$ of subsets of $X$ by

$$\mathcal{C} = \bigcup_{n \in \mathbb{N}} \mathcal{C}_n.$$

The family of sets $\mathcal{C}$ is countable and for every $x_a, x_b \in X$ one also has

$$\langle x_a, x_b \rangle \in R \quad \text{iff} \quad \text{for every } B \in \mathcal{C} \text{ it holds } x_a \in B \text{ implies } x_b \in B.$$

The family $\mathcal{C}$ is therefore a witness of the weakly smoothness of $R$. Note, Lemma 122 (1) already shows that the relation $R$ is weakly smooth. Using that $\mathcal{C}_n \subseteq \mathcal{C}_{n+1}$ for all $n \in \mathbb{N}$, it is easy to see that $\mathcal{C}$ also forms a lattice on $X$.

Let $\mu_a, \mu_b \in Prob[X]$ be such that $\langle \mu_a, \mu_b \rangle \in (R_n)^{\mathrm{wgt}}$ for every $n \in \mathbb{N}$. For every $B \in \mathcal{C}$ there exists $n \in \mathbb{N}$ with $B \in \mathcal{C}_n$ and therefore, Theorem 39 yields $\mu_a(B) \leq \mu_b(B)$. Applying again Theorem 39, we hence finally derive $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. ☐

In case the set $X$ in the previous lemma is finite, the claim is trivial. Indeed, if $X$ is finite, there there exists $i \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ with $n \geq i$ it holds $R_i = R_n$. The latter observation directly yields the claim. In the general setting involving an uncountable Polish space $X$, however, it might hold $R_n \supsetneq R_{n+1}$ for every $n \in \mathbb{N}$ that causes the mathematical difficulties of the presented result.

There is an alternative proof for Lemma 123 relying on Lemma 3.11 in [Les10]. The cited lemma of [Les10] is completely analogous to our previous lemma, however, considers relations that are closed in the product space (rather than weakly smooth). It turns out that there is a suitable Polish topology on $X$ such that for every $n \in \mathbb{N}$ the set $R_n$ is closed in the corresponding product topology on $X \times X$ (see Chapter 13 in [Kec95]). As the this topology can be chosen in such a why that the induced Borel sigma algebras on $X$ concerning the original topology and the modified topology are the same, Lemma 123 is indeed implied by Lemma 3.11 in [Les10]. However, instead, we exploit the nice properties of weakly smooth relations to provide a direct argument circumventing the detour involving [Les10]. Moreover, our proof is orthogonal to [Les10] where one relies on results for converging sequences of probability measures on Polish spaces (see also Example 2 (6)).

*Proof of Theorem 120.* For every $n \in \mathbb{N}$ define the relation $R_n \subseteq Path \times Path$ as follows: for every paths $\pi_a = s_{a0}\, act_{a1}\, s_{a1} \ldots$ and $\pi_b = s_{b0}\, act_{b1}\, s_{b1} \ldots$,

$$\langle \pi_a, \pi_b \rangle \in R_n \quad \text{iff} \quad \langle \pi_{a|n}, \pi_{b|n} \rangle \in R^{\text{path},n}$$

where the finite paths $\pi_{a|n}, \pi_{b|n} \in Path_n$ are given as follows: in case $n = 0$ let $\pi_{a|n} = s_{a0}$ and $\pi_{a|n} = s_{b0}$ and in the other case where $n > 0$ let $\pi_{a|n} = s_{a0}\, act_{a1}\, s_{a1} \ldots act_{an}\, s_{an}$ and $\pi_{b|n} = s_{b0}\, act_{b1}\, s_{b1} \ldots act_{bn}\, s_{bn}$. We have $R_n \supseteq R_{n+1}$ for every $n \in \mathbb{N}$ as well as $R^{\text{path}} = \bigcap_{n \in \mathbb{N}} R_n$. Since the relation $R$ is required to be weakly smooth, for every $n \in \mathbb{N}$ it is easy to see that the relation $R_n$ is weakly smooth.

Let $\mathfrak{S}_a$ be a $\mu_a$-scheduler. Applying Theorem 77, there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that for all $n \in \mathbb{N}$,

$$\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b] \rangle \in (R^{\text{path},n})^{\text{wgt}}.$$

By Lemma 123, we can conclude the claim of theorem, i.e., the statement $\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R^{\text{path}})^{\text{wgt}}$, provided for every $n \in \mathbb{N}$ one has $\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R_n)^{\text{wgt}}$.

Let $n \in \mathbb{N}$ be a natural number. We show $\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R_n)^{\text{wgt}}$ now. Let $W_n$ be a weight function for $(Pr_n[\mathfrak{S}_a], R^{\text{path},n}, Pr_n[\mathfrak{S}_b])$. Abbreviate

$$X = (Act \times Sta)^{\omega}.$$

Relying on the Disintegration theorem (see Section 2.1), there exist two Borel functions $f_{a,n} \colon Path_n \to Prob[X]$ and $f_{b,n} \colon Path_n \to Prob[X]$ with

$$Pr[\mathfrak{S}_a] = Pr_n[\mathfrak{S}_a] \rtimes f_{a,n} \quad \text{and} \quad Pr[\mathfrak{S}_b] = Pr_n[\mathfrak{S}_b] \rtimes f_{b,n}.$$

Introduce the function $f_n\colon Path_n \times Path_n \to Prob[X \times X]$,

$$f_n(\hat{\pi}_a, \hat{\pi}_b) = f_{a,n}(\hat{\pi}_a) \otimes f_{b,n}(\hat{\pi}_a).$$

The function $f_n$ is Borel by Example 6. Moreover, introduce the Borel function $\xi_n$ with domain $Path_n \times Path_n \times X \times X$ and codomain $Path \times Path$ as follows: for every $\hat{\pi}_a, \hat{\pi}_b \in Path_n$ and $\tilde{\pi}_a, \tilde{\pi}_b \in X$ let

$$\xi_n(\hat{\pi}_a, \hat{\pi}_b, \tilde{\pi}_a, \tilde{\pi}_b) = \langle \hat{\pi}_a \diamond \tilde{\pi}_a, \hat{\pi}_b \diamond \tilde{\pi}_b \rangle.$$

Define the probability measure $W \in Prob[Path \times Path]$ by

$$W = (\xi_n)_\sharp (W_n \rtimes f_n).$$

We claim that $W$ is a weight function for $(Pr[\mathfrak{S}_a], R_n, Pr[\mathfrak{S}_b])$.

It is shown that $W$ is a coupling of $(Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b])$ first. Let $\hat{\Pi}_{a,n} \subseteq Path_n$ and $\tilde{\Pi}_a \subseteq X$ be Borel sets. One has

$$\xi_n^{-1}((\hat{\Pi}_{a,n} \times \tilde{\Pi}_a) \times Path) = \hat{\Pi}_{a,n} \times Path_n \times \tilde{\Pi}_a \times X$$

and therefore,

$$\begin{aligned}
& W((\hat{\Pi}_{a,n} \times \tilde{\Pi}_a) \times Path) \\
= & \int_{\hat{\Pi}_{a,n} \times \tilde{\Pi}_a} f_n(\hat{\pi}_a, \hat{\pi}_b)(\tilde{\Pi}_a \times X)\, dW_n(\hat{\pi}_a, \hat{\pi}_b) \\
= & \int_{\hat{\Pi}_{a,n} \times \tilde{\Pi}_a} f_{a,n}(\hat{\pi}_a)(\tilde{\Pi}_a)\, dW_n(\hat{\pi}_a, \hat{\pi}_b) \\
= & \int_{\hat{\Pi}_{a,n}} f_{a,n}(\hat{\pi}_a)(\tilde{\Pi}_a)\, dPr_n(\hat{\pi}_a) \\
= & Pr[\mathfrak{S}_a](\hat{\Pi}_{a,n} \times \tilde{\Pi}_a).
\end{aligned}$$

By Carathéodory extension theorem (see Section 2.1), for every Borel set $\Pi_a \subseteq Path$ it holds $W_n(\Pi_a \times Path) = Pr[\mathfrak{S}_a](\Pi_a \times Path)$. One analogously justifies the identity $W_n(Path \times \Pi_b) = Pr[\mathfrak{S}_b](Path \times \Pi_b)$ for every Borel set $\Pi_b \subseteq Path$. As a consequence, the probability measure $W$ is a coupling of $(Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b])$.

Thanks to Remark 38, the sets $R_n$ and $R^{\mathrm{path},n}$ are Borel in $Path \times Path$ and $Path_n \times Path_n$, respectively. To conclude that $W$ is a weight function for $(Pr[\mathfrak{S}_a], R_n, Pr[\mathfrak{S}_b])$, it remains to show $W(R_n) = 1$. Since

$$\xi_n^{-1}(R_n) = R^{\mathrm{path},n} \times X \times X,$$

we immediately obtain

$$W(R_n) = \int_{R^{\mathrm{path},n}} f_n(\hat{\pi}_a, \hat{\pi}_b)(X \times X)\, dW_n(\hat{\pi}_a, \hat{\pi}_b) = W_n(R^{\mathrm{path},n}) = 1.$$

Putting things together, $W$ is a weight function for $(Pr[\mathfrak{S}_a], R_n, Pr[\mathfrak{S}_b])$. □

## 5.6 Logical characterisation extended

Whereas the logical characterisation in Section 5.4 is restricted to the two temporal logics APCTL$_\bullet$ and APCTL$_\circ$, the following generalisation of Theorem 118 includes the comparable expressive logics $\exists$APCTL$^*$ and APCTL$^*$. Besides this, the proof goes beyond the presented techniques in [FKP17] that yielded the key ingredients in the former Section 5.4.

**Theorem 124.** *Let $\mathcal{T} = (Sta, Act, \to)$ be an STS, $\mathcal{A}$ be an action event family, and rew be a reward function. Assume that $\mathcal{T}$ is non-blocking, image-finite, and Borel concerning the hit sigma algebra. For every states $s_a, s_b \in Sta$ the following two chains of equivalences hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\mathrm{sou}} s_b$ *iff* $s_a \preceq_\bullet s_b$ *iff* $s_a \preceq^{\exists} s_b$.

(2) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$ *iff* $s_a \simeq_\circ s_b$ *iff* $s_a \simeq^{*} s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth.* ⌟

Statements (1) and (2) are accompanied with the same difficulty as both logics APCTL$^*$ and $\exists$APCTL$^*$ involve the exists modality ranging over schedulers. To be more precise, consider states $s_a, s_b \in Sta$ such that such that $s_a \preceq s_b$. We show $s_a \preceq^{\exists} s_b$ by an induction over the construction of formulas. Let $\tau$ be an APCTL$^*$ path-measure formula such that $s_a \models \exists\tau$. Consider an $s_a$-scheduler $\mathfrak{S}_a$ with $Pr[\mathfrak{S}_a] \models \tau$. To derive $s_b \models \exists\tau$, the task is to provide an $s_b$-scheduler $\mathfrak{S}_b$ such that $Pr[\mathfrak{S}_b] \models \tau$. To get a feeling for mathematical difficulty, the same discussions as in Chapter 4 can be conducted as schedulers are required to form Borel functions. Fortunately, Theorem 118 tells us that the simulation preorder yields a weakly smooth relation. Consequently, we can rely on Corollary 121 to obtain a convenient $s - b$scheduler. The complete proof of Theorem 124 can be found below.

**Proof of section's main result.** The following first two auxiliary lemmas extend simple observation for probability measures to the corresponding outer-measure functions.

**Lemma 125.** *Let $X$ be a measurable space, $\mu \in Prob[X]$ be a probability measure, and $M \subseteq X$ be an arbitrary set, and $B \subseteq X$ be a Borel set. If $\mu(B) = 1$, then it holds*

$$\mu^{\mathrm{out}}(M) = \mu^{\mathrm{out}}(M \cap B).$$

*Proof.* For every sets $M_1, M_2 \subseteq X$ the following inequality for outer measures is standard (see also Lemma 1.5.5 in [Bog07]):

$$|\mu^{\text{out}}(M_1) - \mu^{\text{out}}(M_2)| \leq \mu^{\text{out}}((M_1 \setminus M_2) \cup (M_2 \setminus M_1)).$$

Relying on $(M \setminus (B \cap M)) \cup ((B \cap M) \setminus M) = M \setminus B$, we hence obtain

$$|\mu^{\text{out}}(M) - \mu^{\text{out}}(M \cap B)| \leq \mu^{\text{out}}(M \setminus B).$$

Assuming $\mu(B) = 1$, it easily follows $\mu^{\text{out}}(M \setminus B) = 0$ as one has $M \setminus B \subseteq X \setminus B$ as well as $\mu(X \setminus B) = 0$. From this we finally derive the identity $\mu^{\text{out}}(M) = \mu^{\text{out}}(M \cap B)$ □

**Lemma 126.** *Let $X$ be a measurable space, $\mu_a, \mu_b \in Prob[X]$, and $R \subseteq X \times X$ be a relation such that $\langle \mu_a, \mu_b \rangle \in R^{\text{wgt}}$. Then the following two statements hold:*

(1) $\mu_a^{\text{out}}(M) \leq \mu_b^{\text{out}}(M)$ *for every upper $R$-stable sets $M \subseteq X$.*

(2) $\mu_a^{\text{out}}(M) = \mu_b^{\text{out}}(M)$ *for every $R$-stable sets $M \subseteq X$.*

*Proof.* Let $W$ be a weight function for $(\mu_a, R, \mu_b)$. First of all, we justify that for every sets $M_a, M_b \subseteq X$ the following two identities hold:

$$W^{\text{out}}(M_a \times X) = \mu_a^{\text{out}}(M_a) \quad \text{and} \quad W^{\text{out}}(X \times M_b) = \mu_b^{\text{out}}(M_b).$$

Let $M_a \subseteq X$ be a set. For every measurable set $B \subseteq X \times X$ with $M_a \times X \subseteq B$ there exists a measurable set $B_a \subseteq X$ such that $B = B_a \times X$ and $M_a \subseteq B_a$. This insight entails

$$\begin{aligned}
& W^{\text{out}}(M_a \times X) \\
= \; & \inf\{W(B) \; ; \; B \subseteq X \times X \text{ measurable set with } M_a \times X \subseteq B\} \\
= \; & \inf\{W(B_a \times X) \; ; \; B_a \subseteq X \text{ measurable set with } M_a \subseteq B_a\} \\
= \; & \inf\{\mu_a(B_a) \; ; \; B_a \subseteq X \text{ measurable set with } M_a \subseteq B_a\} \\
= \; & \mu_a^{\text{out}}(M_a).
\end{aligned}$$

One analogously shows the identity $W^{\text{out}}(X \times M_b) = \mu_b^{\text{out}}(M_b)$ for every $M_b \subseteq X$.

The remaining proof for the lemma is as in Remark 34. Nevertheless, let us recall the argument for statement (1). To this end let $M \subseteq X$ be an upper $R$-stable set. Moreover, suppose a measurable set $R' \subseteq X \times X$ with $W(R') = 1$ and $R' \subseteq R$. It obviously holds $R' \cap (M \times X) \subseteq X \times M$. According to Lemma 125, one moreover has $W^{\text{out}}(M \times X) = W^{\text{out}}(R' \cap (M \times X))$. Putting things together, we obtain

$$\mu_a^{\text{out}}(M) = W^{\text{out}}(M \times X) = W^{\text{out}}(R' \cap (M \times X)) \leq W^{\text{out}}(X \times M) = \mu_b^{\text{out}}(M).$$

This finishes the argumentation. □

The following lemma yields the key ingredient for our proof of Theorem 124.

**Lemma 127.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS, $\mathcal{A}$ be an action event family, and rew be a reward function. Then for every states $s_a, s_b \in Sta$ the following two statements hold:*

(1) *$s_a \preceq^{\exists} s_b$ if there is a simulation $R$ that is a weakly smooth relation with $\langle s_a, s_b \rangle \in R$.*

(2) *$s_a \simeq^{*} s_b$ if there is a bisimulation $R$ that is a smooth relation with $\langle s_a, s_b \rangle \in R$.*

*Proof.* We concentrate on statement (1) as the argument for claim (2) is completely analogous. For this purpose let $s_a, s_b \in Sta$ be states and $R$ be a simulation that is a weakly smooth relation with $\langle s_a, s_b \rangle \in R$. We show $s_a \preceq^{\exists} s_b$ by an induction over the construction $\exists$APCTL$^*$ state formulas. Relying on Corollary 121 and Lemma 126, it suffices to show that for every $\exists$APCTL$^*$ state formula $\sigma$ the set $[\![\sigma]\!]$ is upper $R$-stable, that for every $\exists$APCTL$^*$ path-measure formula $\tau$ the set $[\![\tau]\!]$ is upper $(R^{\mathrm{path}})^{\mathrm{wgt}}$-stable and that for every $\exists$APCTL$^*$ path formula $\upsilon$ the set $[\![\upsilon]\!]$ is upper $R^{\mathrm{path}}$-stable. Again, we proceed by an induction over the construction of formulas. In what follows we regard two selected cases.

Let $\tau$ be an $\exists$APCTL$^*$ path-measure formula such that the set $[\![\tau]\!]$ is upper $(R^{\mathrm{path}})^{\mathrm{wgt}}$-stable. Consider the $\exists$APCTL$^*$ state formula $\sigma = \exists \tau$. Let $\langle s_a, s_b \rangle \in R \cap ([\![\sigma]\!] \times Sta)$. It follows $s_a \models \sigma$ and hence, there exists an $s_a$-scheduler $\mathfrak{S}_a$ such that $Pr[\mathfrak{S}_a] \models \tau$. According to Corollary 121, there exists an $s_b$-scheduler $\mathfrak{S}_b$ with

$$\langle Pr[\mathfrak{S}_a], Pr[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path}})^{\mathrm{wgt}}.$$

Using that the set $[\![\tau]\!]$ is upper $(R^{\mathrm{path}})^{\mathrm{wgt}}$-stable and as we have $Pr[\mathfrak{S}_a] \models \tau$, we conclude $Pr[\mathfrak{S}_b] \models \tau$. It therefore follows $s_b \models \sigma$, i.e., one has $s_b \in [\![\sigma]\!]$. Consequently, the set $[\![\sigma]\!]$ is upper $R$-stable.

Let $\upsilon$ be an $\exists$APCTL$^*$ path formula such that the set $[\![\upsilon]\!]$ is upper $R^{\mathrm{path}}$-stable. Consider $q \in \mathbb{Q} \cap [0,1]$ and define the $\exists$APCTL$^*$ path-measure formula $\tau = \mathbb{P}_{>q}[\upsilon]$. Moreover, let $\langle \chi_a, \chi_b \rangle \in (R^{\mathrm{path}})^{\mathrm{wgt}} \cap ([\![\tau]\!] \times Prob[Path])$. We have $\chi_a \models \tau$, i.e., it holds $\chi_a^{\mathrm{out}}([\![\upsilon]\!]) > q$. Since the set $[\![\upsilon]\!]$ is upper $R^{\mathrm{path}}$-stable, Lemma 126 yields

$$q < \chi_a^{\mathrm{out}}([\![\upsilon]\!]) \leq \chi_b^{\mathrm{out}}([\![\upsilon]\!]).$$

We obtain $\chi_b \models \tau$ and thus, we can conclude that set $[\![\tau]\!]$ is $(R^{\mathrm{path}})^{\mathrm{wgt}}$-stable. □

*Proof of Theorem 124.* We concentrate on statement (1) as the argument for (2) is completely analogous. Let $s_a, s_b \in Sta$ be states. Clearly, $s_a \preceq^{\exists} s_b$ implies $s_a \preceq_{\bullet} s_b$. According to Theorem 118, it suffices to show that $s_a \preceq s_b$ implies $s_a \preceq^{\exists} s_b$. Assume $s_a \preceq s_b$. By Theorem 118, we also know that the relation $\preceq$ is weakly smooth. Hence, according to Lemma 127, we derive $s_a \preceq^{\exists} s_b$ (see also Remark 116). □

## 5.7 Logics for simple stochastic transition systems

In a simple STS $(Sta, Act, \rightarrow)$ every transition involves only one action (see Section 3.1), more precisely, for every transition $s \rightarrow \varphi$ there exists an action $act \in Act$ such that $\varphi(\{act\} \times Sta) = 1$. The developed results in this chapter, in particular, Corollary 119 and Theorem 124, clearly apply also for simple STSs. This section provides an adapted version of these main theorems where the assumed image finiteness can be slightly relaxed:

**Theorem 128.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a simple STS where the set Act is countable, $\mathcal{A}$ be an action event family, and rew be a reward function. Moreover, assume that $\mathcal{T}$ is non-blocking, point-wise image-finite, and Borel concerning the hit sigma algebra. Then for every states $s_a, s_b \in Sta$ the following two chains of equivalences hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\mathrm{sou}} s_b$ *iff* $s_a \preceq_\bullet s_b$ *iff* $s_a \preceq^{\exists} s_b$.

(2) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$ *iff* $s_a \simeq_\circ s_b$ *iff* $s_a \simeq^{*} s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth.* ⌟

Observe that the STS in the previous theorem may not be image finite. Indeed, the conditions on the STS cover the case where $Enabled[s]$ is countable for every state $s \in Sta$. Consequently, Theorem 128 yields an improvement of Theorem 124 for the subclass of simple STSs where the action space is countable. The proof steps and involved ideas for Theorem 128 are the same as for Theorem 124. Indeed, no essential new mathematical ideas are required for the argument. Nevertheless, some more details are summarised at the end of this section.

**Borel concerning the hit sigma algebra.** Before we devote ourselves to a proof of Theorem 128, let us investigate the property of being Borel concerning the hit sigma algebra in the context of simple STSs. Roughly speaking, a simple STS is Borel concerning the hit sigma algebra precisely when the corresponding condition in Definition 115 holds point-wise, i.e., for every individual action:

**Theorem 129.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a simple STS where the set Act is countable. Then the following two statements are equivalent:*

(1) *$\mathcal{T}$ is Borel concerning the hit sigma algebra.*

(2) *For every act $\in$ Act and every open set $O_{Sta} \subseteq Prob[Sta]$ the following set is Borel in Sta,*

$$\{s \in Sta \,;\, Enabled[s, act] \cap O_{Sta} \neq \varnothing\}.$$

*Proof.* Our argument below requires some additional insights on the Polish topology on $Prob[X]$ where $X$ is a Polish space (see also Example 2 (6)). We summarise this auxiliary material first.

Let $X$ be a Polish space. For every closed set $C \subseteq X$ the set $\{\mu \in Prob[X] \; ; \; \mu(C) < 1\}$ is open (see Corollary 15.6 in [AB06]). Besides this, as we have a metric on $Prob[X]$, we can safely speak about convergent sequences of probability measures. Let $\mu \in Prob[X]$ be a probability measure and $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures $\mu_n \in Prob[X]$. According to the Portmanteau theorem (see Theorem 17.20 in [Kec95]), the sequences $(\mu_n)_{n\in\mathbb{N}}$ converges in $Prob[X]$ to the limit $\mu$ precisely when for every closed set $C \subseteq X$ it holds $\limsup_{n\in\mathbb{N}} \mu_n(C) \leq \mu(C)$.

Let $Y$ is a Polish space and $f\colon X \to Y$ a continuous function. Then the pushforward function $f_\sharp$ is continuous. This can be shown by an application of the Portmanteau theorem. Indeed, it is easy to see that for every convergent sequence $(\mu_n)_{n\in\mathbb{N}}$ in $Prob[X]$ with limit $\mu \in Prob[X]$ the sequence $(f_\sharp(\mu_n))_{n\in\mathbb{N}}$ converges in $Prob[Y]$ to the limit $f_\sharp(\mu)$.

(1) *implies* (2). Let $act \in Act$ be an action and $O_{Sta} \subseteq Prob[Sta]$ be an open set. To obtain the implication from (1) to (2), it suffices to provide an open set $O_{act} \subseteq Prob[Act \times Sta]$ such that the following identity holds:

$$\{s \in Sta \; ; \; Enabled[s, act] \cap O_{Sta} \neq \varnothing\} = \{s \in Sta \; ; \; Enabled[s] \cap O_{act} \neq \varnothing\}.$$

Recall, as the Polish space $Act$ is equipped with the discrete topology (see Definition 2 (1)), the set $Act \setminus \{act\}$ is closed in $Act$. Hence, the set $O_{Act,act}$ is open in $Prob[Act]$ where

$$O_{Act,act} = \{\alpha \in Prob[Act] \; ; \; \alpha(Act \setminus \{act\}) < 1\}.$$

Define the continuous functions $\xi_{Act}\colon Act \times Sta \to Act$, $\xi_{Act}(act', s) = act'$ as well as $\xi_{Sta}\colon Act \times Sta \to Sta$, $\xi_{Sta}(act', s) = s$. Hence, the corresponding pushforward functions $(\xi_{Act})_\sharp$ and $(\xi_{Sta})_\sharp$ are continuous. It follows that the set $O_{act}$ is open in $Prob[Act \times Sta]$ where we define

$$O_{act} = (\xi_{Act})_\sharp^{-1}(O_{Act,act}) \cap (\xi_{Sta})_\sharp^{-1}(O_{Sta}).$$

In order to see that for every state $s \in Sta$ it holds $Enabled[s, act] \cap O_{Sta} \neq \varnothing$ iff $Enabled[s] \cap O_{act} \neq \varnothing$, it suffices to observe the following statement. Let $act' \in Act$ be an action and $\mu \in Prob[Sta]$ be a probability measure. Define $\varphi \in Prob[Act \times Sta]$ by $\varphi = Dirac[act'] \otimes \mu$. Then we have that

$$\varphi \in Enabled[s] \cap O_{act} \quad \text{iff} \quad \mu \in Enabled[s, act] \cap O_{Sta} \text{ and } act = act'.$$

Indeed, the latter equivalence follows immediately from the observation that it holds $\mathit{Dirac}[\mathit{act}'] \in O_{\mathit{Act},\mathit{act}}$ precisely when $\mathit{act} = \mathit{act}'$.

(2) *implies* (1). Let $\mathit{act} \in \mathit{Act}$. Define the function $\xi_{\mathit{act}} \colon \mathit{Prob}[\mathit{Sta}] \to \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$,

$$\xi_{\mathit{act}}(\mu) = \mathit{Dirac}[\mathit{act}] \otimes \mu.$$

We show that the function $\xi_{\mathit{act}}$ is continuous first. To this end let $\mu \in \mathit{Prob}[\mathit{Sta}]$ as well as $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures $\mu_n \in \mathit{Prob}[\mathit{Sta}]$ that converges in $\mathit{Prob}[\mathit{Sta}]$ to the limit $\mu$. Let $C \subseteq \mathit{Act} \times \mathit{Sta}$ be a closed set. Then it is easy to see that $\mathit{Section}[C, \mathit{act}, \cdot]$ is closed in $\mathit{Sta}$. Using that the sequence $(\mu_n)_{n\in\mathbb{N}}$ converges to $\mu$, the Portmanteau theorem yields the inequality

$$\limsup_{n\in\mathbb{N}} \mu_n(\mathit{Section}[C, \mathit{act}, \cdot]) \leq \mu(\mathit{Section}[C, \mathit{act}, \cdot]).$$

Applying again the Portmanteau theorem and as $\xi_{\mathit{act}}(\mu_n)(C) = \mu_n(\mathit{Section}[C, \mathit{act}, \cdot])$ for every $n \in \mathbb{N}$ and $\xi_{\mathit{act}}(\mu)(C) = \mu(\mathit{Section}[C, \mathit{act}, \cdot])$, the sequence $(\xi_{\mathit{act}}(\mu_n))_{n\in\mathbb{N}}$ converges in $\mathit{Prob}[Y]$ to the limit $\xi_{\mathit{act}}(\mu)$. From this we derive that the function $\xi_{\mathit{act}}$ is indeed continuous.

Let $O \subseteq \mathit{Prob}[\mathit{Act} \times \mathit{Sta}]$ be an open set. Define the set $S \subseteq \mathit{Sta}$ by

$$S = \bigcup_{\mathit{act}' \in \mathit{Act}} \{s \in \mathit{Sta} \,;\, \mathit{Enabled}[s, \mathit{act}'] \cap (\xi_{\mathit{act}'})^{-1}(O) \neq \varnothing\}.$$

For every $\mathit{act}' \in \mathit{Act}$ the set $(\xi_{\mathit{act}'})^{-1}(O)$ is open in $\mathit{Sta}$. Assuming statement (2) holds and as the set $\mathit{Act}$ is countable, it hence follows that the set $S$ is Borel in $\mathit{Sta}$. Using that the STS $\mathcal{T}$ is simple, we moreover have the identity

$$S = \{s \in \mathit{Sta} \,;\, \mathit{Enabled}[s] \cap O \neq \varnothing\}.$$

From this it directly follows that the STS $\mathcal{T}$ is Borel concerning the hit sigma algebra. □

**Corollary 130.** *Let $\mathcal{T} = (\mathit{Sta}, \mathit{Act}, \to)$ be a simple STS where the set Act is countable and such that for every $s \in \mathit{Sta}$ and $\mathit{act} \in \mathit{Act}$ the set $\mathit{Enabled}[s, \mathit{act}]$ is closed in $\mathit{Prob}[\mathit{Sta}]$. Moreover, assume that $\mathcal{T}$ is Borel concerning the hit sigma algebra. Then $\mathcal{T}$ is Souslin.*

*Proof.* For every action $\mathit{act} \in \mathit{Act}$ define the set-valued $F_{\mathit{act}} \colon \mathit{Sta} \rightsquigarrow \mathit{Prob}[\mathit{Sta}]$,

$$F_{\mathit{act}}(s) = \mathit{Enabled}[s, \mathit{act}].$$

As for every state $s \in Sta$ and action $act \in Act$ the set $Enabled[s, act]$ is closed in $Prob[Sta]$, Theorem 129 together with Remark 23 yield that for every action $act \in Act$ the set $Rel[F_{act}]$ is Souslin in $Sta \times Prob[Sta]$.

For every $act \in Act$ define the function $\xi_{act} \colon Sta \times Prob[Sta] \to Sta \times Prob[Act \times Sta]$,

$$\xi_{act}(s, \mu) = \langle s, Dirac[act] \otimes \mu \rangle.$$

For every $act \in Act$ the set $\xi_{act}(Rel[F_{act}])$ is Borel in $Sta \times Prob[Act \times Sta]$ according to Example 6 and Remark 10 (5). Since

$$\to = \bigcup_{act \in Act} \xi_{act}(Rel[F_{act}]),$$

the set $\to$ is Souslin in $Sta \times Prob[Act \times Sta]$ applying Remark 10 (4). Consequently, the STS under consideration is Souslin. $\square$

**Corollary 131.** *Let $\mathcal{T} = (Sta, Act, \to)$ be a simple STS where the set Act is countable. Assume that $\mathcal{T}$ is non-blocking, point-wise image-finite, and Borel concerning the hit sigma algebra. For every $s_a, s_b \in Sta$ the two statements below hold:*

(1) $s_a \preceq s_b$ *implies* $s_a \leq^{\mathrm{tr}} s_b$.

(2) $s_a \simeq s_b$ *implies* $s_a =^{\mathrm{tr}} s_b$.

*Proof.* The following argument is basically the same as for Corollary 119. By Remark 4, we can safely assume some action event family on *Act*. According to Corollary 130, the STS $\mathcal{T}$ is Souslin. Hence, Theorem 128 and Corollary 83 yield the claim. $\square$

**Proof of section's main result.** As we have already mentioned, a proof of Theorem 128 is analogous to our argument for Theorem 124. The following material summarises some intermediary steps. Our first lemma adapts Theorem 117:

**Lemma 132.** *Let $\mathcal{T} = (Sta, Act, \to)$ be a simple STS where the set Act is countable and $\mathcal{A}$ be an action event family. Moreover, assume that $\mathcal{T}$ is non-blocking, point-wise image-finite, and Borel concerning the hit sigma algebra. Then the following two statements hold:*

(1) *The set $[\![\sigma]\!]$ is Borel in Sta for every APCTL$_\bullet$ state formula $\sigma$.*

(2) *The set $[\![\sigma]\!]$ is Borel in Sta for every APCTL$_\circ$ state formula $\sigma$.*

*Proof.* The following argumentation is a simple adaption of our proof for Theorem 117. Thanks to Remark 19 and Theorem 129, for every action $act \in Act$ there exists a family $(f_{act,n})_{n\in\mathbb{N}}$ of Borel functions $f_{act,n} \colon Sta \to Prob[Sta]$ such that for every $s \in Sta$,

$$Enabled[s, act] = \{f_{act,n}(s) \; ; \; n \in \mathbb{N}\}.$$

For every $act \in Act$ and $n \in \mathbb{N}$ define $f'_{act,n} \colon Sta \to Prob[Act \times Sta]$,

$$f'_{act,n}(s) = Dirac[act] \otimes f_{act,n}(s).$$

By Example 6, for every $act \in Act$ and $n \in \mathbb{N}$ the function $f'_{act,n}$ is Borel.

Let $k \in \mathbb{N}$ and define the APCTL$_\bullet$ state formula $\sigma$ by

$$\sigma \quad = \quad \exists(\mathbb{P}_{>q_0}[v_0] \;\wedge\; \ldots \;\wedge\; \mathbb{P}_{>q_k}[v_k])$$

where for every $i \in \{0, \ldots, k\}$, $q_i$ is a rational number in $[0, 1]$ and $v_i$ is the APCTL$_\bullet$ path-measure formula given by

$$v_i \quad = \quad \bigcirc(A_{i0} \wedge \sigma_{i0}) \;\vee\; \ldots \;\vee\; \bigcirc(A_{ik_i} \wedge \sigma_{ik_i})$$

for some natural number $k_i \in \mathbb{N}$, $A_{i0}, \ldots, A_{ik_i} \in Bool[\mathcal{A}]$ and APCTL$_\bullet$ state formula $\sigma_{i0}, \ldots, \sigma_{ik_i}$. Assume that for all $i \in \{0, \ldots, k\}$ the sets $[\![\sigma_{i0}]\!], \ldots, [\![\sigma_{ik_i}]\!]$ are Borel in $Sta$. Inspecting the syntax of APCTL$_\bullet$ state formulas and relying on an induction over the construction of APCTL$_\bullet$ state formulas, it suffices to show that the set $[\![\sigma]\!]$ is Borel in $Sta$ in order to conclude the lemma.

For every $act \in Act$, $n \in \mathbb{N}$ and $i \in \{0, \ldots, k\}$ define the function $g_{act,ni} \colon Sta \to [0, 1]$,

$$g_{act,ni}(s) = f'_{act,n}(s)((A_{i0} \times [\![\sigma_{i0}]\!]) \cup \ldots \cup (A_{ik_i} \times [\![\sigma_{ik_i}]\!])).$$

Inspecting Example 114, it is easy to see that

$$[\![\sigma]\!] = \bigcup_{act\in Act} \bigcup_{n\in\mathbb{N}} \bigcap_{i\in\{0,\ldots,k\}} (g_{act,ni})^{-1}((q_i, 1]).$$

Using the techniques of Remark 5, for every $act \in Act$, $n \in \mathbb{N}$, and $i \in \{0, \ldots, k\}$ the function $g_{act,ni}$ is Borel and therefore, it directly follows that the set $[\![\sigma]\!]$ is Borel in $Sta$. This shows statement (1). Claim (2) follows analogously. $\square$

The following lemma is in the spirit of Theorem 118:

**Lemma 133.** *Let $\mathcal{T} = (Sta, Act, \to)$ be a simple STS where the set Act is countable and $\mathcal{A}$ be an action event family. Moreover, assume that $\mathcal{T}$ is non-blocking, point-wise image-finite, and Borel concerning the hit sigma algebra. For every states $s_a, s_b \in Sta$ the following two statements hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\mathrm{sou}} s_b$ *iff* $s_a \preceq_\bullet s_b$.

(2) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$ *iff* $s_a \simeq_\circ s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth.*

*Proof.* Using Lemma 132, the argument is exactly the same as for Theorem 118 (see also [FKP17]). Again, the key ingredients are given by Theorems 39 and 40. $\square$

## 5.8 Expected values of payoff functions

Intuitively, a payoff function assigns a real number to every trace of an STS. Thus, whereas a logical formula either hold in a state or not, expected values of payoff functions provide quantitative information for a state. In what follows we shows that also these quantitative properties are preserved by the simulation preorder and the bisimulation equivalence under the Souslin requirements under consideration. As a consequence, the modal logics in this chapter can be extended with an expectation operator such that Theorems 124 and 128 are still maintained.

**Definition 134.** Let $\mathcal{T} = (Sta, Act, \to)$ be an STS. A *payoff function* (*for* $\mathcal{T}$) is a Borel function

$$pay\colon Act^\omega \to \mathbb{R} \cup \{-\infty, +\infty\}.$$

Let *pay* be a payoff function. For every probability measure $\mu \in Prob[Sta]$ and $\mu$-scheduler $\mathfrak{S}$ the $\mathfrak{S}$-*expected payoff* (*of pay*) is defined by

$$Ex[\mathfrak{S}, pay] = \int pay(\sigma)\, dPrTrace[\mathfrak{S}](\sigma),$$

i.e., the $\mathfrak{S}$-expected payoff is given by the expected value of the function *pay* with respect to the trace distribution $PrTrace[\mathfrak{S}]$. ⌟

We remark that the $\mathfrak{S}$-expected payoff of a payoff function *pay* might be $-\infty$ or $+\infty$. However, it holds $-\infty < Ex[\mathfrak{S}, pay] < +\infty$ precisely when the function *pay* is integrable with respect to the probability measure $PrTrace[\mathfrak{S}]$. For instance, provided there are real numbers $r_l, r_u \in \mathbb{R}$ such that $PrTrace[\mathfrak{S}](\{\sigma \in Act^\omega ; r_l \leq pay(\sigma) \leq r_u\}) = 1$, the function *pay* is integrable with respect to the probability measure $PrTrace[\mathfrak{S}]$, in particular, it follows $-\infty < Ex[\mathfrak{S}, pay] < +\infty$. For instance, provided the STS under consideration is augmented with a reward function, the accumulated reward until reaching a goal or the long-run average (mean payoff) of an infinite path yield natural examples for payoff functions. The following example of a payoff function does not involve any reward function:

**Example 135.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an STS and assume that response is an action contained in *Act*. Introduce the function $pay\colon Act^\omega \rightarrow \mathbb{R} \cup \{-\infty+, \infty\}$ as follows: for every $\sigma = act_1\, act_2 \ldots$ define

$$pay(\pi) = \sup\{\inf\{j \in \mathbb{N} ; act_{1+i+j} = \text{response}\} ; i \in \mathbb{N}\}.$$

Consider a state $s$ and an $s$-scheduler $\mathfrak{S}$ with the following property: the set consisting of every trace $act_1\, act_2 \ldots \in Act^\omega$ such that $act_i = \text{response}$ for infinitely many $i \in \mathbb{N} \setminus \{0\}$ has probability one with respect to the trace distribution $PrTrace[\mathfrak{S}]$. The value $Ex[\mathfrak{S}, pay]$ intuitively represents the expected maximal number of steps between the response action if the STS under consideration is governed by the scheduler $\mathfrak{S}$. ⌟

**Proposition 136.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a Souslin STS, pay be a payoff function, $R$ be a Souslin simulation, and $\mu_a, \mu_b \in Prob[Sta]$ be so that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Then for every $\mu_a$-scheduler $\mathfrak{S}_a$ there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that*

$$Ex[\mathfrak{S}_a, pay] = Ex[\mathfrak{S}_b, pay].$$

*Proof.* Let $\mathfrak{S}_a$ be a $\mu_a$-scheduler. By Corollary 78, there exists a $\mu_b$-scheduler $\mathfrak{S}_b$ such that for every $n \in \mathbb{N}$ it holds

$$\langle Pr_n[\mathfrak{S}_a], Pr_n[\mathfrak{S}_b] \rangle \in (R^{\mathrm{path},n})^{\mathrm{wgt}}.$$

By Lemma 81, it immediately follows $PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b]$. This insight directly yields the claim. □

# 6 Parallel composition based on spans and couplings

We present a generic parallel-composition operator for simple STSs based on spans and couplings. Our operator only refers to simple STSs for the same reasons as discussed in Section 4.3.3 of [Seg95]. The new 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.

The assumption that STSs to be composed behave stochastically independent of each other is not adequate in every situation. We illustrate this continuing the running example from Section 3.1 modelling a simple cooling system of a server. The server can operate in different energy modes. For instance, hardware components can be shut down to save energy costs in phases of low workloads and can be set in a high performance mode in busy phases. Thus, the energy consumption can be adapted to meet all the service level agreements while saving energy cost.

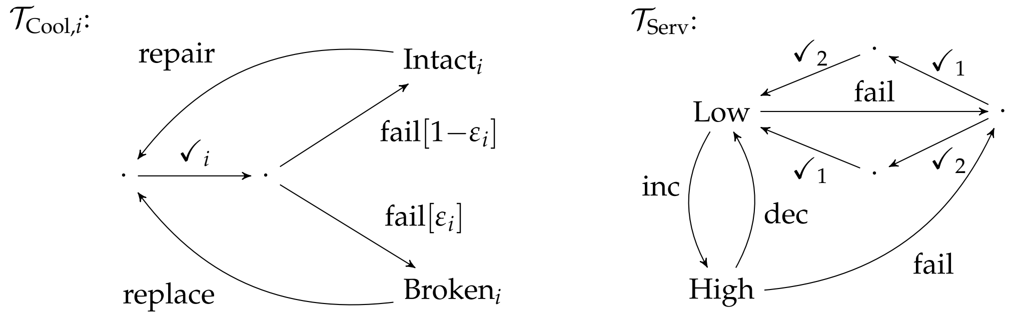


Figure 6.1: Cooling systems $\mathcal{T}_{\mathrm{Cool},1}$ and $\mathcal{T}_{\mathrm{Cool},2}$ of a server $\mathcal{T}_{\mathrm{Serv}}$.

Consider the STSs depicted in Figure 6.1 modelling two cooling systems $\mathcal{T}_{\mathrm{Cool},1}$ and $\mathcal{T}_{\mathrm{Cool},2}$ dissipating the heat produced by the server $\mathcal{T}_{\mathrm{Serv}}$. A failure in the systems affects all the components at the same time, which is formalised by means of the synchronisation action fail. Intuitively, the action fail models a common cause failure, e.g., triggered by power fluctuations. When fail is executed, the components are interrupted immediately such that the cooling system $\mathcal{T}_{\mathrm{Cool},i}$ needs to be replaces with probability $\varepsilon_i$. The cooling system $\mathcal{T}_{\mathrm{Cool},i}$ can be repaired with probabilities $1 - \varepsilon_i$.

The behaviour of the cooling systems $\mathcal{T}_{\mathrm{Cool},1}$ and $\mathcal{T}_{\mathrm{Cool},2}$ is coupled by the internal power and cooling strategy of the server $\mathcal{T}_{\mathrm{Serv}}$. In case the server operates in the energy-saving state labelled by Low, the server typically stresses only one cooling system and internally alternates between them. The latter ensures that the cooling components wearing out equally during a certain time span that in turn lowers maintenance costs. As a consequence, the common failure affects either $\mathcal{T}_{\mathrm{Cool},1}$ or $\mathcal{T}_{\mathrm{Cool},2}$ depending on which device is actually working. We can thus rely on the assumption that the cooling systems behave stochastically independent within a failure in the energy-saving mode, which is formalised by

$$[\mathrm{fail}]\colon\ \mathrm{Low} \implies \mathbb{P}(\mathrm{Broken}_1 \wedge \mathrm{Broken}_2) = \mathbb{P}(\mathrm{Broken}_1) \cdot \mathbb{P}(\mathrm{Broken}_2).$$

In contrast, if the server operates in the energy-consuming state labelled by High, the cooling systems are stressed equally to dissipate the increased heat. Here, it is highly probable that a failure affects the cooling systems equally. More precisely, either both cooling systems need to be replaced or none of them with high probability: with a probability greater than 0.9, either both cooling systems are broken or both are intact, i.e.,

$$[\mathrm{fail}]\colon\ \mathrm{High} \implies \mathbb{P}((\mathrm{Broken}_1 \wedge \mathrm{Broken}_2) \vee (\mathrm{Intact}_1 \wedge \mathrm{Intact}_2)) > 0.9.$$

In our example we abstract from the precise operational behaviour of the internal power-management strategy of the server controlling the interplay of the cooling systems. On the one hand, we have only vague knowledge on the precise implementation of this internal strategy and on the other hand, we want to abstract from the complex internal behaviour to keep the model simple and manageable. Therefore, stochastic dependencies caused by the power-management strategy are formalised in a declarative manner by the above two symbolic expressions for the action fail.

The main feature of the span-coupling composition operator to be developed in this chapter is that stochastic dependencies between components can be declaratively specified by specific couplings of probability measures. In contrast to the standard composition operator, the new operator takes all the possible couplings between components into account instead of considering only the independent one. Thus, the new operator does not rely on the assumption that components interact stochastically independent. To the best of our knowledge, couplings have not been used as a declarative modelling formalism in a compositional framework for operational systems.

Besides couplings, our composition operator relies on spans to characterise the global state space of the STSs to be composed. Spans allow for arbitrary sets and associated projections functions as well as induce a generic notion for couplings. As a consequence, our composition operator in particular covers stochastic transition systems with shared variables. As a consequence, our framework can be instantiated in several ways so that standard composition operators for stochastic systems are covered, e.g., [Seg95, D'A99, DK05, CSKN05, Cat05]. Moreover, compatibility results with composition operators for modelling formalisms from the literature such as probabilistic timed and rectangular hybrid automata [Spr01, KNSS02, Spr11, ZSR$^+$12, Spr15] can be obtained.

As a main result of this chapter, we show Theorem F (see Chapter 1) stating that simulation preorder and bisimulation equivalence are congruences with respect to our newly developed composition operator under some side constraints. The mathematical challenge associated to the proof is to construct appropriate probability measures that yield a certain span coupling and that are compatible with a specific relation. The key feature of our proof is given by the disintegration theorem (see Section 2.1), which, roughly speaking, makes the coupling structure of probability measures on a product space accessible from a mathematical point of view.

## 6.1 Standard compositional framework

We provide a brief introduction on the standard compositional framework for simple STSs where probability measures occurring in STSs to be composed are put together by means of their product measure. After that, we present a first easy congruence result with respect to the standard composition operator. The following material contains no new contributions and can be seen as foundations for further considerations in this chapter on a general compositional framework.

**Definition 137.** Two Polish spaces $Act_1$ and $Act_2$ are called *composable* provided their union $Act_1 \cup Act_2$ constitutes a Polish space when equipped with the union topology. Recall, a subset $O \subseteq Act_1 \cup Act_2$ is open with respect to the union topology iff both sets $O \cap Act_1$ and $O \cap Act_2$ are open in $Act_1$ and $Act_2$, respectively. We refer to two STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ as *composable* provided the respective action spaces are composable. $\lrcorner$

The previous definition is motivated by the fact that the action space in a composition is obtained by the union of the respective action spaces of the STSs to be composed. The requirement in the definition ensures that the action space of the resulting STS constitutes a

Polish space. Recall, Example 2 (4) shows that the union of disjoint Polish spaces is Polish. However, the Polish spaces $Act_1$ and $Act_2$ in Definition 137 are not necessarily disjoint. For practical applications the requirement in Definition 137 is harmless that is also illustrated by the following example:

**Example 138.** Let $Q_1$ and $Q_2$ be countable sets disjoint from $\mathbb{R}_{\geq 0}$. Then $Act_1 = Q_1 \cup \mathbb{R}_{\geq 0}$ and $Act_2 = Q_2 \cup \mathbb{R}_{\geq 0}$ form Polish spaces by Example 2 (4). Moreover, it is easy to see that $Act_1$ and $Act_2$ are composable. Indeed, as the sets $Q_1$ and $Q_2$ are countable, we immediately obtain that the Polish spaces $Q_1$ and $Q_2$ are composable. Since the union $Q_1 \cup Q_2$ is also disjoint from $\mathbb{R}_{\geq 0}$, the claim follows from Example 2 (4). ⌟

**Definition 139.** Let $\mathcal{T}_1 = (Sta_1, Act_1, \rightarrow_1)$ and $\mathcal{T}_2 = (Sta_2, Act_2, \rightarrow_2)$ be simple STSs and $Sync \subseteq Act_1 \cap Act_2$ be a set. If $\mathcal{T}_1$ and $\mathcal{T}_2$ are composable, the *(standard) composition of* $(\mathcal{T}_1, \mathcal{T}_2)$ *concerning Sync* is given by the STS

$$\mathcal{T}_1 \parallel_{\otimes, Sync} \mathcal{T}_2 = (Sta_1 \times Sta_2, Act_1 \cup Act_2, \rightarrow)$$

whose transition relation $\rightarrow$ is defined by the following three inference rules:

(1) For every $act_1 \in Act_1 \setminus Sync$, $s_1 \in Sta_1$, $s_2 \in Sta_2$, and $\mu_1 \in Prob[Sta_1]$:

$$\frac{s_1 \rightarrow_1 \langle act_1, \mu_1 \rangle}{\langle s_1, s_2 \rangle \rightarrow \langle act_1, \mu_1 \otimes Dirac[s_2] \rangle}.$$

(2) For every $act_2 \in Act_2 \setminus Sync$, $s_1 \in Sta_1$, $s_2 \in Sta_2$, and $\mu_2 \in Prob[Sta_2]$:

$$\frac{s_2 \rightarrow_2 \langle act_2, \mu_2 \rangle}{\langle s_1, s_2 \rangle \rightarrow \langle act_2, Dirac[s_1] \otimes \mu_2 \rangle}.$$

(3) For every $act \in Sync$, $s_1 \in Sta_1$, $s_2 \in Sta_2$, $\mu_1 \in Prob[Sta_1]$, and $\mu_2 \in Prob[Sta_2]$:

$$\frac{s_1 \rightarrow_1 \langle act, \mu_1 \rangle \quad \text{and} \quad s_2 \rightarrow_2 \langle act, \mu_2 \rangle}{\langle s_1, s_2 \rangle \rightarrow \langle act, \mu_1 \otimes \mu_2 \rangle}.$$

⌟

When composing two STSs by means of the operator introduced in the previous definition, we implicitly assume that the considered STSs are composable. Actions in $Act_1 \setminus Sync$ and $Act_2 \setminus Sync$ are viewed as local and can be taken by the respective STS in an autonomous fashion (see inference rules (1) and (2)). In particular, an execution of an action contained in $Act_1 \setminus Sync$ does not affect the local state of the STS $\mathcal{T}_2$ and accordingly for the actions in $Act_2 \setminus Sync$. In contrast to that, those action that are contained in the set $Sync$ are synchronisation actions meaning that they need to be performed by $\mathcal{T}_1$ and $\mathcal{T}_2$ at the same time (see inference rule (3)). Involved distributions over the corresponding state spaces are combined using the respective product measures. Consequently, if $S_1 \subseteq Sta_1$ and $S_2 \subseteq Sta_2$ are Borel sets, the two events that the STS $\mathcal{T}_1$ enters a state in $S_1$ and that the STS $\mathcal{T}_2$ enters a state in $S_2$ are stochastically independent from each other with respect to every probability measure occurring in the standard composition of $(\mathcal{T}_1, \mathcal{T}_2)$.

**Definition 140.** Let $\mathcal{T}_a = (Sta_a, Act, \rightarrow_a)$ and $\mathcal{T}_b = (Sta_b, Act, \rightarrow_b)$ be simple STSs such that their respective action spaces are the same. A *simulation for* $(\mathcal{T}_a, \mathcal{T}_b)$ is a relation $R \subseteq Sta_a \times Sta_b$ with the following property: for every states $s_a \in Sta_a$ and $s_b \in Sta_b$, action $act \in Act$, and probability measure $\mu_a \in Prob[Sta_a]$ with

$$\langle s_a, s_b \rangle \in R \quad \text{and} \quad s_a \rightarrow_a \langle act, \mu_a \rangle$$

there exists a probability measure $\mu_b \in Prob[Sta_b]$ such that

$$s_b \rightarrow_b \langle act, \mu_b \rangle \quad \text{and} \quad \langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}.$$

A relation $R \subseteq Sta_a \times Sta_b$ is called a *bisimulation for* $(\mathcal{T}_a, \mathcal{T}_b)$ provided the relation $R$ is a simulation for $(\mathcal{T}_a, \mathcal{T}_b)$ and its inverse relation $R^{-1}$ is a simulation for $(\mathcal{T}_b, \mathcal{T}_a)$. ⌟

In the previous chapters simulation and bisimulation has been considered as a relation between states of one single STSs (see, e.g., Definition 47 as well as Remark 49). Compared with this, Definition 140 introduces these concepts as a relation between STSs such that the branching-time behaviour of different STSs can be compared. In order to achieve a coherent comparison, the STSs are assumed to involve the same action spaces. For two simple STSs $\mathcal{T}_a = (Sta_a, Act, \rightarrow_a)$ and $\mathcal{T}_b = (Sta_b, Act, \rightarrow_b)$ as well as corresponding states $s_a \in Sta_a$ and $s_b \in Sta_b$ we furthermore define

$$\begin{aligned} \mathcal{T}_a, s_a \preceq \mathcal{T}_b, s_b \quad &\text{iff} \quad \text{there is a simulation } R \text{ for } (\mathcal{T}_a, \mathcal{T}_b) \text{ with } \langle s_a, s_b \rangle \in R, \\ \mathcal{T}_a, s_a \simeq \mathcal{T}_b, s_b \quad &\text{iff} \quad \text{there is a bisimulation } R \text{ for } (\mathcal{T}_a, \mathcal{T}_b) \text{ with } \langle s_a, s_b \rangle \in R. \end{aligned}$$

The following result shows that simulation and bisimulation are congruences with respect to the standard composition operator for STS:

**Proposition 141.** *Consider four STSs $\mathcal{T}_{a1}$, $\mathcal{T}_{b1}$, $\mathcal{T}_{a2}$, and $\mathcal{T}_{b2}$ and denote the respective state spaces by $Sta_{a1}$, $Sta_{b1}$, $Sta_{a2}$, and $Sta_{b2}$. Assume that the action space of $\mathcal{T}_{a1}$ and $\mathcal{T}_{b1}$ are given by the*

*Polish space* $Act_1$ *and that the action space of* $\mathcal{T}_{a2}$ *and* $\mathcal{T}_{b2}$ *are given by the Polish space* $Act_2$*. Let* $Sync \subseteq Act_1 \cap Act_2$ *and define the two STSs*

$$\mathcal{T}_a = \mathcal{T}_{a1} \|_{\otimes,Sync} \mathcal{T}_{a2} \quad \textit{and} \quad \mathcal{T}_b = \mathcal{T}_{b1} \|_{\otimes,Sync} \mathcal{T}_{b2}.$$

*For every* $s_{a1} \in Sta_{a1}$*,* $s_{b1} \in Sta_{b1}$*,* $s_{a2} \in Sta_{a2}$*, and* $s_{b2} \in Sta_{b2}$ *the two statements below hold:*

(1) $\mathcal{T}_{a1}, s_{a1} \preceq \mathcal{T}_{b1}, s_{b1}$ *and* $\mathcal{T}_{a2}, s_{a2} \preceq \mathcal{T}_{b2}, s_{b2}$ *implies* $\mathcal{T}_a, \langle s_{a1}, s_{a2}\rangle \preceq \mathcal{T}_b, \langle s_{b1}, s_{b2}\rangle$.

(2) $\mathcal{T}_{a1}, s_{a1} \simeq \mathcal{T}_{b1}, s_{b1}$ *and* $\mathcal{T}_{a2}, s_{a2} \simeq \mathcal{T}_{b2}, s_{b2}$ *implies* $\mathcal{T}_a, \langle s_{a1}, s_{a2}\rangle \simeq \mathcal{T}_b, \langle s_{b1}, s_{b2}\rangle$.

*Proof.* The argument is standard. Nevertheless, let us illustrate a proof for the two claims. Let $R_1$ be a simulation for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $R_2$ be a simulation for $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$. Define the Polish spaces $Sta_a = Sta_{a1} \times Sta_{a2}$ and $Sta_b = Sta_{b1} \times Sta_{b2}$. Introduce the relation

$$R = \{\langle\langle s_{a1}, s_{a2}\rangle, \langle s_{b1}, s_{b2}\rangle\rangle \in Sta_a \times Sta_b \; ; \; \langle s_{a1}, s_{b1}\rangle \in R_1 \text{ and } \langle s_{a2}, s_{b2}\rangle \in R_2\}.$$

Then $R$ is a simulation for $(\mathcal{T}_a, \mathcal{T}_b)$. This follows easily from the following observation. Let $\mu_{a1} \in Prob[Sta_{a1}]$, $\mu_{b1} \in Prob[Sta_{b1}]$, $\mu_{a2} \in Prob[Sta_{a2}]$, and $\mu_{b2} \in Prob[Sta_{b2}]$. Consider weight functions $W_1$ and $W_2$ for $(\mu_{a1}, R_1, \mu_{b1})$ and $(\mu_{a2}, R_2, \mu_{b2})$, respectively. Then $W$ is a weight function for $(\mu_{a1} \otimes \mu_{a2}, R, \mu_{b1} \otimes \mu_{b2})$ where $W \in Prob[Sta_a \times Sta_b]$ is given as follows (see also Carathéodory extension theorem in Section 2.1): for every Borel set $B_{a1} \subseteq Sta_{a1}$, $B_{b1} \subseteq Sta_{b1}$, $B_{a2} \subseteq Sta_{a2}$, and $B_{b2} \subseteq Sta_{b2}$ let

$$W((B_{a1} \times B_{a2}) \times (B_{b1} \times B_{b2})) = W_1(B_{a1} \times B_{b1}) \otimes W_2(B_{a2} \times B_{b2}).$$

From this claim (1) immediately follows and claim (2) can be derived analogously. □

The importance of a congruence result for a compositional framework like the previous proposition is also accompanied by the question which properties are preserved by the simulation preorder and the bisimulation equivalence. The material below can be seen as folklore (see, e.g., Section 7.1.1 in [BK08] and Section 7.6 in [Pan09]). More precisely, to rely on the preservation results presented in Chapters 4 and 5, the following observation relating the notions in Definitions 47 and 140 is crucial.

**Remark 142.** Let $\mathcal{T}_a = (Sta_a, Act, \to_a)$ and $\mathcal{T}_b = (Sta_b, Act, \to_b)$ be simple STSs that have the same action space $Act$. For simplicity assume that the state spaces $Sta_a$ and $Sta_b$ are disjoint. The *union of* $(\mathcal{T}_a, \mathcal{T}_b)$ is defined by the simple STS

$$\mathcal{T}_a \cup \mathcal{T}_b = (Sta_a \cup Sta_b, Act, \to)$$

whose transition relation $\to$ is given by the following two inference rules:

(1) For every $s_a \in Sta_a$, $act \in Act$, and $\mu_a \in Prob[Sta_a \cup Sta_b]$:

$$\frac{\mu_a(Sta_a) = 1 \quad \text{and} \quad s_a \rightarrow_a \langle act_a, \mu'_a \rangle}{s_a \rightarrow \langle act_a, \mu_a \rangle}.$$

(2) For every $s_b \in Sta_b$, $act \in Act$, and $\mu_b \in Prob[Sta_a \cup Sta_b]$:

$$\frac{\mu_b(Sta_b) = 1 \quad \text{and} \quad s_b \rightarrow_b \langle act_b, \mu'_b \rangle}{s_b \rightarrow \langle act_b, \mu_b \rangle}.$$

Here, $\mu'_a \in Prob[Sta_a]$ is the probability measure induced by $\mu_a$, i.e., for every Borel set $S_a \subseteq Sta_a$ it holds $\mu'_a(S_a) = \mu_a(S_a)$, and analogously for $\mu'_b$. Intuitively, the STS $\mathcal{T}_a \cup \mathcal{T}_b$ simply combines the two STSs $\mathcal{T}_a$ and $\mathcal{T}_b$ into one.

For every states $s_a \in Sta_a$ and $s_b \in Sta_b$ the following two equivalences hold:

$$s_a \preceq_{\mathcal{T}_a \cup \mathcal{T}_b} s_b \quad \text{iff} \quad \mathcal{T}_a, s_a \preceq \mathcal{T}_b, s_b,$$
$$s_a \simeq_{\mathcal{T}_a \cup \mathcal{T}_b} s_b \quad \text{iff} \quad \mathcal{T}_a, s_a \simeq \mathcal{T}_b, s_b.$$

Here, the relations $\preceq_{\mathcal{T}_a \cup \mathcal{T}_b}$ and $\simeq_{\mathcal{T}_a \cup \mathcal{T}_b}$ denote the simulation preorder and the bisimulation equivalence for the STS $\mathcal{T}_a \cup \mathcal{T}_b$, respectively. A precise proof for these equivalence is technical but easy (see also Remark 49). We omit the details here and only remark that it suffices to show the following two observations. If $R \subseteq (Sta_a \cup Sta_b) \times (Sta_a \cup Sta_b)$ is a simulation for $\mathcal{T}_a \cup \mathcal{T}_b$, then the relation $R \cap (Sta_a \times Sta_b)$ is a simulation for $(\mathcal{T}_a, \mathcal{T}_b)$. Vice versa, if $R \subseteq Sta_a \times Sta_b$ is a simulation for $(\mathcal{T}_a, \mathcal{T}_b)$, then $R$ is also a relation in $(Sta_a \cup Sta_b) \times (Sta_a \cup Sta_b)$ that additionally forms a simulation for $\mathcal{T}_a \cup \mathcal{T}_b$. ⌟

Relying on the previous remark, the results of the previous chapters turn out to be applicable for the comparison of two STSs. We illustrate this by means of the following example. Let $\mathcal{T}_a = (Sta_a, Act, \rightarrow_a)$ and $\mathcal{T}_b = (Sta_b, Act, \rightarrow_b)$ be two Souslin STSs and $s_a \in Sta_a$ and $s_b \in Sta_b$ be respective states. For simplicity assume that $Sta_a \cap Sta_b = \varnothing$. Consider a simulation $R$ for $(\mathcal{T}_a, \mathcal{T}_b)$ such that $\langle s_a, s_b \rangle \in R$ and so that the set $R$ is Souslin in $Sta_a \times Sta_b$. Then for every $s_a$-scheduler for $\mathcal{T}_a$, say $\mathfrak{S}_a$, there exists an $s_b$-scheduler for $\mathcal{T}_b$, say $\mathfrak{S}_b$, such that

$$PrTrace[\mathfrak{S}_a] = PrTrace[\mathfrak{S}_b],$$

i.e., every trace distribution in the STS $\mathcal{T}_a$ originating in state $s_a$ can be mimicked by a trace distribution in the STS $\mathcal{T}_b$ originating in state $s_b$. This claim follows directly from Corollary 83 as well as Remark 142. Again, as in Remark 142, the precise argument is technical but easy. The basic observations are as follows: first of all, the STS $\mathcal{T}_a \cup \mathcal{T}_b$ is Souslin, and moreover, the set $R$ is Souslin in the Polish space $(Sta_a \cup Sta_b) \times (Sta_a \cup Sta_b)$. A similar discussion can be conducted for the results in Chapter 5 concerning the proven logical characterisation of the simulation preorder and the bisimulation equivalence.

## 6.2 Spans and span couplings

The following definition for spans is borrowed from category theory where the notion of a span yields a generalisation of the concept of a relation between objects. Throughout this thesis, we use spans as a modelling formalism to specify the global state space of two STSs to be composed.

$$\begin{tikzcd} & X \arrow[dl, dash, "\rho_1"'] \arrow[dr, dash, "\rho_2"] & \\ X_1 & & X_2 \end{tikzcd}$$

Figure 6.2: Illustration of a span.

**Definition 143.** A *span* is defined as a tuple

$$Sp = (X, X_1, X_2, \rho_1, \rho_2)$$

such that $X$, $X_1$, and $X_2$ are three Polish spaces as well as $\rho_1 \colon X \to X_1$ and $\rho_2 \colon X \to X_2$ are two Borel functions. $\lrcorner$

Figure 6.2 provides an illustration of a span. For every $x \in X$ we also write $x_{|1}$ and $x_{|2}$ instead of $\rho_1(x)$ and $\rho_2(x)$, respectively. Moreover, the *span relation* is given by the relation $Rel[Sp] \subseteq X_1 \times X_2$ defined as follows:

$$Rel[Sp] = \{\langle x_1, x_2 \rangle \in X_1 \times X_2 \text{ ; there is } x \in X \text{ with } x_1 = x_{|1} \text{ and } x_2 = x_{|2}\}.$$

Given two STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ with state spaces $Sta_1 = X_1$ and $Sta_2 = X_2$, respectively, the set $X$ intuitively stands for the global state space of $\mathcal{T}_1$ and $\mathcal{T}_2$ where $\rho_1$ and $\rho_2$ yield corresponding projection functions from $X$ to $Sta_1$ and $Sta_2$, respectively. Thus, considering two states $s_1 \in Sta_1$ and $s_2 \in Sta_2$, every $x \in X$ such that $\rho_1(x) = s_1$ and $\rho_2(x) = s_2$ represents a global state resulting from the local states $s_1$ and $s_2$ of the respective STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ to be composed.

**Example 144.** The following examples provide natural instances of spans. We remark that all the sets appearing in the following enumeration form Polish spaces by Example 2.

(1) Let $X_1$ and $X_2$ be Polish spaces. Denote the natural projections from the Cartesian product $X_1 \times X_2$ onto the components $X_1$ and $X_2$ by $\rho_1$ and $\rho_2$, respectively, i.e., $\rho_1(x_1, x_2) = x_1$ and $\rho_2(x_1, x_2) = x_2$ for every $x_1 \in X_1$ and $x_2 \in X_2$. Then the *Cartesian span for* $(X_1, X_2)$ is the span $(X_1 \times X_2, X_1, X_2, \rho_1, \rho_2)$.

(2) Let $X$ be a Polish space. Suppose that $\rho$ is the identity function on $X$, i.e., for every $x \in X$ it holds $\rho(x) = x$. The *identity span for* $X$ is the span $(X, X, X, \rho, \rho)$.

(3) Consider finite sets of variables $Var_1$ and $Var_2$. Let $\rho_1$ and $\rho_2$ be the functions projecting a variable evaluation over $Var_1 \cup Var_2$ onto $Var_1$ and $Var_2$, respectively, i.e., for every $e \in Eval[Var_1 \cup Var_2]$ it holds $\rho_1(e) = e_{|Var_1}$ and $\rho_2(e) = e_{|Var_2}$. The *variable span for* $(Var_1, Var_2)$ is given by $(Eval[Var_1 \cup Var_2], Eval[Var_1], Eval[Var_2], \rho_1, \rho_2)$.

(4) Pick two finite alphabets $\Sigma_1$ and $\Sigma_2$. Moreover, let $\rho_1$ and $\rho_2$ be the functions projecting a word over the alphabet $\Sigma_1 \cup \Sigma_2$ onto its consecutive sequences of letters in $\Sigma_1$ and $\Sigma_2$, respectively. Then the *finite-words span for* $(\Sigma_1, \Sigma_2)$ is defined by the tuple $((\Sigma_1 \cup \Sigma_2)^*, \Sigma_1^*, \Sigma_2^*, \rho_1, \rho_2)$.

(5) Consider a span $Sp = (X, X_1, X_2, \rho_1, \rho_2)$. Recall, according to Remark 8 (1), the two pushforward functions $(\rho_1)_\sharp$ and $(\rho_2)_\sharp$ are Borel. Hence, we can safely define the *probabilistic version of Sp* by the span $(Prob[X], Prob[X_1], Prob[X_2], (\rho_1)_\sharp, (\rho_2)_\sharp)$.

⌟

The variable span for $(Var_1, Var_2)$ in Example 144 (3) can be seen as a generalisation of the spans in (1), i.e., $Var_1 \cap Var_2 = \varnothing$, and in (2), i.e., $Var_1 = Var_2$. Intuitively, the presented definition for a variable span inherently declares the variables in $Var_1 \cap Var_2$ as shared and the remaining ones contained in $Var_1 \setminus Var_2$ or $Var_2 \setminus Var_1$ as local. Let us illustrate the versatility of spans by means of the following example of a span:

**Example 145.** Pick three different variables $v_1$, $v_2$, and $v$. Consider the variant of the variable span given by the span $(Eval[\{v_1, v_2\}], Eval[\{v\}], Eval[\{v\}], \rho_1, \rho_2)$ such that for every $e \in Eval[\{v_1, v_2\}]$ it holds $\rho_1(e)(v) = e(v_1)$ and $\rho_2(e)(v) = e(v_2)$. Intuitively, given two STSs that both control a variable with name $v$, the above span formalises the fact that $v$ is not shared between these two STSs, i.e., the global state space specified by the span $Sp$ incorporates the two copies $v_1$ and $v_2$ of $v$. ⌟

The definition of a span coupling is straightforward:

**Definition 146.** Let $Sp = (X, X_1, X_2, \rho_1, \rho_2)$ be a span as well as $\mu_1 \in Prob[X_1]$ and $\mu_2 \in Prob[X_2]$ be probability measures. A probability measure $\mu \in Prob[X]$ is called a *span coupling of* $(\mu_1, \mu_2)$ *(concerning* $Sp$*)* if the following two conditions hold:

(1) $(\rho_1)_\sharp(\mu) = \mu_1$, i.e., for every Borel set $B_1 \subseteq X_1$ one has $\mu(\rho_1^{-1}(B_1)) = \mu_1(B_1)$ .

(2) $(\rho_2)_\sharp(\mu) = \mu_2$, i.e., for every Borel set $B_2 \subseteq X_2$ one has $\mu(\rho_2^{-1}(B_2)) = \mu_2(B_2)$.

⌟

For every probability measure $\mu \in Prob[X]$ we use $\mu_{|1}$ and $\mu_{|2}$ as shorthand notations for $(\rho_1)_\sharp(\mu)$ and $(\rho_2)_\sharp(\mu)$, respectively. Obviously, Definitions 146 collapses with Definition 29 considering the special case where the underlying span is a Cartesian span. Inspecting for instance the identity span or the variable span, there may exists no span coupling of given probability measures. In the remainder of this section we establish a connection between the concepts of a span coupling and a weight function that finally provides an intuitive characterisation for the existence of span couplings.

**Lemma 147.** *Let* $Sp = (X, X_1, X_2, \rho_1, \rho_2)$ *be a span,* $W \in Prob[X_1 \times X_2]$ *be a probability measure, and* $B \subseteq X_1 \times X_2$ *be a Borel set such that* $B \subseteq Rel[Sp]$ *and* $W(B) = 1$*. Then there is a Borel set* $B' \subseteq X_1 \times X_2$ *and a Borel function* $f\colon X_1 \times X_2 \to X$ *with the following properties:*

(1) $B' \subseteq B$ *and* $\mu(B') = 1$*.*

(2) *For every* $\langle x_1, x_2 \rangle \in B'$ *it holds* $f(x_1, x_2)_{|1} = x_1$ *and* $f(x_1, x_2)_{|2} = x_2$*.*

*Proof.* A proof can be obtained by a standard application of a measurable-selection principle. More precisely, define the set-valued function $F\colon X_1 \times X_2 \rightsquigarrow X$,

$$F(x_1, x_2) = \{x \in X \,;\, x_{|1} = x_1 \text{ and } x_{|2} = x_2\}.$$

The function $\rho\colon X \to X_1 \times X_2, \rho(x) = \langle x_{|1}, x_{|2} \rangle$ is Borel, and therefore the set $Graph[\rho]$ is Borel in $X \times (X_1 \times X_2)$ (see Remark 10 (6)). From this we derive that the set $Rel[F]$ is Borel and also Souslin in $(X_1 \times X_2) \times X$ (see Remark 10 (2)). Moreover, for every $\langle x_1, x_2 \rangle \in B$ the set $F(x_1, x_2)$ is not empty. We can hence apply Theorem 21 that yields the existence of a Borel $W$-selection of $F$. Thus, there are a Borel set $B'' \subseteq X_1 \times X_2$ and a Borel function $f\colon X_1 \times X_2 \to X$ such that $W(B'') = 1$ and $f(x_1, x_2) \in F(x_1, x_2)$ for every $\langle x_1, x_2 \rangle \in B''$. Defining $B' = B \cap B''$, the set $B'$ and the function $f$ satisfy requirements (1) and (2). □

**Theorem 148.** *Let $Sp = (X, X_1, X_2, \rho_1, \rho_2)$ be a span and $\mu_1 \in Prob[X_1]$ and $\mu_2 \in Prob[X_2]$ be probability measures. Then we have the following equivalence:*

$$\textit{there is a span coupling of } (\mu_1, \mu_2) \quad \textit{iff} \quad \langle \mu_1, \mu_2 \rangle \in Rel[Sp]^{\mathrm{wgt}}.$$

*Proof.* Define the Borel function $\rho\colon X \to X_1 \times X_2, \rho(x) = \langle x_{|1}, x_{|2} \rangle$. By Remark 10 (1), the set $Rel[Sp]$ is Souslin in $X_1 \times X_2$. Thus, relying on Lemma 13, it is easy to see that for every span coupling $\mu$ of $(\mu_1, \mu_2)$ the probability measure $\rho_\sharp(\mu)$ forms a weight function for $(\mu_1, Rel[Sp], \mu_2)$. Hence, if there is span coupling of $(\mu_1, \mu_2)$, then $\langle \mu_1, \mu_2 \rangle \in Rel[Sp]^{\mathrm{wgt}}$. This shows the first implication of the claimed equivalence.

For the reverse implication consider a weight function $W$ for $(\mu_1, Rel[Sp], \mu_2)$. Thus, there exists a Borel set $B \subseteq X_1 \times X_2$ such that $B \subseteq Rel[Sp]$ and $W(B) = 1$. According to Lemma 147, there exist are a Borel set $B' \subseteq X_1 \times X_2$ and a Borel function $f\colon X_1 \times X_2 \to X$ such that $B' \subseteq Rel[Sp]$ and $W(B') = 1$ as well as $f(x_1, x_2)_{|1} = x_1$ and $f(x_1, x_2)_{|2} = x_2$ for every $\langle x_1, x_2 \rangle \in B'$. It is easy to see that $f_\sharp(W)$ is a span coupling of $(\mu_1, \mu_2)$. □

The previous theorem provides a tight connection between the existence of span couplings and weight functions concerning the span relation. This yields another nice application of weight functions. Moreover, our foundational results in Sections 2.6 and 2.7 concerning weight liftings of relations turn out to be applicable in the study of span couplings:

**Corollary 149.** *Let $Sp = (X, X_1, X_2, \rho_1, \rho_2)$ be a span, $\mu_1 \in Prob[X_1]$, and $\mu_2 \in Prob[X_2]$. Consider a Polish space $Y$ and two Borel functions $f_1\colon X_1 \to Y$ and $f_2\colon X_2 \to Y$ such that*

$$Rel[Sp] = \{\langle x_1, x_2 \rangle \in X_1 \times X_2 \; ; \; f_1(x_1) = f_2(x_2)\}.$$

*Then the following equivalence holds:*

$$\textit{there is a span coupling of } (\mu_1, \mu_2) \quad \textit{iff} \quad (f_1)_\sharp(\mu_1) = (f_2)_\sharp(\mu_2).$$

*Proof.* The claim is a consequence of Theorem 148 together with Corollary 41. □

In the previous corollary we assume that the span relation is smooth (see also Remark 37). It turns out that this requirement is harmless in the context of this chapter having for instance a variable span in mind:

**Example 150.** Let $Var_1$ and $Var_2$ be two finite sets of variables and consider the variable span $Sp$ for $(Var_1, Var_2)$ (see Example 144 (3)). Define $SVar = Var_1 \cap Var_2$. Intuitively, the variables in $SVar$ represent shared variables. It is easy to see that

$$Rel[Sp] = \{\langle e_1, e_2 \rangle \in Eval[Var_1] \times Eval[Var_2] \; ; \; e_{1|SVar} = e_{2|SVar}\},$$

i.e., the span relation is smooth by Remark 37. According to Corollary 149, for every $\eta_1 \in Prob[Eval[Var_1]]$ and $\eta_2 \in Prob[Eval[Var_2]]$ the following equivalence holds:

$$\text{there is a span coupling of } (\eta_1, \eta_2) \quad \text{iff} \quad \eta_{1|SVar} = \eta_{2|SVar}.$$

As a consequence, there exists a span coupling of $(\eta_1, \eta_2)$ precisely when the corresponding restrictions onto the shared variables are the same. This is a statement one naturally expects in the context of variables. ⌟

## 6.3 Span-coupling composition operator

This section is devoted to our new composition operator for simple STSs relying on spans and span couplings. Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be STSs and denote their corresponding state spaces by $Sta_1$ and $Sta_2$, respectively. To improve our notions, we drop the Polish spaces $Sta_1$ and $Sta_2$ from the tuple $(Sta, Sta_1, Sta_2, \rho_1, \rho_2)$ defining a span and simply refer to a triple $(Sta, \rho_1, \rho_2)$ as a span for $(\mathcal{T}_1, \mathcal{T}_2)$.

**Definition 151.** Let $\mathcal{T}_1 = (Sta_1, Act_1, \rightarrow_1)$ and $\mathcal{T}_2 = (Sta_2, Act_2, \rightarrow_2)$ be simple STSs, $Sync \subseteq Act_1 \cap Act_2$ be a set of actions, and $Sp = (Sta, \rho_1, \rho_2)$ be a span for $(\mathcal{T}_1, \mathcal{T}_2)$. Provided the STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ are composable, the *span-coupling composition of* $(\mathcal{T}_1, \mathcal{T}_2)$ *concerning* $(Sp, Sync)$ is defined by the STS

$$\mathcal{T}_1 \parallel_{Sp,Sync} \mathcal{T}_2 = (Sta, Act_1 \cup Act_2, \rightarrow),$$

whose transition relation $\rightarrow$ is given by the following three inference rules:

(1) For every $s \in Sta$, $act_1 \in Act_1 \setminus Sync$, and $\mu \in Prob[Sta]$:

$$\frac{s_{|1} \rightarrow_1 \langle act_1, \mu_{|1} \rangle \quad \text{and} \quad \mu_{|2} = Dirac[s_{|2}]}{s \rightarrow \langle act_1, \mu \rangle}.$$

(2) For every $s \in Sta$, $act_2 \in Act_2 \setminus Sync$, and $\mu \in Prob[Sta]$:

$$\frac{\mu_{|1} = Dirac[s_{|1}] \quad \text{and} \quad s_{|2} \rightarrow_2 \langle act_2, \mu_{|2} \rangle}{s \rightarrow \langle act_2, \mu \rangle}.$$

(3) For every $s \in Sta$, $act \in Act$, and $\mu \in Prob[Sta]$:

$$\frac{s_{|1} \rightarrow_1 \langle act, \mu_{|1} \rangle \quad \text{and} \quad s_{|2} \rightarrow_2 \langle act, \mu_{|2} \rangle}{s \rightarrow \langle act, \mu \rangle}.$$

┘

Intuitively, the set of actions *Sync* stands for the set of all synchronisation actions and the span *Sp* specifies the global state space for the STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ to be composed. As for the standard compositional framework, the new composition operator requires that the STSs under consideration are composable (see Definition 137). However, this assumption is harmless (see Example 138) and is assumed implicitly when composing STSs. To simplify notions, we write

$$\mathcal{T}_1 \parallel_{CartSp,Sync} \mathcal{T}_2, \quad \mathcal{T}_1 \parallel_{IdSp,Sync} \mathcal{T}_2, \quad \text{and} \quad \mathcal{T}_1 \parallel_{VarSp,Sync} \mathcal{T}_2$$

provided the considered span is given by the respective Cartesian span, identity span, and variable span, respectively (see Example 144). Clearly, if one focuses on the identity span for $(\mathcal{T}_1, \mathcal{T}_2)$, then the state spaces of $\mathcal{T}_1$ and $\mathcal{T}_2$ are necessarily the same.

Actions not contained in *Sync* are seen as local for the respective STS and can be executed autonomously. Inspecting the inference rule in (1), the condition $\mu_{|2} = Dirac[s_{|2}]$ ensures that the state space of the STS $\mathcal{T}_2$ is not affected when the STS $\mathcal{T}_1$ takes the local action $act_1$. The same discussion applies for the inference rule (2). As for the standard composition operator, according to the inference rule (3), a synchronisation action can be executed if both communication partners are ready to perform the action.

To refer to our initial motivation presented in Section 1, our composition operator introduced in Definition 151 does not involve any assumptions concerning stochastic dependencies between $\mathcal{T}_1$ and $\mathcal{T}_2$. Considering for instance the case where the span *Sp* is Cartesian, i.e., $Sta = Sta_1 \times Sta_2$, we obtain the following insight: if $s_1 \rightarrow_1 \langle act, \mu_1 \rangle$ and $s_2 \rightarrow_2 \langle act, \mu_2 \rangle$ with $act \in Sync$, then for every coupling $\mu$ of $(\mu_1, \mu_2)$ it holds $\langle s_1, s_2 \rangle \rightarrow \langle act, \mu \rangle$ . Hence, we take all possible couplings of $(\mu_1, \mu_2)$ into account. In comparison to that, the standard composition operator only includes the independent coupling $\mu_1 \otimes \mu_2$ (see Definition 139). However, the standard composition can be expressed by our compositional framework as illustrated below:

**Example 152.** Let $\mathcal{T}_1 = (Sta_1, Act_1, \rightarrow_1)$ and $\mathcal{T}_2 = (Sta_2, Act_2, \rightarrow_2)$ be simple STSs and $Sync \subseteq Act_1 \cap Act_2$ be a set of actions. Assume that $\mathcal{T}_1$ and $\mathcal{T}_2$ are composable. Define $Sta = Sta_1 \times Sta_2$ and $Act = Act_1 \cup Act_2$. Introduce the simple STS $\mathcal{T}_\otimes = (Sta, Act, \rightarrow)$ whose transition relation $\rightarrow$ is given by the following inference rule: for every $s \in Sta$, $act \in Act$, and $\mu \in Prob[Sta]$:

$$\frac{\mu = \mu_1 \otimes \mu_2 \text{ for some } \mu_1 \in Prob[Sta_1] \text{ and } \mu_2 \in Prob[Sta_2]}{s \rightarrow \langle act, \mu \rangle}.$$

Then we have the identity below (see also Remark 31):

$$\mathcal{T}_1 \parallel_{\otimes,Sync} \mathcal{T}_2 = (\mathcal{T}_1 \parallel_{CartSp,Sync} \mathcal{T}_2) \parallel_{IdSp,Act} \mathcal{T}_\otimes.$$

Intuitively, the STS $\mathcal{T}_\otimes$ encodes the fact that all the relevant couplings in the standard composition of $\mathcal{T}_1$ and $\mathcal{T}_2$ are given by the respective product measures. ⌟

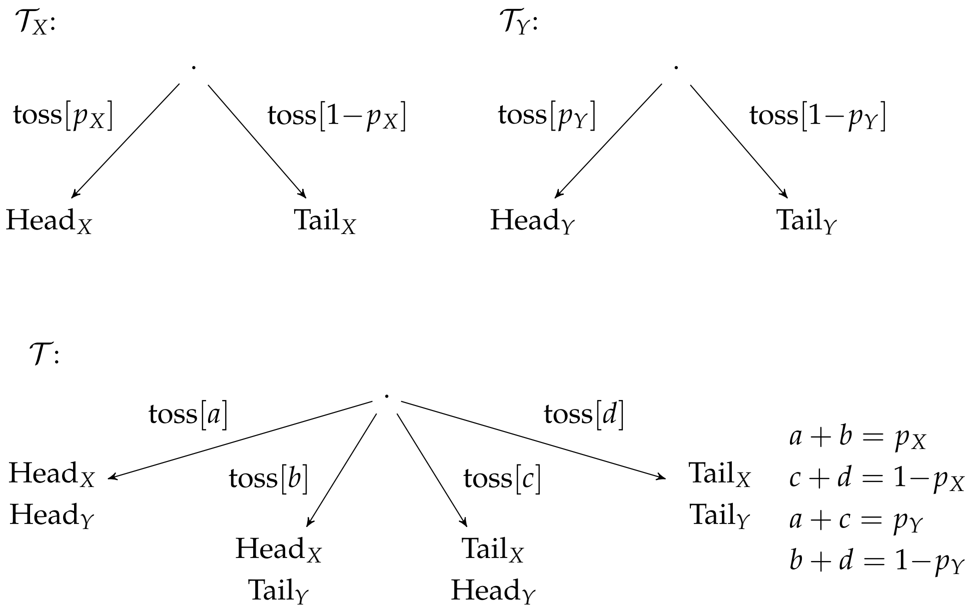


Figure 6.3: Simple STSs modelling coin tosses.

**Example 153.** We return to our coin-tossing example from Section 2.5 that discusses different couplings modelling a simultaneous toss of two coins. Take the notions as in Example 30 concerning a coin-toss experiment with the two coins $\mathrm{Coin}_X$ and $\mathrm{Coin}_Y$. The simple STSs $\mathcal{T}_X$ and $\mathcal{T}_Y$ depicted in Figure 6.3 model the toss of the individual coins $\mathrm{Coin}_X$ and $\mathrm{Coin}_Y$, respectively. The span-coupling composition of $(\mathcal{T}_X, \mathcal{T}_Y)$, i.e., the STS $\mathcal{T}$ illustrated in Figure 6.3, formally given by

$$\mathcal{T} = \mathcal{T}_X \parallel_{CartSp,\{\mathrm{toss}\}} \mathcal{T}_Y,$$

models the experiment of tossing the two coins $\mathrm{Coin}_X$ and $\mathrm{Coin}_Y$ simultaneously where we do not rely on any further stochastic information concerning their interplay. In particular, it is not assumed that the coins are tossed independently from each other. Note, as demonstrated by Example 152, an additional STS may encode available stochastic information to refine the model for the simultaneous coin toss. ⌟

**Example 154.** Let $Var_1 = \{\text{cool}\}$ and $Var_2 = \{\text{temp}, \text{cool}\}$. Recalling our running example concerning a cooling system of a server (see introduction of this chapter and also Section 3.1), the variable temp may refer to the internal temperature of the server and the variable cool may represent the heat dissipated by the cooling system. Consider the STSs $\mathcal{T}_1 = (Eval[Var_1], \mathbb{R}_{\geq 0}, \rightarrow_1)$ and $\mathcal{T}_2 = (Eval[Var_2], \mathbb{R}_{\geq 0}, \rightarrow_2)$ modelling the continuous evolution of the cooling system and the server, respectively.

Assume the STS $\mathcal{T}_1$ for the cooling system is determined by the the differential inclusion

$$dot[\text{cool}] \in [0, 5].$$

Intuitively, this means that the heat dissipated over time can be represented by every differentiable function whose slope is between zero and five at each point in time. Formally, for every $e \in Eval[Var]$ and $t \in \mathbb{R}_{\geq 0}$ it holds $e \rightarrow_1 \langle t, Dirac[e'] \rangle$ if there exists a differentiable function $f \colon \mathbb{R}_{\geq 0} \rightarrow Eval[Var_1]$ such that for every $t' \in [0, t]$ one has that $dot[f](t')(\text{cool}) \in [0, 5]$. Here, $dot[f]$ is the function with domain $\mathbb{R}_{\geq 0}$ and codomain $Eval[Var_1]$ given by the first derivative of $f$ concerning the time.

The STS $\mathcal{T}_2$ for the server is governed by the differential equation

$$dot[\text{temp}] = \text{temp} - \text{cool},$$

i.e., for every $e, e' \in Eval[Var_2]$ and $t \in \mathbb{R}_{\geq 0}$ we have $e \rightarrow_2 \langle t, Dirac[e'] \rangle$ if there exists a differentiable function $g \colon \mathbb{R}_{\geq 0} \rightarrow Eval[Var_2]$ such that for every $t' \in [0, t]$ it holds $dot[g](t')(\text{temp}) = g(t')(\text{temp}) - g(t')(\text{cool})$. Again, $dot[g]$ denotes the derivate of $g$ with respect to time. Note, the STS $\mathcal{T}_2$ includes no equation constraining the derivative of the variable cool. Indeed, this variable is controlled and determined by the cooling system and, roughly speaking, the server has no knowledge on how the flow of the actual cooling system looks like relying on the principle of separation the concerns.

The span-coupling composition for $(\mathcal{T}_1, \mathcal{T}_2)$ given by

$$\mathcal{T} = \mathcal{T}_1 \parallel_{VarSp, \mathbb{R}_{\geq 0}} \mathcal{T}_2$$

represents the differential inclusion system comprising of the two equations $dot[\text{cool}] \in [0, 5]$ and $dot[\text{temp}] = \text{temp} - \text{cool}$. The variable span under consideration declares the variable cool as shared. The components synchronise over the time. ⌟

Although the latter example involving a variable span involves no proper distributions, let us refer to Example 150. As a consequence of this example, a synchronisation action of two STSs with shared variables can only be executed provided the shared variables are affected equally.

## 6.4 Closure properties for Souslin and image-compact systems

We have investigated subclasses of STSs in Chapters 4 and 5. Concerning a proper theory on composition that fits well with our previous achievements, the question whether these subclasses are closed under composition is crucial. Besides others, we show that the composition of two reactive Souslin STSs also yields a Souslin STS. The latter is also of great practical importance: large and complex STSs satisfying certain measurability conditions can be obtained by the composition of small and manageable STSs with the specific property under consideration.

**Theorem 155.** *Let $\mathcal{T}_1 = (Sta_1, Act_1, \rightarrow_1)$ and $\mathcal{T}_2 = (Sta_2, Act_2, \rightarrow_2)$ be reactive STSs. Pick a span $Sp = (Sta, \rho_1, \rho_2)$ for $(\mathcal{T}_1, \mathcal{T}_2)$ as well as Borel set $Sync \subseteq Act_1 \cup Act_2$. Define*

$$\mathcal{T} = \mathcal{T}_1 \parallel_{Sp,Sync} \mathcal{T}_2.$$

*If the STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ are Souslin, then the STS $\mathcal{T}$ is also Souslin.*

*Proof.* The argument is straightforward and can be easily derived from Remark 68 □

The previous result is related to those subclass of STSs investigated in Chapter 4 concerning trace-distribution relations. In the remainder of this chapter, we provide a similar result referring to the STSs in Chapter 5 being Borel concerning the hit sigma algebra.

**Definition 156.** A span $(X, X_1, X_2, \rho_1, \rho_2)$ is called *proper* provided the functions $\rho_1$ and $\rho_2$ are continuous and moreover, for every compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$ the set $(\rho_1)^{-1}(K_1) \cap (\rho_2)^{-1}(K_2)$ is compact in $X$. ⌟

**Example 157.** The requirements on a proper span are harmless in the context of this thesis. Indeed, it is easy to see that every Cartesian span, every variable span, and every identity span is proper (see Example 144). Moreover, the variant of the variable span in Example 145 is also proper. Vice versa, it easy to give a span that is not proper. Consider for instance the span $Sp = (\mathbb{R} \times \mathbb{R} \times \mathbb{R}, \mathbb{R}_1, \mathbb{R}_2, \rho_1, \rho_2)$ such that for every $r_1, r_2, r_3 \in \mathbb{R}$ it holds $\rho_1(r_1, r_2, r_3) = r_1$ and $\rho_2(r_1, r_2, r_3) = r_2$. For every compact sets $K_1, K_2 \subseteq \mathbb{R}$ the set $(\rho_1)^{-1}(K_1) \cap (\rho_2)^{-1}(K_2) = K_1 \times K_2 \times \mathbb{R}$ is obviously not compact and hence, the span $Sp$ is not proper. ⌟

Continuing the previous example, we remark that every $Sp = (X, X_1, X_2, \rho_1, \rho_2)$ is proper provided the functions $\rho_1$ and $\rho_2$ are continuous and the the Polish space $X$ is compact. Indeed, for every compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$, using the continuity of $\rho_1$ and $\rho_2$, the set $(\rho_1)^{-1}(K_1) \cap (\rho_2)^{-1}(K_2)$ is closed in $X$ and therefore, by the compactness of $X$, also compact in $X$.

The following lemma shows that the probabilistic version of a proper span is proper:

**Lemma 158.** *Let $Sp = (X, X_1, X_2, \rho_1, \rho_2)$ be a proper span as well as $P_1 \subseteq Prob[X_1]$ and $P_2 \subseteq Prob[X_2]$ be compact sets. Define*

$$P = \{\mu \in Prob[X] \; ; \; \mu \text{ span coupling of } (\mu_1, \mu_2) \text{ for some } \mu_1 \in P_1 \text{ and } \mu_2 \in P_2\}.$$

*Then the set $P$ is compact in $Prob[X]$.*

*Proof.* The argument for the latter claim is the same as for the case where $P_1$ and $P_2$ are singleton sets (see also the proof of Lemma 3.5 in [Les10]). In fact, the claim basically follows from Prokhorov's theorem characterising the relatively compact subsets of probability measures (see Theorems 5.1 and 5.2 in [Bil99]). Recalling the theorem, given a Polish space $Y$ and a subset $P_Y \subseteq Prob[Y]$, Prokhorov's theorem states that the set $P_Y$ is compact in $Prob[Y]$ precisely when for every $\varepsilon \in \mathbb{R}_{>0}$ there is a compact set $K_Y \subseteq Y$ such that for every $\mu \in P_Y$ it holds $\mu(K_Y) \geq 1 - \varepsilon$.

Since the functions $\rho_1$ and $\rho_2$ are continuous, the pushforward functions $(\rho_1)_\sharp$ and $(\rho_2)_\sharp$ are continuous as well (see introduction of the proof of Theorem 129). It holds

$$P = (\rho_1)_\sharp^{-1}(P_1) \cap (\rho_2)_\sharp^{-1}(P_2).$$

Consequently, the set $P$ is closed in $Prob[X]$. To show that $P$ is compact in $Prob[X]$, it hence remains to show that $P$ is relatively compact in $Prob[X]$. Let $\varepsilon \in \mathbb{R}_{>0}$. According to Prokhorov's theorem theorem, there are compact sets $K_1 \subseteq X_1$ and $K_2 \subseteq X_2$ such that for every $\mu_1 \in P_2$ and $\mu_2 \in P_2$,

$$\mu_1(K_1) \geq 1 - \frac{\varepsilon}{2} \quad \text{and} \quad \mu_2(K_2) \geq 1 - \frac{\varepsilon}{2}.$$

Define the set $K \subseteq X$ by

$$K = (\rho_1)^{-1}(K_1) \cap (\rho_2)^{-1}(K_2).$$

As the span $Sp$ is proper, the set $K$ is compact in $X$. Moreover, for every $\mu \in P$ it holds

$$\mu(X \setminus K)$$

$$
\begin{aligned}
&= \mu((\rho_1)^{-1}(X_1 \setminus K_1) \cup (\rho_2)^{-1}(X_2 \setminus K_2)) \\
&\leq \mu((\rho_1)^{-1}(X_1 \setminus K_1)) + \mu((\rho_2)^{-1}(X_2 \setminus K_2)) \\
&= \mu_{|1}(X_1 \setminus K_1) + \mu_{|2}(X_2 \setminus K_2) \\
&\leq \varepsilon
\end{aligned}
$$

Consequently, applying Prokhorov's theorem, the set $P$ is relatively compact in $Prob[X]$ and therefore, compact in $Prob[X]$. □

The span-coupling composition operator takes all the possible couplings of components into account and hence, the composition of two reactive point-wise image-finite STSs is not image-finite in general. The following definition weakens the condition on the image-finiteness that yields a class of STSs related to Theorem 128 of Chapter 5 that is closed under composition

**Definition 159.** An STS $\mathcal{T} = (Sta, Act, \rightarrow)$ is called *point-wise image-compact* if for every state $s \in Sta$ and action $act \in Act$ the set $Enabled[s, act]$ is compact in $Prob[Sta]$. In particular, every point-wise image-finite STS is point-wise image-compact. ⌟

**Theorem 160.** *Let $\mathcal{T}_1 = (Sta_1, Act_1, \rightarrow_1)$ and $\mathcal{T}_2 = (Sta_2, Act_2, \rightarrow_2)$ be reactive STSs such that the sets $Act_1$ and $Act_2$ are countable. Pick a proper span $Sp = (Sta, \rho_1, \rho_2)$ for $(\mathcal{T}_1, \mathcal{T}_2)$ as well as set $Sync \subseteq Act_1 \cup Act_2$. Define*

$$\mathcal{T} = \mathcal{T}_1 \parallel_{Sp,Sync} \mathcal{T}_2.$$

*Provided the STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ are point-wise image-compact and Borel concerning the hit sigma algebra, the STS $\mathcal{T}$ is also point-wise image-compact and Borel concerning the hit sigma algebra.*

*Proof.* Let $act \in Act$. As the set $Act$ is countable, we can rely on Theorem 129. It hence suffices to show that for the set valued function $F_{act} \colon Sta \rightarrow Prob[Sta]$,

$$F_{act}(s) = Enabled[\mathcal{T}, s, act]$$

the following two properties hold: First, the set-valued function $F_{act}$ is Borel concerning the hit sigma algebra, i.e., the set $\{s \in Sta \; ; \; F_{act}(s) \cap O \neq \varnothing\}$ is Borel in $Sta$ for every open set $O \subseteq Prob[Sta]$ and second, the set $F_{act}(s)$ is compact in $Prob[Sta]$ for every $s \in Sta$. Here, Lemma 18.4 in [AB06] shows that $F_{act}$ is Borel concerning the hit sigma algebra and moreover, Lemma 158 provides a proof for the fact that $F_{act}(s)$ is compact in $Prob[Sta]$ for every $s \in Sta$. □

**Remark 161.** Inspecting Lemma 18.4 in [AB06] carefully, we even obtain the following result where the notions are as in Theorem 160. The STS $\mathcal{T}$ is Borel concerning the hit sigma algebra provided the following two statement holds:

(1) For every state $s_1 \in Sta_1$ and action $act_1 \in Act_1$ the set $Enabled[\mathcal{T}_1, s_1, act_1]$ is closed in $Prob[Sta_1]$ and accordingly, for every state $s_2 \in Sta_2$ and action $act_2 \in Act_2$ the set $Enabled[\mathcal{T}_2, s_2, act_2]$ is closed in $Prob[Sta_2]$.

(2) For every $s_1 \in Sta_1$, $s_2 \in Sta_2$, and $act \in Sync$ the set $Enabled[\mathcal{T}_1, s_1, act]$ is compact in $Prob[Sta_1]$ or the set $Enabled[\mathcal{T}_2, s_2, act]$ is compact in $Prob[Sta_2]$.

Thus, the compactness requirement only appears for synchronisation actions and even here, only one of the corresponding sets is assumed to be compact. However, assuming just the previous properties (1) and (2), we cannot conclude that the STS $\mathcal{T}$ is point-wise image-compact in general. ⌟

## 6.5 Construction of span couplings

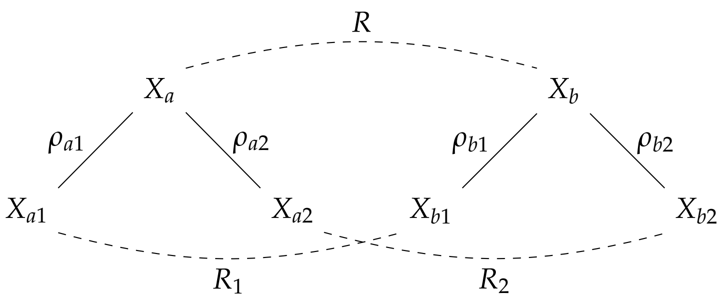


Figure 6.4: Schematic overview of the section's setting.

The following material provides a sufficient and necessary condition for the existence of span couplings with specific properties. To be more precise and to illustrate the motivation of this section, let $Sp_a = (X_a, X_{a1}, X_{a2}, \rho_{a1}, \rho_{a2})$ and $Sp_b = (X_b, X_{b1}, X_{b2}, \rho_{b1}, \rho_{b2})$ be spans as well as $R_1 \subseteq X_{a1} \times X_{b1}$ and $R_2 \subseteq X_{a2} \times X_{b2}$ be relations. Define the relation

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

Moreover, pick probability measures $\mu_a \in Prob[X_a]$, $\mu_{b1} \in Prob[X_{b1}]$, and $\mu_{b2} \in Prob[X_{b2}]$ with the following two properties:

$$\langle \mu_{a|1}, \mu_{b1} \rangle \in (R_1)^{\mathrm{wgt}} \quad \text{and} \quad \langle \mu_{a|2}, \mu_{b2} \rangle \in (R_2)^{\mathrm{wgt}}.$$

We refer to Figure 6.4 for an overview on the introduced notions. Intuitively, the relations $R_1$, $R_2$, and $R$ connect the components of the two spans under consideration where the relation $R$ yields a lifting of two relations $R_1$ and $R_2$. In this section we focus on a condition that ensures the existence of a probability measure $\mu_b \in Prob[X_b]$ that satisfies the following two conditions:

$$\mu_b \text{ is a span coupling of } (\mu_{b1}, \mu_{b2}) \quad \text{and} \quad \langle \mu_a, \mu_b \rangle \in R^{\text{wgt}}.$$

Roughly speaking, we are hence interested in a condition such that the span coupling $\mu_a$ with respect to the span $Sp_a$ can be transferred to a span coupling $\mu_b$ with respect to the span $Sp_b$ while respecting the given relations. The special case where $\mu_a = Dirac[x_a]$, $\mu_{b1} = Dirac[x_{b1}]$, and $\mu_{b2} = Dirac[x_{b2}]$ for some $x_a \in X_a$, $x_{b1} \in X_{b1}$, and $x_{b2} \in X_{b2}$ motivates the following definition of a compatibility criterion:

**Definition 162.** Let $Sp_a = (X_a, X_{a1}, X_{a2}, \rho_{a1}, \rho_{a2})$ and $Sp_b = (X_b, X_{b1}, X_{b2}, \rho_{b1}, \rho_{b2})$ be spans and $R_1 \subseteq X_{a1} \times X_{b1}$ and $R_2 \subseteq X_{a2} \times X_{b2}$ be relations. We say that the relations $R_1$ and $R_2$ are *compatible with $Sp_a$ and $Sp_b$* provided the following property holds: for every $x_a \in X_a$, $x_{b1} \in X_{b1}$, and $x_{b2} \in X_{b2}$ such that

$$\langle x_{a|1}, x_{b1} \rangle \in R_1 \quad \text{and} \quad \langle x_{a|2}, x_{b2} \rangle \in R_2$$

there exists $x_b \in X_b$ with

$$x_{b|1} = x_{b1} \quad \text{and} \quad x_{b|2} = x_{b2}.$$

⌟

**Theorem 163.** *Let $Sp_a = (X_a, X_{a1}, X_{a2}, \rho_{a1}, \rho_{a2})$ and $Sp_b = (X_b, X_{b1}, X_{b2}, \rho_{b1}, \rho_{b2})$ be spans as well as $R_1 \subseteq X_{a1} \times X_{b1}$ and $R_2 \subseteq X_{a2} \times X_{b2}$ be relations. Define the relation*

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

*If the relations $R_1$ and $R_2$ are compatible with $Sp_a$ and $Sp_b$, then for every probability measures $\mu_a \in Prob[X_a]$, $\mu_{b1} \in Prob[X_{b1}]$, and $\mu_{b2} \in Prob[X_{b2}]$ the two statements below are equivalent:*

(1) $\langle \mu_{a|1}, \mu_{b1} \rangle \in (R_1)^{\text{wgt}}$ *and* $\langle \mu_{a|2}, \mu_{b2} \rangle \in (R_2)^{\text{wgt}}$.

(2) *There exists a span coupling $\mu_b$ of $(\mu_{b1}, \mu_{b2})$ such that $\langle \mu_a, \mu_b \rangle \in R^{\text{wgt}}$.*

⌟

A precise proof of the theorem can be found below. The result provides a characterisation of the existence of a specific span coupling under a compatibility assumption. It even turns out that the involved compatibility assumption given by Definition 162 is best possible by means of the following observation: assuming that for every probability measures $\mu_a \in Prob[X_a]$, $\mu_{b1} \in Prob[X_{b1}]$, and $\mu_{b2} \in Prob[X_{b2}]$ the statements (1) and (2) of Theorem 163 are equivalent, then it easily follows that the relations $R_1$ and $R_2$ are compatible with $Sp_a$ and $Sp_b$. Indeed, given $x_a \in X_a$, $x_{b1} \in X_{b1}$, and $x_{b2} \in X_{b2}$ as in Definition 162, then one can simply consider the assumed equivalence of (1) and (2) for the corresponding Dirac measures.

**Corollary 164.** *Let $Sp_a = (X_a, X_{a1}, X_{a2}, \rho_{a1}, \rho_{a2})$ and $Sp_b = (X_b, X_{b1}, X_{b2}, \rho_{b1}, \rho_{b2})$ be spans as well as $R_1 \subseteq X_{a1} \times X_{b1}$ and $R_2 \subseteq X_{a2} \times X_{b2}$ be two relations. Define*

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

*If the span $Sp_b$ is Cartesian, then for every probability measures $\mu_a \in Prob[X_a]$, $\mu_{b1} \in Prob[X_{b1}]$, and $\mu_{b2} \in Prob[X_{b2}]$ the following two statements are equivalent:*

(1) $\langle \mu_{a|1}, \mu_{b1}\rangle \in (R_1)^{\mathrm{wgt}}$ *and* $\langle \mu_{a|2}, \mu_{b2}\rangle \in (R_2)^{\mathrm{wgt}}$.

(2) *There exists a span coupling $\mu_b$ of $(\mu_{b1}, \mu_{b2})$ such that $\langle \mu_a, \mu_b\rangle \in R^{\mathrm{wgt}}$.*

*Proof.* Assuming the span $Sp_b$ is Cartesian, it directly follows that the relations $R_1$ and $R_2$ are compatible with $Sp_a$ and $Sp_b$ and thus, Theorem 163 directly yields the claim. $\square$

In the previous corollary we restrict ourselves to the important case where the span $Sp_b$ is Cartesian. We emphasise that even a direct proof of the corollary without using Theorem 163 is not immediate. Indeed, the main challenge of constructing a specific coupling as sketched in the introduction of this chapter remains and hence, a direct argument basically avoids technicalities we deal with in our proof of the general theorem below.

Let us present a sketch how to obtain a convenient span coupling for (2) assuming statement (1) holds. This also provides a feeling for the proof of the general result given by Theorem 163. Consider the notions as in Corollary 164. Assume statement (1) and let $W_1$ and $W_2$ be weight functions for $\langle \mu_{a|1}, \mu_{b1}\rangle \in (R_1)^{\mathrm{wgt}}$ and $\langle \mu_{a|2}, \mu_{b2}\rangle \in (R_2)^{\mathrm{wgt}}$, respectively. According to the disintegration theorem (see Section 2.1), there are two Borel functions $f_1 \colon X_{a1} \to Prob[X_{b1}]$ and $f_2 \colon X_{a2} \to Prob[X_{b2}]$ such that

$$W_1 = \mu_{a|1} \rtimes f_1 \quad \text{and} \quad W_2 = \mu_{a|2} \rtimes f_2.$$

Based on the functions $f_1$ and $f_2$, define the function $f\colon X_a \to Prob[X_b]$,

$$f(x_a) = f_1(x_{a|1}) \otimes f_2(x_{a|2}).$$

By Example 6, the function $f$ is Borel and hence, we can safely define the probability measures $W = \mu_a \rtimes f$, i.e., for every Borel sets $B_a \subseteq X_a$ and $B_b \subseteq X_b$,

$$W(B_a \times B_b) = \int_{B_a} f(x_a)(B_b)\, d\mu_a(x_a),$$

and $\mu_b = Post[\mu_a, f]$, i.e., for every Borel sets $B_a \subseteq X_a$,

$$\mu_b(B_b) = \int f(x_a)(B_b)\, d\mu_a(x_a).$$

Thus, for every Borel set $B_b \subseteq X_b$ it holds $\mu_b(B_b) = W(X_a \times B_b)$. It turns out that $\mu_b$ is a span coupling of $(\mu_{b1}, \mu_{b2})$ that additionally satisfies $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. The latter is precisely proven below. From this statement (2) of Corollary 164 follows.

Roughly speaking, to recapitulating the previous argument, the disintegration theorem makes the coupling structure of the weight functions $W_1$ and $W_2$ accessible from a mathematical point of view. Note, a similar construction is applied in the illustrated proof of the gluing lemma in Section 3.4.

**Proof of section's main result.** The following lemma summarises the presented key observation for Corollary 164 presented above:

**Lemma 165.** *Let $Sp_a = (X_a, X_{a1}, X_{a2}, \rho_{a1}, \rho_{a2})$ and $Sp_b = (X_b, X_{b1}, X_{b2}, \rho_{b1}, \rho_{b2})$ be spans and $R_1 \subseteq X_{a1} \times X_{b1}$ and $R_2 \subseteq X_{a2} \times X_{b2}$ be relations. Define*

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

*For every probability measures $\mu_a \in Prob[X_a]$, $\mu_{b1} \in Prob[X_{b1}]$, and $\mu_{b2} \in Prob[X_{b2}]$ the following two statements are equivalent:*

(1) *There are Borel functions $f_1\colon X_{a1} \to Prob[X_{b1}]$ and $f_2\colon X_{a2} \to Prob[X_{b2}]$ satisfying the two conditions below:*

   a) *The probability measure $\mu_{a|1} \rtimes f_1$ is a weight function for $(\mu_{a|1}, R_1, \mu_{b1})$ and accordingly, $\mu_{a|2} \rtimes f_2$ is a weight function for $(\mu_{a|2}, R_2, \mu_{b2})$.*

   b) *There exists a Borel set $B_a \subseteq X_a$ such that $\mu_a(B_a) = 1$ and so that for every $x_a \in B_a$ there is a span coupling of $(f_1(x_{a|1}), f_2(x_{a|2}))$.*

(2) *There exists a span coupling* $\mu_b$ *of* $(\mu_{b1}, \mu_{b2})$ *such that* $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$.

*Proof.* (1) *implies* (2). Let $f_1 \colon X_{a1} \to \mathit{Prob}[X_{b1}]$ and $f_2 \colon X_{a2} \to \mathit{Prob}[X_{b2}]$ be Borel functions as well as $B_a \subseteq X_a$ be a Borel set as in (1). Defining

$$W_1 = \mu_{a|1} \rtimes f_1 \quad \text{and} \quad W_2 = \mu_{a|2} \rtimes f_2,$$

the following properties hence hold: $W_1$ and $W_2$ are weight functions for $(\mu_{a|1}, R_1, \mu_{b1})$ and $(\mu_{a|2}, R_2, \mu_{b2})$, respectively, as well as one has $\mu_a(B_a) = 1$ and for every $x_a \in B_a$ there is a span coupling of $(f_1(x_{a|1}), f_2(x_{a|2}))$.

Define the function $f' \colon X_a \to \mathit{Prob}[X_{b1}] \times \mathit{Prob}[X_{b2}]$,

$$f'(x_a) = \langle f_1(x_{a|1}), f_2(x_{a|2}) \rangle.$$

According to Lemma 13, there exists a Borel set $P' \subseteq \mathit{Prob}[X_{b1}] \times \mathit{Prob}[X_{b2}]$ such that

$$P \subseteq f'(B_a) \quad \text{and} \quad (f')_\sharp(\mu_a)(P) = 1.$$

Here, as $P \subseteq f'(B_a)$, for every $\langle \mu'_{b1}, \mu'_{b2} \rangle \in P$ there exists a span coupling of $(\mu'_{b1}, \mu'_{b2})$. Applying Lemma 147, there are a Borel function $g \colon \mathit{Prob}[X_{b1}] \times \mathit{Prob}[X_{b2}] \to \mathit{Prob}[X_b]$ and a Borel set $P' \subseteq \mathit{Prob}[X_{b1}] \times \mathit{Prob}[X_{b2}]$ such that the following properties hold:

$$P' \subseteq P \quad \text{and} \quad (f')_\sharp(\mu_a)(P') = 1$$

as well as for every $\langle \mu'_{b1}, \mu'_{b2} \rangle \in P'$,

$$g(\mu'_{b1}, \mu'_{b2}) \text{ is a span coupling of } (\mu'_{b1}, \mu'_{b2}).$$

Based on the Borel functions $f'$ and $g$, define the Borel function $f \colon X_a \to \mathit{Prob}[X_b]$,

$$f(x_a) = g(f'(x_a)).$$

Introduce the probability measure $\mu_b \in \mathit{Prob}[X_b]$ by $\mu_b = \mathit{Post}[\mu_a, f]$, i.e., for every Borel set $B_b \subseteq X_b$,

$$\mu_b(B_b) = \int f(x_a)(B_b)\, d\mu_a(x_a).$$

We show that the probability measure $\mu_b$ is a span coupling of $(\mu_{b1}, \mu_{b2})$ that additionally satisfies the condition $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$ that finally shows (2).

Let $\xi_1 \colon X_a \times X_b \to X_{a1} \times X_{b1}$ and $\xi_2 \colon X_a \times X_b \to X_{a2} \times X_{b2}$ be given as follows: for every $x_a \in X_a$ and $x_b \in X_b$ let

$$\xi_1(x_a, x_b) = \langle x_{a|1}, x_{b|1} \rangle \quad \text{and} \quad \xi_2(x_a, x_b) = \langle x_{a|2}, x_{b|2} \rangle.$$

Define the probability measure $W \in \mathit{Prob}[X_a \times X_b]$ by

$$W = \mu_a \rtimes f.$$

Then, the following two identities hold:

$$W_1 = (\xi_1)_\sharp(W) \quad \text{and} \quad W_2 = (\xi_2)_\sharp(W).$$

Let us see why. Suppose Borel sets $B_{a1} \subseteq X_{a1}$ and $B_{b1} \subseteq X_{b1}$. Relying on the identity $\xi_1^{-1}(B_{a1} \times B_{b1}) = \rho_{a1}^{-1}(B_{a1}) \times \rho_{b1}^{-1}(B_{b1})$, we obtain

$$(\xi_1)_\sharp(W)(B_{a1} \times B_{b1}) = \int_{\rho_{a1}^{-1}(B_{a1})} f(x_a)(\rho_{b1}^{-1}(B_{b1}))\, d\mu_a(x_a).$$

Recalling the properties of the set $P'$, for every pair $\langle \mu'_{b1}, \mu'_{b2} \rangle \in P'$ we have the identity $g(\mu'_{b1}, \mu'_{b2})_{|1} = \mu'_{b1}$. Thus, for every $x_a \in (f')^{-1}(P')$ it holds $g(f'(x_a))_{|1} = f_1(x_{a|1})$. Since $\mu_a((f')^{-1}(P')) = (f')_\sharp(\mu_a)(P') = 1$, we therefore obtain

$$(\xi_1)_\sharp(W)(B_{a1} \times B_{b1}) = \int_{\rho_{a1}^{-1}(B_{a1})} f_1(x_{a|1})(B_{b1})\, d\mu_a(x_a).$$

and consequently,

$$(\xi_1)_\sharp(W)(B_{a1} \times B_{b1}) = \int_{B_{a1}} f_1(x_{a1})(B_{b1})\, d\mu_{a|1}(x_{a1}) = W_1(B_{a1} \times B_{b1}).$$

Carathéodory uniqueness theorem (see Section 2.1) therefore yields $(\xi_1)_\sharp(W) = W_1$. The second statement $(\xi_2)_\sharp(W) = W_2$ can be proven in the same way.

Relying on the obtained identities, we show that $\mu_b$ is a span coupling of $(\mu_{b1}, \mu_{b2})$. For every Borel set $B_{b1} \subseteq X_{b1}$ it holds

$$W(X_a \times \rho_{b1}^{-1}(B_{b1})) = W(\xi_1^{-1}(X_{a1} \times B_{b1})) = W_1(X_{a1} \times B_{b1}) = \mu_{b1}(B_{b1})$$

and therefore, we obtain

$$\mu_{b|1}(B_{b1}) = \mu_b(\rho_{b1}^{-1}(B_{b1})) = W(X_a \times \rho_{b1}^{-1}(B_{b1})) = \mu_{b1}(B_{b1}).$$

One analogously justifies $\mu_{b|2} = \mu_{b2}$ and hence, $\mu_b$ is a span coupling of $(\mu_{b1}, \mu_{b2})$.

It remains to show that $W$ is a weight function for $(\mu_a, R, \mu_b)$. Clearly, $W$ is a coupling of $(\mu_a, \mu_b)$. Let $R'_1 \subseteq X_{a1} \times X_{b1}$ and $R'_2 \subseteq X_{a2} \times X_{b2}$ be Borel sets so that

$$W_1(R'_1) = 1, \quad R'_1 \subseteq R_1, \quad W_2(R'_2) = 1, \quad \text{and} \quad R'_2 \subseteq R_2.$$

Define

$$R' = \xi_1^{-1}(R'_1) \cap \xi_2^{-1}(R'_2).$$

It holds $W(\xi_1^{-1}(R'_1)) = W_1(R'_1) = 1$ and $W(\xi_2^{-1}(R'_2)) = W_1(R'_2) = 1$. From this we conclude $W(R') = 1$. It is easy to see that $R' \subseteq R$. Consequently, $W$ is a weight function for $(\mu_a, R, \mu_b)$. This finishes the proof of the implication from (1) to (2).

(2) *implies* (1). Let $\mu_b$ be a span coupling of $(\mu_{b1}, \mu_{b2})$ such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. Suppose a weight function $W$ for $(\mu_a, R, \mu_b)$. As in the first part of the proof introduce the Borel functions $\xi_1 \colon X_a \times X_b \to X_{a1} \times X_{b1}$ and $\xi_2 \colon X_a \times X_b \to X_{a2} \times X_{b2}$ as follows: for every $x_a \in X_a$ and $x_b \in X_b$ let

$$\xi_1(x_a, x_b) = \langle x_{a|1}, x_{b|1} \rangle \quad \text{and} \quad \xi_2(x_a, x_b) = \langle x_{a|2}, x_{b|2} \rangle.$$

Define the probability measures $W_1 \in \mathit{Prob}[X_{a1} \times X_{b1}]$ and $W_2 \in \mathit{Prob}[X_{a2} \times X_{b2}]$ by

$$W_1 = (\xi_1)_\sharp(W) \quad \text{and} \quad W_2 = (\xi_2)_\sharp(W).$$

Using Lemma 13, it is easy to see that $W_1$ and $W_2$ are a weight functions for $(\mu_{a|1}, R_1, \mu_{b1})$ and $(\mu_{a|2}, R_2, \mu_{b2})$, respectively. Relying on the disintegration theorem (see Section 2.1), there are three Borel functions $f_1 \colon X_{a1} \to \mathit{Prob}[X_{b1}]$, $f_2 \colon X_{a2} \to \mathit{Prob}[X_{b2}]$, and $f \colon X_a \to \mathit{Prob}[X_b]$ so that

$$W_1 = \mu_{a|1} \rtimes f_1, \quad W_2 = \mu_{a|2} \rtimes f_2, \quad \text{and} \quad W = \mu_a \rtimes f.$$

To conclude statement (1), it suffices to show that there exists a Borel set $B_a \subseteq X_a$ with $\mu_a(B_a) = 1$ and such that for every $x_a \in B_a$ the probability measure $f(x_a)$ is a span coupling of $(f_1(x_{a|1}), f_2(x_{a|2}))$ .

Let $\mathcal{G}_{b1}$ be a countable generator of the Borel sigma algebra on $X_{b1}$ such that $\mathcal{G}_{b1}$ is closed under finite intersections (see Remark 4). Pick Borel sets $B_{b1,0}, B_{b1,1}, B_{b1,2}, \ldots$ of $X_{b1}$ such that $\mathcal{G}_{b1} = \{B_{b1,0}, B_{b1,1}, B_{b1,2}, \ldots\}$. Moreover, define the family of sets

$$\mathcal{B}_a = \{\rho_{a1}^{-1}(B_{a1}) \; ; \; B_{a1} \subseteq X_{a1} \text{ is a Borel set}\}.$$

It is easy to see that $\mathcal{B}_a$ is a sigma algebra on $X_a$ that is contained in the Borel sigma algebra on $X_a$. Hence, $\mu_a$ can be also viewed as probability measure concerning the newly introduced sigma algebra $\mathcal{B}_a$ on $X_a$.

Since $W_1 = (\xi_1)_\sharp(W)$, $W_1 = \mu_{a|1} \rtimes f_1$, and $W = \mu_a \rtimes f$, for every $n \in \mathbb{N}$ and every Borel sets $B_{a1} \subseteq X_{a1}$ it holds

$$\int_{\rho_{a1}^{-1}(B_{a1})} f(x_a)_{|1}(B_{b1,n})\, d\mu_a(x_a) = \int_{\rho_{a1}^{-1}(B_{a1})} f_1(x_{a|1})(B_{b1,n})\, d\mu_a(x_a).$$

Thus, by a standard result from measure theory (see Folgerung 9.2.5 in [Sch08]), for every $n \in \mathbb{N}$ there exists a set $B_{a,n} \in \mathcal{B}_a$ with $\mu_a(B_{a,n}) = 1$ and such that for every $x_a \in B_{a,n}$ the following identity holds:

$$f(x_a)_{|1}(B_{b1,n}) = f_1(x_{a|1})(B_{b1,n}).$$

Define

$$B'_a = \bigcap_{n\in\mathbb{N}} B_{a,n}.$$

As for every $n \in \mathbb{N}$ the set $B_{a,n}$ is Borel in $X_a$, we have that the set $B'_a$ is Borel in $X_a$. It holds $\mu_a(B'_a) = 1$, Moreover, Carathéodory uniqueness theorem (see Section 2.1) yields the identity $f(x_a)_{|1} = f_1(x_{a|1})$ for every $x_a \in B'_a$.

One analogously proves the existence of a Borel set $B''_a \subseteq X_a$ such that $\mu_a(B''_a) = 1$ and $f(x_a)_{|2} = f_2(x_{a|2})$ for every $x_a \in B''_a$. Defining $B_a = B'_a \cap B''_a$, we obtain $\mu_a(B_a) = 1$ and moreover, $f(x_a)$ is a span coupling of $(f_1(x_{a|1}), f_2(x_{a|2}))$ for every $x_a \in B_a$. □

*Proof of Theorem 163.* The implication from (2) to (1) follows immediately from Lemma 165. The reminder of this proof is devoted to the reverse implication. Assume statement (1) holds. Let $W_1$ and $W_2$ be weight functions for $(\mu_{a|1}, R_1, \mu_{b1})$ and $(\mu_{a|2}, R_2, \mu_{b2})$, respectively. By the disintegration theorem (see Section 2.1), there are Borel functions $f_1\colon X_{a1} \to Prob[X_{b1}]$ and $f_2\colon X_{a2} \to Prob[X_{b2}]$ such that

$$W_1 = \mu_{a|1} \rtimes f_1 \quad \text{and} \quad W_2 = \mu_{a|2} \rtimes f_2.$$

Consider Borel sets $R'_1 \subseteq X_{a1} \times X_{b1}$ and $R'_2 \subseteq X_{a2} \times X_{b2}$ so that

$$W_1(R'_1) = 1, \quad R'_1 \subseteq R_1, \quad W_2(R'_2) = 1, \quad \text{and} \quad R'_2 \subseteq R_2.$$

Sections of Borel sets in product spaces are Borel (see Section 2.1). According to Lemma 15, there hence exists a Borel set $B_{a1} \subseteq X_{a1}$ so that

$$\mu_{a|1}(B_{a1}) = 1 \quad \text{and} \quad f_1(x_{a1})(Section[R', x_{a1}, \cdot]) = 1 \text{ for every } x_{a1} \in B_{a1}.$$

Analogously, there exists a Borel set $B_{a2} \subseteq X_{a2}$ with

$$\mu_{a|2}(B_{a2}) = 1 \quad \text{and} \quad f_2(x_{a2})(Section[R', x_{a2}, \cdot]) = 1 \text{ for every } x_{a2} \in B_{a2}.$$

Define

$$B_a = \rho_{a1}^{-1}(B_{a1}) \cap \rho_{a2}^{-1}(B_{a2}).$$

Then we have $\mu_a(B_a) = 1$.

It remains to show that for every $x_a \in B_a$ there is a span coupling of $(f_1(x_{a|1}), f_2(x_{a|2}))$. Let $x_a \in B_a$. Define the set $B_b \subseteq X_{b1} \times X_{b2}$ by

$$B_b = Section[R'_1, x_{a|1}, \cdot] \times Section[R'_2, x_{a|2}, \cdot].$$

Since $R'_1$ and $R'_2$ are Borel sets in $X_{a1} \times X_{b1}$ and $X_{a2} \times X_{b2}$, respectively, the set $B_b$ is Borel in $X_{b1} \times X_{b2}$ (see also Section 2.1). It follows $f_1(x_{a|1}) \otimes f_2(x_{a|2})(B_b) = 1$. Moreover, using the assumptions of the corollary, we have the inclusion $B_b \subseteq Rel[Sp_b]$. It therefore follows that $f_1(x_{a|1}) \otimes f_2(x_{a|2})$ is a weight function for $(f_1(x_{a|1}), Rel[Sp_b], f_2(x_{a|2}))$. According to Theorem 148, there exists a span coupling of $(f_1(x_{a|1}), f_2(x_{a|2}))$. □

## 6.6 Congruence property of simulation and bisimulation

This section provides results stating that the simulation preorder and the bisimulation equivalence are congruences with respect to the span-coupling composition operator. Throughout this section we consider the four simple STSs

$$\mathcal{T}_{a1} = (Sta_{a1}, Act_1, \rightarrow_{a1}), \quad \mathcal{T}_{a2} = (Sta_{a2}, Act_2, \rightarrow_{a2}),$$
$$\mathcal{T}_{b1} = (Sta_{b1}, Act_1, \rightarrow_{b1}), \quad \mathcal{T}_{b2} = (Sta_{b2}, Act_2, \rightarrow_{b2}).$$

Note, the action spaces of the STSs $\mathcal{T}_{a1}$ and $\mathcal{T}_{b1}$ are the same and accordingly, the action spaces of $\mathcal{T}_{a2}$ and $\mathcal{T}_{b2}$ coincide.

**Theorem 166.** *Consider a set $Sync \subseteq Act_1 \cap Act_2$, a span $Sp_a = (Sta_a, \rho_{a1}, \rho_{a2})$ for $(\mathcal{T}_{a1}, \mathcal{T}_{a2})$, as well as a span $Sp_b = (Sta_b, \rho_{b1}, \rho_{b2})$ for $(\mathcal{T}_{b1}, \mathcal{T}_{b2})$. Moreover, let $R_1$ and $R_2$ be a simulations for $(\mathcal{T}_{a1}, \mathcal{T}_{a2})$ and $(\mathcal{T}_{b1}, \mathcal{T}_{b2})$, respectively. Assume that the relations $R_1$ and $R_2$ are compatible with $Sp_a$ and $Sp_b$. Define the two STSs*

$$\mathcal{T}_a = \mathcal{T}_{a1} \parallel_{Sp_a, Sync} \mathcal{T}_{a2} \quad \text{and} \quad \mathcal{T}_b = \mathcal{T}_{b1} \parallel_{Sp_b, Sync} \mathcal{T}_{b2}.$$

*Then the relation $R$ is a simulation for $(\mathcal{T}_a, \mathcal{T}_b)$ where*

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

*Proof.* Assume that the relations $R_1$ and $R_2$ are compatible with $Sp_a$ and $Sp_b$. Let $s_a \in Sta_a$, $s_b \in Sta_b$, $act \in Act_1 \cup Act_2$, and $\mu_a \in Prob[Sta_a]$ be such that

$$\langle s_a, s_b \rangle \in R \quad \text{and} \quad s_a \rightarrow_a \langle act, \mu_a \rangle.$$

Our task is to show that there exists a probability measure $\mu_b \in Prob[Sta_b]$ with the two properties $s_b \rightarrow_b \langle act, \mu_b \rangle$ and $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$.

We consider the case where $act \in Sync$ first. It follows $s_{a|1} \rightarrow_{a1} \langle act, \mu_{a|1} \rangle$ and $s_{a|2} \rightarrow_{a2} \langle act, \mu_{a|2} \rangle$. As $R_1$ and $R_2$ are simulations for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$, respectively, there are $\mu_{b1} \in Prob[Sta_{b1}]$ and $\mu_{b2} \in Prob[Sta_{b2}]$ such that

$$\begin{aligned} &s_{b|1} \rightarrow_{b1} \langle act, \mu_{b1} \rangle \quad \text{and} \quad \langle \mu_{a|1}, \mu_{b1} \rangle \in (R_1)^{\mathrm{wgt}}, \\ &s_{b|2} \rightarrow_{b2} \langle act, \mu_{b2} \rangle \quad \text{and} \quad \langle \mu_{a|2}, \mu_{b2} \rangle \in (R_2)^{\mathrm{wgt}}. \end{aligned}$$

By Theorem 163, there exists a span coupling $\mu_b$ of $(\mu_{b1}, \mu_{b2})$ such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. It moreover holds $s_b \rightarrow_b \langle act, \mu_b \rangle$

Assume $act \in Act_1 \setminus Sync$ now. Then it holds $s_{a|1} \rightarrow_{a1} \langle act, \mu_{a|1} \rangle$ and $\mu_{a|2} = Dirac[s_{a|2}]$. Since $R_1$ is a simulation for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and as we have $\langle s_{a|1}, s_{b|1} \rangle \in R_1$, there is a probability measure $\mu_{b1} \in Prob[Sta_{b1}]$ such that

$$s_{b|1} \rightarrow_{b1} \langle act, \mu_{b1} \rangle \quad \text{and} \quad \langle \mu_{a|1}, \mu_{b1} \rangle \in (R_1)^{\mathrm{wgt}}.$$

Relying on the identity $\mu_{a|2} = Dirac[s_{a|2}]$, we derive $\langle \mu_{a|2}, Dirac[s_{b|2}] \rangle \in (R_2)^{\mathrm{wgt}}$. We can hence apply Theorem 163 that yields a span coupling $\mu_b$ of $(\mu_{b1}, Dirac[s_{b|2}])$ such that $\langle \mu_a, \mu_b \rangle \in R^{\mathrm{wgt}}$. It moreover holds $s_b \rightarrow_b \langle act, \mu_b \rangle$. The remaining case where the action $act$ is contained in $Act_2 \setminus Sync$ can be treated as the previous case. □

Relying on the established general theorem, we obtain the following result for the Cartesian case without any further side constraints:

**Corollary 167.** *Let $Sync \subseteq Act_1 \cap Act_2$ and define the two STSs*

$$\mathcal{T}_a = \mathcal{T}_{a1} \parallel_{CartSp,Sync} \mathcal{T}_{a2} \quad \textit{and} \quad \mathcal{T}_b = \mathcal{T}_{b1} \parallel_{CartSp,Sync} \mathcal{T}_{b2}.$$

*For every $s_{a1} \in Sta_{a1}$, $s_{b1} \in Sta_{b1}$, $s_{a2} \in Sta_{a2}$, and $s_{b2} \in Sta_{b2}$ the two implications below hold:*

*(1)* $\mathcal{T}_{a1}, s_{a1} \preceq \mathcal{T}_{b1}, s_{b1}$ *and* $\mathcal{T}_{a2}, s_{a2} \preceq \mathcal{T}_{b2}, s_{b2}$ *implies* $\mathcal{T}_a, \langle s_{a1}, s_{a2} \rangle \preceq \mathcal{T}_b, \langle s_{b1}, s_{b2} \rangle$.

*(2)* $\mathcal{T}_{a1}, s_{a1} \simeq \mathcal{T}_{b1}, s_{b1}$ *and* $\mathcal{T}_{a2}, s_{a2} \simeq \mathcal{T}_{b2}, s_{b2}$ *implies* $\mathcal{T}_a, \langle s_{a1}, s_{a2} \rangle \simeq \mathcal{T}_b, \langle s_{b1}, s_{b2} \rangle$.

*Proof.* The claim is a direct consequence of Theorem 166 (see also Corollary 164). □

## 6.7 Declarative modelling of stochastic dependencies

As discussed in the introduction of this chapter by means of an example of two cooling systems dissipating the heat produced by an energy-adaptive server (see also Figure 6.1), there are situations where one wants to abstract from the precise operational behaviour causing specific stochastic dependencies. We present an application of the span-coupling composition operator illustrating how to include additional stochastic information in a declarative manner.

**Definition 168.** Let $AP$ be a finite set. Given an STS $\mathcal{T}$ with state space $Sta$, an *AP-labelling function* (*for* $\mathcal{T}$) is a Borel function $Lab\colon Sta \rightsquigarrow AP$. ⌟

Intuitively, the elements of $AP$ stand for atomic proposition expressing facts known about states of the STS under consideration. An $AP$-labelling function assigns those atomic propositions to a state that are satisfied in that state. To keep the theory in this section simple and to avoid extensive measure-theoretic considerations, we only focus on finite sets of atomic propositions. As a consequence, the powerset of $AP$ is finite and hence, trivially forms a Polish space whose induced Borel sigma algebra is given by the discrete sigma algebra (see Example 3 (1)). It in particular follows that a function $Lab\colon Sta \rightsquigarrow AP$ is Borel if for every atomic proposition $ap$ the set $\{s \in Sta\ ;\ ap \in Lab(s)\}$ is Borel in $Sta$.

**Compositions based on coupling-constraint functions.** Let $\mathcal{T}_1 = (Sta_1, Act_1, \rightarrow_1)$ and $\mathcal{T}_2 = (Sta_2, Act_2, \rightarrow_2)$ be two simple STSs as well as $Sync \subseteq Act_1 \cap Act_2$. Moreover, let $AP$ be a finite set as well as $Lab_1$ and $Lab_2$ be $AP$-labelling functions for $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively. Define the Borel function $Lab\colon Sta_1 \times Sta_2 \rightsquigarrow AP$,

$$Lab(s_1, s_2) = Lab_1(s_1) \cup Lab_2(s_2).$$

Moreover, pick a function

$$\mathfrak{C}\colon 2^{AP} \times Sync \rightsquigarrow Prob[2^{AP}]$$

to which we also refer as a coupling-constraint function. The function $\mathfrak{C}$ represents the additional stochastic information caused by the interaction of $\mathcal{T}_1$ and $\mathcal{T}_2$. To be more precise, given transitions $s_1 \rightarrow_1 \langle act, \mu_1 \rangle$ and $s_2 \rightarrow_2 \langle act, \mu_2 \rangle$ with $act \in Sync$, only those couplings of $(\mu_1, \mu_2)$ are considered in a composition of $\mathcal{T}_1$ and $\mathcal{T}_2$ with respect to $\mathfrak{C}$ that are compatible with some probability measure contained in the set $\mathfrak{C}(Lab(s_1, s_2), act)$. As in Example 152, this additional available stochastic information is encoded in a third STS controlling the interplay of the STSs $\mathcal{T}_1$ and $\mathcal{T}_2$: introduce the simple STS

$$\mathcal{T}_{\mathfrak{C}} = (Sta_1 \times Sta_2, Act, \rightarrow_{\mathfrak{C}})$$

whose transition relation $\rightarrow_{\mathfrak{C}}$ is given by the following inference rule: for every states $s_1 \in Sta_1$ and $s_2 \in Sta_2$, action $act \in Act$, and probability measure $\mu \in Prob[Sta]$,

$$\frac{act \in Sync \quad \text{and} \quad Lab_\sharp(\mu) \in \mathfrak{C}(Lab(s_1, s_2), act)}{\langle s_1, s_2 \rangle \rightarrow_{\mathfrak{C}} \langle act, \mu \rangle}.$$

Following Example 152, this leads to the following definition:

**Definition 169.** The notions are given as before. Then the *coupling composition of* $(\mathcal{T}_1, \mathcal{T}_2)$ *concerning* $(\mathfrak{C}, Sync)$ is given by the STS

$$\mathcal{T}_1 \parallel_{\mathfrak{C},Sync} \mathcal{T}_2 = (\mathcal{T}_1 \parallel_{CartSp,Sync} \mathcal{T}_2) \parallel_{IdSp,Sync} \mathcal{T}_{\mathfrak{C}}.$$

┘

Roughly speaking, stochastic information referring to the interaction of $\mathcal{T}_1$ and $\mathcal{T}_2$ are encoded by the third STS $\mathcal{T}_{\mathfrak{C}}$, which sorts out all the couplings being not compatible with the coupling-constraint function $\mathfrak{C}$ under consideration. Thus, stochastic information concerning the interaction of $\mathcal{T}_1$ and $\mathcal{T}_2$ are specified declaratively by the function $\mathfrak{C}$. We demonstrate this by means of the following example:

**Example 170.** We continue our running example from the introduction of this chapter involving two cooling systems modelled by the STSs $\mathcal{T}_{\text{Cool},1}$ and $\mathcal{T}_{\text{Cool},2}$ dissipating the heat produced by the server $\mathcal{T}_{\text{Serv}}$ (see Figure 6.1). Let

$$Sync = \{\text{fail}\}$$

be the set of synchronisation actions as well as

$$AP = \{\text{Low}, \text{High}, \text{Broken}_1, \text{Broken}_2, \text{Intact}_1, \text{Intact}_2\}$$

be the set of atomic propositions. In the introduction of this chapter, depending on the actual performance mode of the server, we have already specified a coupling-constraint function $\mathfrak{C}$ concerning $(AP, Sync)$ in a symbolical way: indeed, the set $\mathfrak{C}(\{\text{Low}\}, \text{fail})$ contains precisely those $\zeta \in Prob[2^{AP}]$ where

$$\zeta(\text{Broken}_1 \wedge \text{Broken}_2) = \zeta(\text{Broken}_1) \cdot \zeta(\text{Broken}_2)$$

and accordingly, the set $\mathfrak{C}(\{\text{High}\}, \text{fail})$ consists precisely of those probability measures $\zeta \in Prob[2^{AP}]$ such that

$$\zeta((\text{Broken}_1 \wedge \text{Broken}_2) \vee (\text{Intact}_1 \wedge \text{Intact}_2)) > 0.9,$$

Moreover, for every $B \in 2^{AP} \setminus \{\{\text{Low}\}, \{\text{High}\}\}$ we have $\mathfrak{C}(B, \text{fail}) = Prob[2^{AP}]$. Recall, if the server operates in the low mode, the cooling systems behave stochastically independent after a failure. In contrast, there is a stochastic dependence when the server is in the high power mode. The composition of the cooling systems and the server with respect to the available stochastic information is given by the STS

$$\mathcal{T} = \big(\mathcal{T}_{\text{Cool},1} \parallel_{CartSp,Sync} \mathcal{T}_{\text{Cool},2}\big) \parallel_{\mathfrak{C},Sync} \mathcal{T}_{\text{Serv}}.$$

When composing the cooling systems in a first step, we do not rely on any stochastic assumptions and take all the possible couplings into account. However, when adding the STS for the server in a next step, we include the available stochastic information induced by the internal power-management strategy of the server. ⌟

Our approach allows for a systematic system design, facilitate the interchangeability and reusability of components, and thus also eases the maintainability. For instance, if there is another device using the batteries with an alternative internal power-management strategy, we can use the STSs $\mathcal{T}_{\text{Cool},1}$ and $\mathcal{T}_{\text{Cool},2}$, solely adapt the STS $\mathcal{T}_{\text{Serv}}$ and the corresponding coupling-constraint function.

**Congruence result.** We extend our congruence property from Section 6.6 to the coupling composition operator relying on a specific coupling-constraint function. For this purpose we consider the four simple STSs

$$\begin{aligned}
\mathcal{T}_{a1} &= (Sta_{a1}, Act_1, \rightarrow_{a1}), \quad \mathcal{T}_{a2} = (Sta_{a2}, Act_2, \rightarrow_{a2}),\\
\mathcal{T}_{b1} &= (Sta_{b1}, Act_1, \rightarrow_{b1}), \quad \mathcal{T}_{b2} = (Sta_{b2}, Act_2, \rightarrow_{b2}).
\end{aligned}$$

Let $AP$ be a finite set as well as $Lab_{a1}$, $Lab_{b1}$, $Lab_{a2}$, and $Lab_{b2}$ be $AP$-labelling functions for $\mathcal{T}_{a1}$, $\mathcal{T}_{a2}$, $\mathcal{T}_{b1}$, and $\mathcal{T}_{b2}$, respectively. Moreover, let $R_1$ be a simulation for $(\mathcal{T}_{a1}, \mathcal{T}_{b1})$ and $R_2$ be a simulation for $(\mathcal{T}_{a2}, \mathcal{T}_{b2})$. Assume that for every states $s_{a1} \in Sta_{a1}$, $s_{a2} \in Sta_{a2}$, $s_{a1} \in Sta_{b1}$, and $s_{a2} \in Sta_{b2}$,

$$\begin{aligned}
\langle s_{a1}, s_{b1}\rangle \in R_1 \quad &\text{implies} \quad Lab_{a1}(s_{a1}) = Lab_{b1}(s_{b1}),\\
\langle s_{a2}, s_{b2}\rangle \in R_2 \quad &\text{implies} \quad Lab_{a2}(s_{a2}) = Lab_{b2}(s_{b2}).
\end{aligned}$$

Intuitively, the latter requirements on $R_1$ and $R_2$ ensure that only those states are related where the same atomic propositions hold. In a state-based setting for simulation and bisimulation, the latter condition is standard (see, e.g., [BK08]). We obtain the following congruence result:

**Proposition 171.** *The notions are given as before. Consider a set $Sync \subseteq Act_1 \cap Act_2$ as well as a function $\mathfrak{C}\colon 2^{AP} \times Sync \rightsquigarrow Prob[2^{AP}]$. Define the STSs*

$$\mathcal{T}_a = \mathcal{T}_{a1} \parallel_{\mathfrak{C},Sync} \mathcal{T}_{a2} \quad \textit{and} \quad \mathcal{T}_b = \mathcal{T}_{b1} \parallel_{\mathfrak{C},Sync} \mathcal{T}_{b2}.$$

*as well as the relation*

$$R = \{\langle s_a, s_b\rangle \in Sta_a \times Sta_b \ ; \ \langle s_{a|1}, s_{b|1}\rangle \in R_1 \textit{ and } \langle s_{a|2}, s_{b|2}\rangle \in R_2\}.$$

*Then the relation $R$ is a simulation for $(\mathcal{T}_a, \mathcal{T}_b)$.*

*Proof.* According to Theorem 166, the relation $R$ is a simulation for $(\mathcal{T}_a', \mathcal{T}_b')$ where

$$\mathcal{T}_a' = \mathcal{T}_{a1} \parallel_{CartSp,Sync} \mathcal{T}_{a2} \quad \text{and} \quad \mathcal{T}_b' = \mathcal{T}_{b1} \parallel_{CartSp,Sync} \mathcal{T}_{b2}.$$

Define the sets $Sta_a = Sta_{a1} \times Sta_{a2}$ and $Sta_b = Sta_{b1} \times Sta_{b2}$. Introduce the functions $Lab_a\colon Sta_a \rightsquigarrow AP$, $Lab_a(s_{a1}, s_{a2}) = Lab_{a1}(s_{a2}) \cup Lab_{a2}(s_{a2})$ and $Lab_b\colon Sta_b \rightsquigarrow AP$, $Lab_b(s_{b1}, s_{b2}) = Lab_{b1}(s_{b1}) \cup Lab_{b2}(s_{b2})$. Clearly, for every $\langle s_a, s_b\rangle \in R$ we have the identity $Lab_a(s_a) = Lab_b(s_b)$. To conclude that $R$ is a simulation for $(\mathcal{T}_a, \mathcal{T}_b)$, it hence suffices to show that for every probability measures $\mu_a \in Prob[Sta_a]$ and $\mu_b \in Prob[Sta_b]$ the following implication holds:

$$\langle \mu_a, \mu_b\rangle \in R^{\mathrm{wgt}} \quad \text{implies} \quad (Lab_a)_\sharp(\mu_a) = (Lab_b)_\sharp(\mu_b).$$

Let $\mu_a \in Prob[Sta_a]$ and $\mu_b \in Prob[Sta_b]$ be such that $\langle \mu_a, \mu_b\rangle \in R^{\mathrm{wgt}}$ as well as $B \subseteq 2^{AP}$ be a set. Since $Lab_a(s_a) = Lab_b(s_b)$ for every $\langle s_a, s_b\rangle \in R$, we obtain

$$R \cap ((Lab_a)^{-1}(B) \times Sta_b) = R \cap (Sta_a \times (Lab_b)^{-1}(B)).$$

Inspecting the argument in Remark 34, this yields

$$\mu_a((Lab_a)^{-1}(B)) = \mu_b((Lab_b)^{-1}(B))$$

that finally justifies $(Lab_a)_\sharp(\mu_a) = (Lab_b)_\sharp(\mu_b)$. $\square$

# 7 Relations to models from the literature

The following material summarises the contribution of this thesis for prominent modelling formalisms from the literature. Despite the measurability requirements investigated throughout this thesis, our investigations extend existing frameworks properly.

We discuss consequences for discrete stochastic systems first covering probabilistic automata [Seg95], discrete Markov decision processes [Put94], and discrete-time as well as continuous-time Markov chains [BHHK03, DP03]. Moreover, as discrete STSs allow for non-deterministic choices between a countable number of alternatives, discrete STS also include continuous-time Markov decision processes [Put94, NK07], interactive Markov chains [Her02], and Markov automata [EHZ10].

Having considered discrete STSs, we focus on purely stochastic STSs (see also Section 4.8) that in particular cover continuous-time Markov chains [BHHK03], semi Markov process [Whi80, LHK01, GJP06], and discrete-time stochastic hybrid automata [AKLP10, AKM11, SA13]. The Souslin condition for purely stochastic as well as the property of being Borel concerning the hit sigma algebra appear natural in the context of purely stochastic STSs (see also Remark 95) and hence, the above mentioned examples of models are covered by the main results of Chapters 4 and 5.

Labelled Markov processes (LMPs) [BDEP97, Des99, DEP02, Pan09] as well as nondeterministic labelled Markov processes (NLMPs) [DTW12, Wol12] yield other prominent subclasses of simple STSs. Assuming the underlying action space is countable, our results of Chapters 4 and 5 are applicable for point-wise image-finite NLMPs and thus also for LMPs. This in particular extends the existing literature by a logical characterisation for the simulation preorder and the bisimulation equivalence in terms of comparable expressive modal logics. Moreover, to the best of our knowledge, simulations has been not considered for NLMPs before. Besides this, we provide an example showing that the class of all NLMPs is not closed under parallel composition. This also influences the investigations on a compositional framework for NLMPs [DLM16].

The stochastic-hybrid-system model in [Hah13, HH13] has an NLMP-like semantics, in particular, the unfolded STS is Borel concerning the hit sigma algebra. A similar discussion applies for HMODEST [BDHK06, HHHK12], a powerful modelling language for stochastic hybrid systems covering several prominent models from the literature (see Table 4 in [BDHK06] and Table 3 in [HHHK12]). The latter yield a nice and powerful modelling language for subclasses of STSs.

Stochastic optimal control [DIY79, BS96] aims to design a strategy for input variables with respect to the presence of stochastic noise satisfying a specified objective while maximising or minimising a given cost functional. This functional may refer to the energy or utility accumulated along sample paths (see also Section 5.8). Stochastic control problems appear in many areas of economy and engineering for the design of, e.g., portfolios of safe and risky assets, consumption-investment strategies, acceleration behaviours for the steering of vehicles, and strategies for job-scheduling scenarios.

A convenient and prominent formalism for modelling the underlying stochastic process of a stochastic control problem is given by controlled Markov processes (CMPs). Interestingly, CMPs are prominent in both the mathematics community, e.g., [DIY79, BS96], and the computer-science community, e.g., [ZEM$^+$14, TMKA16]. CMPs are also formalised in terms of controlled discrete-time stochastic hybrid automata [APLS08, AKLP10, SA13, TMKA13, ZEM$^+$14, TMKA16]. The latter cited articles provide approximate verification techniques, model-checking approaches, as well as corresponding practical experiments. The case studies rely, e.g., on the software tool Faust2 [SGA15] that allows for exporting abstracted models to the prominent probabilistic model checker Prism [KNP11].

We show that CMPs are covered by Souslin STSs and moreover, that for continuous CMPs the Souslin-bisimulation equivalence and the bisimulation equivalence are the same and accordingly, for simulations. This provides an important connection to our contributions of Chapter 4 and, to the best of our knowledge, enables a trace-distribution result for CMPs that has been not known before.

Dynamics obtained by interaction of (digital) sensors with the (continuous) environment lead to hybrid systems. A well-known modeling formalism for hybrid systems is given by hybrid automata [ACH$^+$95, Hen96] consisting of a discrete control structure and flow functions that model the evolution of continuous variables when time passes. The research on stochastic extensions of hybrid automata [Spr01, Pla08, HHHK12, ZSR$^+$12, Hah13] are motivated, e.g., by imperfect sensors that deliver vague information. Undecidability results for stochastic hybrid automata initiate the research on subclasses of the powerful formalism. For instance, the subclass consisting of all (probabilistic) rectangular hybrid automata play an important role within the theory of hybrid systems since many decision problems, e.g., the reachability problem, turn out to be decidable under mild side constraints [ACH$^+$95, Hen96, Kop96, Spr01].

In the last part of this chapter, we introduce stochastic rectangular hybrid automata (SRHAs) and show that such an automaton can be unfolded into a Souslin STSs exploiting the mean value theorem. As a consequence, due to the expressiveness of SRHAs, it follows that many modelling formalisms from the literature are covered by Souslin STSs. This yields a foundation for further work investigating other classes of systems and their respective behavioural relations.

## 7.1 Discrete stochastic transition systems

We emphasise that the action space for discrete STS is not restricted to be a countable set. The precise definition is as follows:

**Definition 172.** We call an STS $\mathcal{T} = (Sta, Act, \rightarrow)$ *discrete* provided it is image finite and the set *Sta* is countable. ⌟

Discrete STSs cover probabilistic automata [Seg95], discrete Markov decision processes [Put94], and discrete-time as well as continuous-time Markov chains [BHHK03, DP03]. Moreover, since discrete STSs include uncountable action spaces and finite non-determinism, discrete STS include continuous-time Markov decision processes [Put94, NK07], interactive Markov chains [Her02], and Markov automata [EHZ10]. Continuous-time Markov decision processes serve, e.g., as a semantical model for stochastic Petri nets [CMBC93].

By Remark 10, it is easy to see that every discrete STS is Souslin. Moreover, we also obtain that every discrete STS is Borel concerning the hit sigma algebra. This yields the following summary of our results:

**Proposition 173.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be a non-blocking discrete STS, $\mathcal{A}$ be an action event family, and rew be a reward function. For every $s_a, s_b \in Sta$ the following statements hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\text{sou}} s_b$ *iff* $s_a \preceq_{\bullet} s_b$ *iff* $s_a \preceq^{\exists} s_b$.

(2) $s_a \preceq s_b$ *implies* $s_a \leq^{\text{tr}} s_b$.

(3) $s_a \simeq s_b$ *iff* $s_a \simeq^{\text{sou}} s_b$ *iff* $s_a \simeq_{\circ} s_b$ *iff* $s_a \simeq^{*} s_b$.

(4) $s_a \simeq s_b$ *implies* $s_a =^{\text{tr}} s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth.*

*Proof.* The proposition follows directly from Corollary 119 and Theorem 124. □

Clearly, for certain subclasses of discrete STSs the previous proposition is not new such as for probabilistic automata and Markov chains. However, we emphasise that the result covers the generative model in the classification of [GSS95] and in this context, to the best of our knowledge, a generic logical characterisation of the simulation preorder and the bisimulation equivalence has been not proven before. Besides this, as the model under consideration includes uncountable action spaces, the result in [NK07] is extended by a complete characterisation of both the simulation preorder and the bisimulation equivalence. Moreover, Proposition 173 (together with Proposition 136 concerning the preservation of expected values) also covers results in [DP03] providing a logical characterisation of the bisimulation equivalence for continuous-time Markov processes.

## 7.2 Purely stochastic systems

According to Theorem 100, the Souslin-bisimulation equivalence and the trace-distribution equivalence are the same for every deterministic purely stochastic Souslin STSs. Thanks to the theory developed in Chapter 5 concerning relations induced by temporal logics, we can extend the latter mentioned result. Recalling Remark 95, a purely stochastic STS $\mathcal{T} = (Sta, Act, \mathfrak{K})$ is Souslin precisely when the corresponding control law $\mathfrak{K}$ constitutes a Borel function. It is easy to see that the latter statement holds precisely when $\mathcal{T}$ is Borel concerning the hit sigma algebra. Hence, we are in the situation of Theorem 124 that finally enables the following result:

**Proposition 174.** *Let $\mathcal{T} = (Sta, Act, \mathfrak{K})$ be a purely stochastic Souslin STS, $\mathcal{A}$ be an action event family, and rew be a reward function. For every $s_a, s_b \in Sta$ the following statement holds:*

(1) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$ *iff* $s_a \simeq_{\circ} s_b$ *iff* $s_a \simeq^{*} s_b$.

(2) $s_a \simeq s_b$ *implies* $s_a =^{\mathrm{tr}} s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth. The implication in statement (2) is an equivalence if the purely stochastic Souslin STS $\mathcal{T}$ is deterministic.*

*Proof.* Every purely stochastic Souslin STS is non-blocking, image-finite, and Borel concerning the hit sigma algebra. Therefore, Theorem 100, Corollary 119 as well as Theorem 124 yield the claims. □

Focusing on purely stochastic Souslin STSs, every satisfaction set of an APCTL$^{*}$ is Borel and hence, one can drop the outer-measure function in Definition 113 where we declare the satisfaction relation for APCTL$^*$ path-measure formulas. The reason is that for purely stochastic Souslin STSs there exists exactly one $s$-scheduler for every state $s$ (see Remark 96) and moreover, the function that maps every state $s$ to the corresponding path measure $Pr[s]$ turns out to be Borel (see also Lemma 101). Thus, one can rely on an easy induction over the construction of APCTL$^*$ formulas in order to show that all the corresponding satisfaction sets are Borel.

**Embedding models from the literature.** The model of (deterministic) purely stochastic Souslin STS serve as a semantical model for prominent stochastic-hybrid-system modelling formalisms from the literature. In what follows we provide an short overview on some important examples:

*Stochastic timed automata.* The stochastic extension of classical timed automata [AD94] given by stochastic timed automata (STAs) [BBB$^+$14, BBCMar] is subsumed by purely stochastic STSs, i.e., the unfolding of every STA forms a purely stochastic Souslin STS. The behaviour of an STA is determined by clocks whose values increase with slope one when (continuous) time passes. Invariant conditions in locations and guards of edges between locations constrain the values for clocks. In an STA both delays in locations and discrete choices over enabled actions are made randomly. Provided the different edges from one location in an STA are labelled by distinct actions, the induced purely stochastic STS obtained by an unfolding is even deterministic.

*Semi Markov processes.* The residence times in locations of a semi Markov process (SMP) [Whi80, LHK01, GJP06] follow an arbitrary distribution, in particular, SMPs cover continuous-time Markov processes. A SMP induces a purely stochastic Souslin STS with an action space $Act = \mathbb{R}_{\geq 0}$ and a naturally given control law $\mathfrak{K}$. Intuitively, the real number $\mathfrak{K}(s)([\underline{t}, \overline{t}] \times S)$ stands for the probability of moving from state $s$ to a state contained in the set $S$ with a delay of at least $\underline{t}$ and at most $\overline{t}$ time units.

*Discrete-time stochastic hybrid automata.* purely stochastic Souslin STSs cover the stochastic extension of hybrid automata given by discrete-time stochastic hybrid automata (DTSHAs) [AKLP10, AKM11, SA13]. Every DTSHA comprises a discrete control graph with locations and jumps in-between as well as real-valued variables. In a location the values of the variables are governed by a stochastic flow, e.g., specified by discrete-time stochastic differential equations. Within a jump, variables can be updated instantaneously with respect to an arbitrary continuous distribution.

*Pure jump Markov processes.* purely stochastic Souslin STSs subsume the model in [EMA16] called pure jump Markov processes (PJMPs). A PJMP evolves in continuous time and has instantaneous discrete movements with exponential distributed random times.

*Stochastic differential equations.* Assuming certain regularity assumption, it is well-known that the solution of a stochastic differential equation (SDE) can be represented by a function $f\colon \mathbb{R} \times \mathbb{R}_{\geq 0} \to Prob[\mathbb{R}]$ satisfying certain requirements, e.g., the Chapman-Kolmogorov equation. Assuming the current state is given by $r$, the value $f(r,t)(B)$ stands for the probability of entering a state contained in the Borel set $B$ after exact time $t$ elapsed. Suppose a Borel function $g\colon \mathbb{R} \to Prob[\mathbb{R}_{\geq 0}]$. Intuitively, the probability measure $g(r)$ specifies the residence time in state $r \in \mathbb{R}$ that might be, e.g., exponential distributed. The composition of the functions $f$ and $g$ yields a single Borel function $\mathfrak{K}\colon \mathbb{R} \to Prob[\mathbb{R}_{\geq 0} \times \mathbb{R}]$ (see also Lemma 1.38 [Kal02]) forming a control law of a purely stochastic Souslin STS. Roughly speaking, the function $\mathfrak{K}$ combines the stochastic evolution of states specified by $f$ and the stochastic timing given by $g$. Consequently, purely stochastic Souslin STS cover general stochastic hybrid automata [BL04, BLB05] (see also [BSA04]), which can be seen as a continuous-time analogue of DTSHAs where the stochastic flows of variables in the discrete locations are specified by SDEs.

**Revisiting Proposition 174.** The major difficulties in proving that the bisimulation equivalence is finer than the trace-distribution equivalence arises from the measurability requirements on schedulers (see discussions in Chapter 4). Recall, we exploit measurable-selection principles to treat these measurability issues. In purely stochastic STSs non-determinism is absent and hence, schedulers can be omitted for the resolution of the non-determinism. It turns out that this fact can be used to substantially simplify the argument for showing that the bisimulation equivalence is subsumed by the trace-distribution equivalence.

Let $\mathcal{T} = (Sta, Act, \to)$ be a a purely stochastic Souslin STS, $s_a, s_b \in Sta$ be states, and $R$ be a bisimulation such that $\langle s_a, s_b \rangle \in R$. According to Theorem 118, we can safely assume that the relation $R$ is smooth. The following material provides a direct argument showing that $s_a =^{\text{tr}} s_b$ without using Corollary 119. In particular, a measurable-selection principle is avoided and thus, we obtain a simpler argument for Proposition 174 (2).

To conclude $s_a =^{\text{tr}} s_b$, relying on Lemma 81 as well as Remark 96, it suffices show that for every $n \in \mathbb{N}$ it holds

$$\langle Pr_n[s_a], Pr_n[s_b] \rangle \in (R^{\text{path},n})^{\text{wgt}}.$$

We perform a induction over the natural numbers. The induction base follows trivially since it holds $Pr_0[s_a] = Dirac[s_a]$, $Pr_0[s_b] = Dirac[s_b]$, and $R^{\text{path},n} = R$. Let $n \in \mathbb{N}$ and assume $\langle Pr_n[\mu_a], Pr_n[\mu_b] \rangle \in (R^{\text{path},n})^{\text{wgt}}$ as an induction hypothesis. As $R$ is smooth and relying on Remark 37, it is easy to see that the two relations $R^{\text{path},n}$ and $R^{Act}$ are smooth.

Moreover, there are two Polish spaces $Y$ and $Z$ as well as two Borel functions $f\colon \mathit{Path}_n \to Y$ and $g\colon \mathit{Act} \times \mathit{Sta} \to Z$ such that the following statements hold: for every finite paths $\hat{\pi}_a, \hat{\pi}_b \in \mathit{Path}_n$,

$$\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R^{\mathrm{path},n} \quad \text{iff} \quad f(\hat{\pi}_a) = f(\hat{\pi}_b)$$

and accordingly, for every action-state pairs $\langle act_a, s_a \rangle, \langle act_b, s_b \rangle \in \mathit{Act} \times \mathit{Sta}$,

$$\langle \langle act_a, s_a \rangle, \langle act_b, s_b \rangle \rangle \in R^{\mathit{Act}} \quad \text{iff} \quad g(act_a, s_a) = g(act_b, s_b).$$

The Borel function $h\colon \mathit{Path}_{n+1} \to Y \times Z$ is introduced as follows: for every finite path $\hat{\pi} \in \mathit{Path}_n$, action $act \in \mathit{Act}$, and state $s \in \mathit{Sta}$ let

$$h(\hat{\pi}\, act\, s) = \langle f(\hat{\pi}), g(act, s) \rangle.$$

Obviously, for every finite paths $\hat{\pi}_a, \hat{\pi}_b \in \mathit{Path}_{n+1}$ it holds

$$\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R^{\mathrm{path},n+1} \quad \text{iff} \quad h(\hat{\pi}_a) = h(\hat{\pi}_b).$$

In particular, relying again on Remark 37, the relation $R^{\mathrm{path},n+1}$ is smooth.

Let $B_X \subseteq X$ and $B_Y \subseteq Y$ be Borel sets. Using that the relation $R$ is a bisimulation and applying Corollary 41 as well as Remark 97, for every finite paths $\hat{\pi}_a, \hat{\pi}_b \in \mathit{Path}_n$ it holds

$$\langle \hat{\pi}_a, \hat{\pi}_b \rangle \in R^{\mathrm{path},n} \quad \text{implies} \quad \mathfrak{K}(\mathrm{Last}(\hat{\pi}_a))(g^{-1}(B_Y)) = \mathfrak{K}(\mathrm{Last}(\hat{\pi}_b))(g^{-1}(B_Y)).$$

The set $f^{-1}(B_X)$ is $R$-stable. Thus, relying on a standard argument from measure theory concerning point-wise approximations of Borel functions by sequences of step functions (see, e.g., Teil III in [Sch08]), for every Borel function $\xi\colon \mathit{Path}_n \times \mathit{Path}_n \to [0,1]$ it holds (see also Remark 34):

$$\int_{f^{-1}(B_X) \times \mathit{Path}_n} \xi(\hat{\pi}_a, \hat{\pi}_b)\ d\, W_n(\hat{\pi}_a, \hat{\pi}_b) = \int_{\mathit{Path}_n \times f^{-1}(B_X)} \xi(\hat{\pi}_a, \hat{\pi}_b)\ d\, W_n(\hat{\pi}_a, \hat{\pi}_b).$$

Putting things together, we hence obtain

$$\begin{aligned}
& \mathit{Pr}_{n+1}[s_a](h^{-1}(B_X \times B_Y)) \\
= \quad & \int_{f^{-1}(B_X)} \mathfrak{K}(\mathrm{Last}(\hat{\pi}_a))(g^{-1}(B_Y))\ d\, \mathit{Pr}_n[s_a](\hat{\pi}_a) \\
= \quad & \int_{f^{-1}(B_X) \times \mathit{Path}_n} \mathfrak{K}(\mathrm{Last}(\hat{\pi}_a))(g^{-1}(B_Y))\ d\, W_n(\hat{\pi}_a, \hat{\pi}_b) \\
= \quad & \int_{\mathit{Path}_n \times f^{-1}(B_X)} \mathfrak{K}(\mathrm{Last}(\hat{\pi}_b))(g^{-1}(B_Y))\ d\, W_n(\hat{\pi}_a, \hat{\pi}_b)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{f^{-1}(B_X)} \mathfrak{K}(\mathrm{Last}(\hat{\pi}_b))(g^{-1}(B_Y)) \; d\,Pr_n[s_b](\hat{\pi}_b) \\
&= Pr_{n+1}[s_b](h^{-1}(B_X \times B_Y)).
\end{aligned}
$$

According to Carathéodory uniqueness theorem (see Section 2.1), we derive

$$h_\sharp(Pr_{n+1}[s_a]) = h_\sharp(Pr_{n+1}[s_b]).$$

Thus, Corollary 41 yields $\langle Pr_{n+1}[\mu_a], Pr_{n+1}[\mu_b]\rangle \in (R^{\text{path},n+1})^{\text{wgt}}$. As discussed before, this shows $s_a =^{\text{tr}} s_b$.

## 7.3 Labelled Markov processes

The original paper [BDEP97] on labelled Markov processes (LMPs) provide fundamental contributions regarding stochastic systems with uncountable state and action spaces. The theory has been adapted and extended in various directions [DEP02, DGJP03, DDLP06, DTW12, FKP17], and is aggregated in the thesis [Des99] as well as the textbook [Pan09]. This section summarises our contributions regarding the theory on LMPs. Let us first recall the definition of the model:

**Definition 175.** A non-blocking simple STS $\mathcal{T} = (Sta, Act, \rightarrow)$ is called a *labelled Markov process* (*LMP*) provided there are a family $(\mathfrak{L}_{act})_{act \in Act}$ of Borel functions

$$\mathfrak{L}_{act} \colon Sta \to Prob[Sta]$$

and a family $(S_{act})_{act \in Act}$ of Borel sets

$$S_{act} \subseteq Sta$$

such that for every $s \in Sta$, $act \in Act$, and $\mu \in Prob[Sta]$ it holds

$$s \rightarrow \langle act, \mu \rangle \quad \text{iff} \quad s \in S_{act} \text{ and } \mathfrak{L}_{act}(s) = \mu.$$

⌟

An important difference concerning the standard literature on LMPs is given by the fact that we focus on probability measures rather than sub-probability measures. Here, a measure $\tilde{\mu}$ on the Polish space $Sta$ is called a sub-probability measure provided one has $\tilde{\mu}(Sta) \leq 1$. Hence, to be precise, our definition of an LMP covers only those systems from the LMP literature that only include sub-probability measures with $\tilde{\mu}(Sta) \in \{0, 1\}$. One motivation of sub-probability measures is that one wants to make a distinction between states where a specific action is enabled and where not (see, e.g., Section 7.4 in [Pan09]). For instance, considering a sub-probability measures $\tilde{\mu}$ with $\tilde{\mu}(Sta) = 0$ such that $s \to \langle act, \tilde{\mu} \rangle$, then this intuitively means that the action *act* is not enabled at $s$. This kind of distinction between states is also included within our formalism since the set $Enabled[s, act]$ may be empty that is precisely the case when $s \notin S_{act}$.

**Proposition 176.** *Let $\mathcal{T} = (Sta, Act, \to)$ be an LMP such that the action space Act is countable. For every states $s_a, s_b \in Sta$ the following four statements hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\text{sou}} s_b$ *iff* $s_a \preceq_{\bullet} s_b$ *iff* $s_a \preceq^{\exists} s_b$.

(2) $s_a \preceq s_b$ *implies* $s_a \leq^{\text{tr}} s_b$.

(3) $s_a \simeq s_b$ *iff* $s_a \simeq^{\text{sou}} s_b$ *iff* $s_a \simeq_{\circ} s_b$ *iff* $s_a \simeq^{*} s_b$.

(4) $s_a \simeq s_b$ *implies* $s_a =^{\text{tr}} s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth.*

*Proof.* According to Theorem 129, every LMP with a countable action space is Borel concerning the hit sigma algebra. As a side remark, thanks to Corollary 130, every LMP with a countable action space is hence also Souslin. The proposition under consideration is thus a summary of Theorem 128 and Corollary 131. □

Let us first note that the notions of simulation and bisimulation used throughout this thesis and the corresponding definitions in the LMP literature coincide. The simple reason is that according to Proposition 176 the simulation relation $\preceq$ is weakly smooth and the bisimulation relation $\simeq$ is smooth and thus, the characterisations of Theorems 39 and 40 are applicable. The latter observation indeed directly implies that the settings of this thesis and the LMP literature are the same. In this thesis the state space of an LMP is required to form a Polish space and hence, we also derive that event bisimulation [DDLP06] and the bisimulation notion of this thesis are the same if the underlying action space of the LMP under consideration is countable.

The recent contribution [FKP17] provides elegant proofs showing that for every states $s_a$ and $s_b$ of an LMP it holds $s_a \preceq s_b$ iff $s_a \preceq_{\bullet} s_b$ and accordingly, $s_a \simeq s_b$ iff $s_a \simeq_{\circ} s_b$. In particular, the facts that the relations $\preceq$ and $\simeq$ are weakly smooth and smooth, respectively, are also not new. In fact, the part of Chapter 5 where we provide a charactersation of the simulation preorder and the bisimulation equivalence in terms of the weak modal logics $\mathrm{APCTL}_\bullet$ and $\mathrm{APCTL}_\circ$, respectively, heavily rely on the techniques in [FKP17] (see Section 5.4 and also Section 5.7). Having the special structure of LMPs in mind, $\mathrm{APCTL}_\bullet$ and $\mathrm{APCTL}_\circ$ and the corresponding modal logics in [FKP17] are basically the same besides the fact that the probability modality in [FKP17] can only be bound probabilities from below. We also refer to Chapter 4 of the thesis [Des99] providing various examples and intuitive explanations of the considered logics for LMPs. To the best of our knowledge, apart from the previous mentioned results, the remaining implications in Proposition 176 properly extend the up to now developed theory for LMPs.

## 7.4 Non-deterministic labelled Markov processes

The theory for LMPs has been conservatively extended in [DTW12, Wol12, DLM16] for non-deterministic labelled Markov processes that include internal non-determinism. The definition of the latter mentioned model is as follows:

**Definition 177.** A non-blocking simple STS $\mathcal{T} = (Sta, Act, \rightarrow)$ is called a *non-deterministic labelled Markov process* (*NLMP*) provided for every action $act \in Act$ the following two conditions hold:

(1) $Enabled[s, act]$ is Borel in $Prob[Sta]$ for every state $s \in Sta$.

(2) For every Borel set $P \subseteq Prob[Sta]$ the following set is Borel in $Sta$,

$$\{s \in Sta \ ; Enabled[s, act] \cap P \neq \varnothing\}.$$

┘

The requirements for NLMPs are closely connected to the property for an STS of being Borel concerning the hit sigma algebra (see Theorem 129). Regarding this issue, let us first state the following remark.

**Remark 178.** Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an NLMP where the action space $Act$ is countable. According to Theorem 129, the NLMP $\mathcal{T}$ is Borel concerning the hit sigma algebra. By Corollary 130, the NLMP $\mathcal{T}$ is Souslin provided for every $s \in Sta$ and $act \in Act$ the set $Enabled[s, act]$ is closed in $Prob[Sta]$. As a consequence, the class of all Souslin STSs covers image-finite NLMPs with countable action spaces. ┘

It is easy to see that image-finite NLMPs cover LMPs (see also Proposition 4.2 in [Wol12]). We refer to Chapter 5 in [Wol12] where modelling formalisms from the literature with an

NLMP semantics are presented. This discussion in particular includes stochastic automata [D'A99, BD04, DK05], probabilistic guarded command language [MM04], and a class of stochastic hybrid systems [FHH$^+$11, Hah13].

**Proposition 179.** *Let $\mathcal{T} = (Sta, Act, \rightarrow)$ be an image-finite NLMP such that the set Act is countable. For every states $s_a, s_b \in Sta$ the following four statements hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\mathrm{sou}} s_b$ *iff* $s_a \preceq_{\bullet} s_b$ *iff* $s_a \preceq^{\exists} s_b$.

(2) $s_a \preceq s_b$ *implies* $s_a \leq^{\mathrm{tr}} s_b$.

(3) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$ *iff* $s_a \simeq_{\circ} s_b$ *iff* $s_a \simeq^{*} s_b$.

(4) $s_a \simeq s_b$ *implies* $s_a =^{\mathrm{tr}} s_b$.

*Moreover, the relation $\preceq$ is weakly smooth and accordingly, the relation $\simeq$ is smooth.*

*Proof.* By Remark 178, the proposition is a summary of Theorem 128 and Corollary 131. □

For the same reasons as in Section 7.3 the bisimulation notion of this thesis and the corresponding one in the NLMP literature are the same for every image-finite NLMP where the action space is countable. This also implies that bisimulation equivalence and the equivalence induced by event bisimulation [DTW12, Wol12] are the same for the mentioned subclass of NLMPs. The result that APCTL$_\circ$ characterises the bisimulation equivalence can be already found in [DTW12, Wol12] (see also our discussions in Section 7.3). However, the remaining implications in Proposition 179 are new. In particular, whereas existing NLMP literature focus on bisimulation relations only, Proposition 179 also includes corresponding results for the simulation preorder. Moreover, we identified an important class of STSs where the Souslin-simulation preorder and the simulation preorder are the same and accordingly, for the corresponding relations induced by Souslin bisimulation and bisimulation.

The second condition on an NLMP in Definition 177 is closed connected to our property on STS of being Borel concerning the hit sigma algebra. However, whereas condition (2) in Definition 177 takes all the Borel subsets of *Prob*[*Sta*] into account, the corresponding condition in this thesis regards only the open subsets of *Prob*[*Sta*] (see also Theorem 129). This seemingly small difference has serious consequences (see also the discussion in Section 18.1 in [AB06] concerning different measurability notions for set-valued functions). First of all, the whole theory in [DTW12, Wol12] heavily relies on the fact that condition (2) in Definition 177 indeed includes all the Borel subsets of *Prob*[*Sta*]. For instance, this is already important for the definition of the semantics of the modal logics in [DTW12, Wol12]. In this sense, we completely rework this existing theory with the weakened assumption (see Theorem 128). Moreover, whereas NLMPs are not closed under composition, a corresponding closure result holds for STSs being Borel concerning the hit sigma algebra (see Theorem 160). The latter is illustrated by the following example:

**Example 180.** Consider the simple STSs $\mathcal{T}_1 = (\mathit{Sta}, \mathit{Act}, \rightarrow_1)$ and $\mathcal{T}_2 = (\mathit{Sta}, \mathit{Act}, \rightarrow_2)$ where $\mathit{Sta} = [0,1]$ and the set $\mathit{Act}$ is a singleton, say $\mathit{Act} = \{\mathit{act}\}$ for some symbol $\mathit{act}$, and the transition relations $\rightarrow_1$ and $\rightarrow_2$ are given as follows:

$$\rightarrow_1 = \{\langle s, \mathit{Dirac}[\mathit{act}] \otimes \mathit{Dirac}[s]\rangle \ ; \ s \in \mathit{Sta}\},$$
$$\rightarrow_2 = \{\langle s, \mathit{Dirac}[\mathit{act}] \otimes \mathit{Dirac}[s']\rangle \ ; \ s, s' \in \mathit{Sta}\}.$$

It is easy to see that $\mathcal{T}_1$ and $\mathcal{T}_2$ are NLMPs. According to Remark 178, the STSs $\mathcal{T}_1$ and $\mathcal{T}_2$ are hence Borel concerning the hit sigma. Define the STS

$$\mathcal{T} = \mathcal{T}_1 \parallel_{\mathit{CartSp},\mathit{Act}} \mathcal{T}_2.$$

Thanks to Theorem 160, we have that the STS $\mathcal{T}$ is also Borel concerning the hit sigma algebra. Indeed, for every $s \in \mathit{Sta}$ the set $\mathit{Enabled}[\mathcal{T}_1, s, \mathit{act}]$ is a singleton and hence compact and moreover, the set $\mathit{Enabled}[\mathcal{T}_2, s, \mathit{act}]$ is equal to $\mathit{Prob}[\mathit{Sta}]$ and thus compact (see Theorem 17.22 in [Kec95]).

Although $\mathcal{T}$ is Borel concerning the hit sigma algebra, the STS is no NLMP. Let us provide an argument for this claim. According to Remarks 10 (1) and (2), there exists a Borel set $B \subseteq \mathit{Sta} \times \mathit{Sta}$ such that the set $M$ is not Borel in $\mathit{Sta}$ where

$$M = \{r \in \mathbb{R} \ ; \ \text{there exists } r' \in \mathbb{R} \text{ with } \langle r, r'\rangle \in B\}.$$

Defining the set

$$\mathit{Dirac}[B] = \{\mathit{Dirac}[r_1] \otimes \mathit{Dirac}[r_2] \ ; \ \langle r_1, r_2\rangle \in B\},$$

we have that $\mathit{Dirac}[B]$ is Borel in $\mathit{Prob}[\mathit{Sta} \times \mathit{Sta}]$ (see, e.g., Corollary 4.5 in [Wol12]). However, the following set is not Borel in $\mathit{Sta} \times \mathit{Sta}$,

$$\{\langle s_1, s_2\rangle \in \mathit{Sta} \times \mathit{Sta} \ ; \ \mathit{Enabled}[\mathcal{T}, \langle s_1, s_2\rangle, \mathit{act}] \cap \mathit{Dirac}[B] \neq \varnothing\} = M \times \mathit{Sta}.$$

As $M \times \mathit{Sta}$ is not Borel in $\mathit{Sta} \times \mathit{Sta}$, it follows that $\mathcal{T}$ is no NLMP. Our example involves no stochastic and hence, it also applies for the standard composition operator. ⌟

## 7.5 Controlled Markov processes

This section focuses on controlled Markov processes, in particular, it is shown that every controlled Markov process induces a Souslin STS. We moreover provide a subclass of controlled Markov processes where the Souslin-simulation preorder and the simulation preorder as well as the Souslin-bisimulation equivalence and the bisimulation equivalence are the same, respectively. As a consequence, our results of Chapter 4 concerning trace-distribution relations are applicable for the model under consideration. The definition below is borrowed from [TMKA16].

**Definition 181.** A *controlled Markov process* (*CMP*) is given by a tuple

$$\mathcal{M} = (Sta, Act, U, Ctr, Adm)$$

comprising of the following components:

(1) $Sta$ is a Polish space (called *state space*).

(2) $Act$ is a Polish space (called *action space*).

(3) $U$ is a Polish space (called *input space*).

(4) $Ctr\colon Sta \times U \to Prob[Act \times Sta]$ is a Borel function (called *control law*).

(5) $Adm\colon Sta \rightsquigarrow U$ is a set-valued function whose induced relation $Rel[Adm]$ is a Souslin set in $Sta \times U$.

⌟

Roughly speaking, the space $U$ consists of all possible inputs controlled by an external agent, the function $Ctr$ formalises the control law of the CMP, and the set-valued function $Adm$ represents the set of all admissible controls in a state, i.e., an input $u \in U$ is *admissible* in a state $s \in Sta$ provided $u \in Adm(s)$. In the context of stochastic control problems, this section focuses solely on modelling purposes of the operational behaviour of the underlying stochastic systems. Hence, the introduced formal model for an CMP includes neither reward functions nor cost functionals to be optimised. However, the formalism can be extended in this direction in a natural way (see also Section 5.8). We introduce the following subclass of CMPs satisfying additional topological properties for the involved spaces and functions.

**Definition 182.** We call a CMP $(Sta, Act, U, Ctr, Adm)$ *continuous* provided the following four conditions are fulfilled:

(1) The Polish space $U$ is a compact.

(2) The function $Ctr$ is continuous.

(3) The set $Rel[Adm]$ is closed in $Sta \times U$.

(4) The set-valued function $Adm$ is lower hemicontinuous, i.e., for every open set $O \subseteq U$ the set $\{s \in Sta \,;\, Adm(s) \cap O \neq \varnothing\}$ is open in $Sta$.

⌟

Whereas requirements (1), (2), and (3) are adopted from Definition 1 of [TMKA16], requirement (4) yields an additional constraint for the formalism studied in [TMKA16]. In fact, lower hemicontinuity is a condition that appears natural in the study of constrained maximisation problems represented by set-valued functions (see, e.g., Sections 17.2 and 17.5 in [AB06]). In examples where the set-valued function $Adm$ models global constraints for control decisions, the lower-hemicontinuity condition is harmless:

**Example 183.** Consider a CMP $\mathcal{M} = (Sta, Act, U, Ctr, Adm)$ such that the function $Ctr$ is continuous and where the input space is given by

$$U = [0, 1] \times [0, 1],$$

Consider a continuous function $f\colon Sta \to [0, 1]$ such that for every state $s \in Sta$ it holds

$$Adm(s) = \{\langle u_1, u_2\rangle \in U \,;\, u_1 + u_2 = f(s)\},$$

i.e., the function $Adm$ may represent possible weightings for the share of resources such as energy, memory, or bandwidth depending on the actual grading formalised by $f$. In many situations the function $f$ is even constant with $f(s) = 1$ for every $s \in Sta$ (see also the example on a small power network in Section 2.2 of [TMKA16]).

We have that the CMP $\mathcal{M}$ under consideration is continuous. Clearly, the Polish space $U$ is compact. By the continuity of the function $f$, it is moreover easy to see that the set $Rel[Adm]$ is closed in $Sta \times U$. Moreover, it easily follows that the set-valued function $Adm$ is lower hemicontinuous. ⌟

The control law *Ctr* of an CMP depends on the actual state of the system and the given input. Intuitively, assuming the actual state of the system is $s$ and an input $u$ admissible in $s$, the successor state of the system is determined by the probability measure $Ctr(s,u)$. More precisely, the semantics of a CMP is defined as follows:

**Definition 184.** Let $\mathcal{M} = (Sta, Act, U, Ctr, Adm)$ be a CMP. The *semantics of* $\mathcal{M}$ is given by the STS

$$[\![\mathcal{M}]\!] = (Sta, Act, \rightarrow)$$

such that for all $s \in Sta$ and $\varphi \in Prob[Act \times Sta]$ the following equivalence holds:

$$s \rightarrow \varphi \quad \text{iff} \quad \text{there exists } u \in U \text{ such that } u \in Adm(s) \text{ and } \varphi = Ctr(s,u).$$

⌟

The introduced STS may involve uncountable non-determinism as every admissible control in a state induces a distribution over action-state pairs. The actions can be seen as a labelling of the corresponding transitions. If all the inputs in $U$ are observable, one may consider CMPs $(Sta, Act, U, Ctr, Adm)$ where $Act = U$ and $Ctr(s,u)(\{u\} \times Sta) = 1$ for every state $s \in Sta$ and control $u \in U$.

**Proposition 185.** *Let $\mathcal{M} = (Sta, Act, U, Ctr, Adm)$ be a CMP. Then the STS $[\![\mathcal{M}]\!]$ is Souslin. If the CMP $\mathcal{M}$ is continuous, then for every states $s_a, s_b \in Sta$ the following four statements concerning the simulation preorder and the bisimulation equivalence of the STS $[\![\mathcal{M}]\!]$ hold:*

(1) $s_a \preceq s_b$ *iff* $s_a \preceq^{\mathrm{sou}} s_b$.

(2) $s_a \preceq s_b$ *implies* $s_a \leq^{\mathrm{tr}} s_b$.

(3) $s_a \simeq s_b$ *iff* $s_a \simeq^{\mathrm{sou}} s_b$.

(4) $s_a \simeq s_b$ *implies* $s_a =^{\mathrm{tr}} s_b$.

*Moreover, the sets $\preceq$ and $\simeq$ are closed in $Sta \times Sta$.*

*Proof.* The argument showing that the STS $[\![\mathcal{M}]\!]$ has the Souslin property is straightforward. Nevertheless, let us present the details. Denote the transition relation of the STS $[\![\mathcal{M}]\!]$ by $\rightarrow$. Define the set $M \subseteq Sta \times U \times Prob[Act \times Sta]$ by

$$M = Graph[Ctr] \cap (Graph[Adm] \times Prob[Act \times Sta])$$

As $Ctr$ is a Borel function, the set $Graph[Ctr]$ is Souslin in $Sta \times U \times Prob[Act \times Sta]$ (see Remark 10 (6)). By the definition of CMPs, the set $Graph[Adm]$ is Souslin in $Sta \times U$. It follows that the set $M$ is Souslin in $Sta \times U \times Prob[Act \times Sta]$ (see Remark 10 (4)). Moreover, it holds

$$\to = \{\langle s, \varphi \rangle \in Sta \times Prob[Act \times Sta] \; ; \; \langle s, u, \varphi \rangle \in M \text{ for some } u \in U\}.$$

Applying Remark 10 (5), we conclude that the STS $[\![\mathcal{M}]\!]$ is Souslin.

Assume that the CMP $\mathcal{M}$ is continuous. Let $s_a, s_b \in Sta$ be states and $R$ be a simulation for $[\![\mathcal{M}]\!]$. Denote the topological of the set $R$ in $Sta \times Sta$ by $\bar{R}$. In the remainder of this proof, we show that $\bar{R}$ is a simulation for $[\![\mathcal{M}]\!]$. Indeed, this suffices to conclude the remaining claims of theorem: first of all, if the topological closure of a simulation is also a simulation, it follows that the simulation preorder $\preceq$ is a closed set in $Sta \times Sta$. As every closed set in $Sta \times Sta$ is in particular Souslin in $Sta \times Sta$ (see Remark 10 (2)), we immediately obtain statement (1). This together with Corollary 83 yield statement (2). The corresponding claims for the bisimulation equivalence can be proven analogously.

For the remainder of this proof denote the metric on $U$ inducing the corresponding Polish topology on $U$ by $dist$. For every $u \in U$ and $\varepsilon \in \mathbb{R}_{>0}$ the open ball in $U$ concerning $dist$ centered at $u$ with radius $\varepsilon$ is denoted by $Ball(u, \varepsilon)$, i.e., we have

$$Ball(u, \varepsilon) = \{u' \in U \; ; \; dist(u, u') < \varepsilon\}.$$

Let $\langle s_a, s_b \rangle \in Sta$ and $\varphi_a \in Prob[Act \times Sta]$ be such that $\langle s_a, s_b \rangle \in \bar{R}$ and $s_a \to \varphi_a$. By the definition of $\bar{R}$, there exists a sequence $(\langle s_{a,n}, s_{b,n} \rangle)_{n \in \mathbb{N}}$ of pairs $\langle s_{a,n}, s_{b,n} \rangle \in R$ converging in $Sta \times Sta$ to the limit $\langle s_a, s_b \rangle$. As we have $s_a \to \varphi_a$ the semantics of CMPs yields a control $u_a \in U$ such that $u_a \in Adm(s_a)$ and $\varphi_a = Ctr(s_a, u_a)$.

Clearly, for every $n \in \mathbb{N}$ the set $Ball(u_a, 1/n)$ is an open set in $U$. Note, for every $n \in \mathbb{N}$ we have that $Adm(s_a) \cap Ball(u_a, 1/n) \neq \varnothing$. Using that $Adm$ is upper hemicontinuous, for every $n \in \mathbb{N}$ the set $\{s'_a \in Sta \; ; \; Adm(s'_a) \cap Ball(u_a, 1/n) \neq \varnothing\}$ is open in $Sta$. This insight and the fact that the sequence $(s_{a,n})_{n \in \mathbb{N}}$ converges in $Sta$ to the limit $s_a$ yield the following statement: there is a natural number $N \in \mathbb{N}$ and a sequence $(u_{a,n})_{n \in \mathbb{N}}$ of elements in $U$ such that for every $n \in \mathbb{N}$ with $n \geq N$ it holds $u_{a,n} \in Adm(s_{a,n})$ and $dist(u_{a,n}, u_a) \leq 1/n$. It follows that the sequence $(u_{a,n})_{n \in \mathbb{N}}$ converges in $U$ to the limit $u_a$ and for every $n \in \mathbb{N}$ with $n \geq N$ we moreover have that $s_{a,n} \to Ctr(s_{a,n}, u_{a,n})$.

As $R$ is a simulation, there exists a sequence $(u_{b,n})_{n \in \mathbb{N}}$ in $U$ such that for every $n \in \mathbb{N}$ with $n \geq N$ it holds $s_{b,n} \to Ctr(s_{b,n}, u_{b,n})$ and $\langle Ctr(s_{a,n}, u_{a,n}), Ctr(s_{b,n}, u_{b,n}) \rangle \in (R^{Act})^{\mathrm{wgt}}$. Using that $U$ is compact, there exists a strictly increasing function $f \colon \mathbb{N} \to \mathbb{N}$ such that the sequences $(u_{b,f(n)})_{n\in\mathbb{N}}$ converges in $U$. Let $u_b \in U$ be the limit of the sequence $(u_{b,f(n)})_{n\in\mathbb{N}}$. Since the set $Rel[Adm]$ is closed in $Sta \times U$, it follows $u_b \in Adm(s_b)$. We therefore obtain $s_b \to Ctr(s_b, u_b)$.

It remains to show $\langle Ctr(s_a, u_a), Ctr(s_b, u_b)\rangle \in (\bar{R}^{Act})^{\mathrm{wgt}}$. Since the function $Ctr$ is continuous and as for every $n \in \mathbb{N}$ with $n \geq f(N)$ it holds

$$\langle Ctr(s_{a,f(n)}, u_{a,f(n)}), Ctr(s_{b,f(n)}, u_{b,f(n)})\rangle \in (R^{Act})^{\mathrm{wgt}} \subseteq (\bar{R}^{Act})^{\mathrm{wgt}},$$

it suffices to show that the set $(\bar{R}^{Act})^{\mathrm{wgt}}$ is closed in $Prob[Act \times Sta] \times Prob[Act \times Sta]$. However, as $R^{Act}$ is closed in $(Act \times Sta) \times (Act \times Sta)$, the latter is stated in Lemma 3.9 in [Les10]. The mentioned lemma is a consequence of general results referring to the convergence of probability measures such as Prokhorov's theorem characterising the relatively compact subsets in $Prob[Act \times Sta]$ (see also Theorems 5.1 and 5.2 in [Bil99]) and the Portmanteau theorem (see also Theorem 17.20 in [Kec95]). □

To the best of our knowledge, the established results referring to trace-distribution relations of an CMP are new and extend existing literature.

## 7.6 Stochastic hybrid systems

We introduce stochastic rectangular hybrid automata (SRHAs) where we follow the standard schema for hybrid automata [Hen96]. It turns out that every SRHA admits a Souslin-STS semantics. As SRHAs cover, e.g., probabilistic timed and hybrid automata [Spr01, KNSS02, Spr11, ZSR$^+$12, Spr15] as well as stochastic automata [D'A99, BD04, DK05], this section provides a powerful high-level modelling formalism for Souslin STSs.

Given a finite set $Var$ of variables, the set $Update[Var]$ denotes the set of all Borel functions with domain $Eval[Var]$ and codomain $Prob[Eval[Var]]$. Intuitively, every function in $Update[Var]$ represents an update for variables: given $upd \in Update[Var]$ and a variable evaluation $e$ for $Var$, a new variable evaluation $e'$ for $Var$ is sampled according to the probability measure $upd(e)$.

**Definition 186.** A *stochastic rectangular hybrid automaton* (*SRHA*) is a tuple

$$\mathcal{R} = (Loc, Var, Act, Flow, Inv, Jump)$$

comprising of the following components:

(1) $Loc$ is a countable set (*location space*),

(2) *Var* is a finite set of variables,

(3) *Act* is a countable set (*action space*) disjoint from $\mathbb{R}_{\geq 0}$,

(4) $\mathit{Flow}\colon \mathit{Loc} \to \mathit{Cond}[\mathit{Var}]$ is a function (*flow function*),

(5) $\mathit{Inv}\colon \mathit{Loc} \to \mathit{Cond}[\mathit{Var}]$ is a function (*invariant function*),

(6) *Jump* is a countable relation (*jump relation*) where

$$\mathit{Jump} \subseteq \mathit{Loc} \times \mathit{Cond}[\mathit{Var}] \times \mathit{Act} \times \mathit{Update}[\mathit{Var}] \times \mathit{Prob}[\mathit{Loc}].$$

Furthermore, for every $l \in \mathit{Loc}$ the sets $\mathit{Flow}(l)$ and $\mathit{Inv}(l)$ are required to be convex. Here, a condition $\mathit{cond} \in \mathit{Cond}[\mathit{Var}]$ is convex if for every $r \in [0,1]$ and $e, e' \in \mathit{Eval}[\mathit{Var}]$ such that $e \models \mathit{cond}$ and $e' \models \mathit{cond}$ it holds $e''_r \models \mathit{cond}$ where the variable evaluation $e''_r \in \mathit{Eval}[\mathit{Var}]$ is defined by $e''(v) = r \cdot e(v) + (1-r) \cdot e'(v)$ for every $v \in \mathit{Var}$. ⌟

Let $\mathcal{R}$ be an SRHA as before. Intuitively, when times passes in a location $l$, then the variables evolves according to a differentiable function whose derivative satisfies the condition $\mathit{Flow}(l)$ for every point in the considered time span. Time can advance in a location $l$ as long as the evaluations of the variables satisfy the invariant $\mathit{Inv}(l)$. Assuming the current location and variable evaluation of the SRHA are given by $l$ and $e$, respectively, a jump $\langle l, \mathit{guard}, \mathit{act}, \mathit{upd}, \lambda\rangle$ is enabled if $e \models \mathit{guard}$. Here, when invoking this jump, the action *act* is executed, the variable evaluation is updated according to the function *upd*, and the successor location is sampled according to $\lambda$. The formal semantics is as follows:

**Definition 187.** Let $\mathcal{R} = (\mathit{Loc}, \mathit{Var}, \mathit{Act}, \mathit{Flow}, \mathit{Inv}, \mathit{Jump})$ be an SRHA. The *semantics of* $\mathcal{R}$ is given by the simple STS

$$[\![\mathcal{R}]\!] = (\mathit{Loc} \times \mathit{Eval}[\mathit{Var}], \mathbb{R}_{\geq 0} \cup \mathit{Act}, \to)$$

where the transition relation $\to$ is given as follows: for every location $l \in \mathit{Loc}$, variable evaluation $e \in \mathit{Eval}[\mathit{Var}]$, action $\mathit{act}' \in \mathbb{R}_{\geq 0} \cup \mathit{Act}$, and probability measure $\mu \in \mathit{Prob}[\mathit{Loc} \times \mathit{Eval}[\mathit{Var}]]$ we have

$$\langle l, e\rangle \to \langle \mathit{act}', \mu\rangle$$

precisely when one of the two conditions below is satisfied:

(1) There are $t \in \mathbb{R}_{\geq 0}$ and a function $f\colon [0,t] \to \mathit{Eval}[\mathit{Var}]$ such that $\mathit{act}' = t$, $f(0) = e$, $\mu = \mathit{Dirac}[\langle l, f(t)\rangle]$, $f(t') \models \mathit{Inv}(l)$ for all $t' \in [0,t]$ and moreover, $f$ is differentiable on $(0,t)$ with $\mathit{dot}[f](t') \models \mathit{Flow}(l)$ for all $t' \in (0,t)$.

(2) There are $act \in Act$, $guard \in Cond[Var]$, $\lambda \in Prob[Loc]$, $upd \in Update[Var]$, and $\eta \in Prob[Eval[Var]]$ such that $act' = act$, $e \models guard$, $\eta \in upd(e)$, $\mu = \lambda \otimes \eta$, and moreover, $\langle l, guard, act, upd, \lambda \rangle \in Jump$.

⌟

Inspecting the definition of the transition relation $\rightarrow$ again, it is possible that a jump of the SRHA leads to a state $\langle l, e \rangle$ with $e \not\models Inv(l)$. However, no time can pass in such a state. Regarding this issue, we also refer to the notion of weak-invariant semantics, also used and discussed in [HHHK12]. Simple syntactical restrictions for an given SRHA ensure that the underlying STS always enters a state $\langle l, e \rangle$ satisfying $e \models Inv(l)$. For instance, one may require that for every jump $\langle l, guard, act, upd, \lambda \rangle \in Jump$ and $e \in Eval[Var]$ such that $e \models Inv(l) \wedge guard$ the following condition holds: for every location $l' \in Loc$,

$$\lambda(\{l'\}) > 0 \quad \text{implies} \quad \eta(Inv(l')) = 1.$$

**Proposition 188.** *For every SRHA $\mathcal{R}$ the STS $[\![\mathcal{R}]\!]$ is Souslin.*

*Proof.* Let $\mathcal{R} = (Loc, Var, Act, Flow, Inv, Jump)$ be an SRHA. Denote the transition relation of the STS $[\![\mathcal{R}]\!]$ by $\rightarrow$ and moreover, define

$$Sta = Loc \times Eval[Var] \quad \text{and} \quad Act' = \mathbb{R}_{\geq 0} \cup Act.$$

We additionally introduce the sets

$$\rightarrow_{\mathbb{R}_{\geq 0}} = \{\langle s, t, \mu \rangle \in Sta \times \mathbb{R}_{\geq 0} \times Prob[Sta] \; ; s \rightarrow \langle t, \mu \rangle\},$$
$$\rightarrow_{Act} = \{\langle s, act, \mu \rangle \in Sta \times Act \times Prob[Sta] \; ; s \rightarrow \langle act, \mu \rangle\}.$$

Our task is to show that the STS $[\![\mathcal{R}]\!]$ is Souslin. According to Remark 68, it suffices to show that the sets $\rightarrow_{\mathbb{R}_{\geq 0}}$ and $\rightarrow_{Act}$ are Souslin in $Sta \times \mathbb{R}_{\geq 0} \times Prob[Sta]$ and $Sta \times Act \times Prob[Sta]$, respectively.

*The set $\rightarrow_{\mathbb{R}_{\geq 0}}$ is Souslin.* We show that the set $\rightarrow_{\mathbb{R}_{\geq 0}}$ is Souslin in $Sta \times \mathbb{R}_{\geq 0} \times Prob[Sta]$. Define the Polish space

$$X_{\mathbb{R}_{\geq 0}} = Eval[Var] \times Sta \times \mathbb{R}_{\geq 0} \times Prob[Sta]$$

The following argument relies on the mean value theorem well-known from undergraduate analysis courses. Introduce the function $g\colon Eval[Var] \times Sta \times \mathbb{R}_{\geq 0} \rightarrow Prob[Sta]$,

$$g(slope, \langle l, e \rangle, t) = Dirac[\langle l, e + t \cdot slope \rangle].$$

Relying on Example 6, it is easy to see that $g$ is a Borel function. Thus, the set $Graph[g]$ is Borel in $X_{\mathbb{R}_{\geq 0}}$ by Remark 10 (6). As the set $Loc$ is countable, it is easy to see that $B$ is Borel in $X_{\mathbb{R}_{\geq 0}}$ where

$$B = \bigcup_{l \in Loc} Flow(l) \times \big(\{l\} \times Inv(l)\big) \times \mathbb{R}_{\geq 0} \times \{Dirac[\langle l, e' \rangle] \; ; e' \models Inv(l)\}.$$

As the sets $Flow(l)$ and $Inv(l)$ are convex for every $l \in Loc$, the mean value theorem yields the following equivalence for every $s \in Sta$, $t \in \mathbb{R}_{\geq 0}$, and $\mu \in Prob[Sta]$:

$$\langle s, t, \mu \rangle \in \rightarrow_{\mathbb{R}_{\geq 0}} \quad \text{iff} \quad \begin{array}{l} \text{there exists } slope \in Eval[Var] \\ \text{such that } \langle slope, s, t, \mu \rangle \in Graph[g] \cap B. \end{array}$$

Finally, applying Remark 10 (1), the set $\rightarrow_{\mathbb{R}_{\geq 0}}$ is Souslin in $Sta \times \mathbb{R}_{\geq 0} \times Prob[Sta]$.

*The set $\rightarrow_{Act}$ is Souslin.* It remains to show that $\rightarrow_{Act}$ is Souslin in $Sta \times Act \times Prob[Sta]$. Define the Polish space

$$X_{Act} = Loc \times Prob[Loc] \times Eval[Var] \times Prob[Eval[Var]] \times Act.$$

Every $jump \in Jump$ induces a Souslin subset $M_{jump}$ of $X_{Act}$ as follows:

$$M_{jump} = \{l\} \times \{\lambda\} \times (Graph[upd] \cap (guard \times Prob[Eval[Var]])) \times \{act\}$$

where $jump = \langle l, guard, act, upd, \lambda \rangle$. By Remark 10 (6), the set $M_{jump}$ is indeed Souslin in $X_{Act}$. Define $M \subseteq X_{Act}$ by

$$M = \bigcup_{jump \in Jump} M_{jump}$$

Since every SRHA includes only countably many jumps, the set $M$ is Souslin in $X_{Act}$ by Remark 10 (4). Introduce the function $h \colon X_{Act} \to Sta \times Act \times Prob[Sta]$,

$$h(l, \lambda, e, \eta, act) = \langle \langle l, e \rangle, act, \lambda \otimes \eta \rangle.$$

The function $h$ is Borel by Example 6. According to Remark 10 (5) and as we have

$$\rightarrow_{Act} = h(M),$$

we finally obtain that the set $\rightarrow_{Act}$ is Souslin in $Sta \times Act \times Prob[Sta]$. $\square$

# 8 Conclusions

In this thesis we have investigated STSs including both an uncountable state space and an uncountable action space (see Chapter 3). Using measurable-selection principles and techniques from descriptive set theory, we have shown that under a Souslin side constraint the simulation preorder is subsumed by the trace-distribution preorder and the bisimulation equivalence is finer that the trace-distribution equivalence (see Chapter 4). This, together with the presented logical characterisation of simulation and bisimulation concerning a weak and a strong modal logic, provide a complete picture on behavioural relations for a large subclass of STSs (see Chapter 5). In order to obtain a logical characterisation covering STSs with uncountable action spaces, the key idea is the new concept of action event families specifying the basic atomic building blocks of the introduced logics.

Furthermore, we have introduced a new parallel-composition operator for STSs appropriate to model stochastic dependencies between components declarative fashion relying on the notion of couplings (see Section 6). Our framework is convenient in situations where one wants to abstract from vaguely or unknown operational behaviour causing specific stochastic dependencies. We have proved that simulation preorder and bisimulation equivalence are congruences with respect to our newly developed parallel operator where the challenge is to construct specific (span) couplings. For this, we have heavily exploited the disintegration theorem.

Our research covers many prominent modelling formalisms from the literature and extend existing contributions for those (see Chapter 7), in particular, for labelled Markov processes, discrete-time stochastic hybrid automata, and controlled Markov processes. Our results yield the basis of many possible directions for further research that are briefly summarised next.

It would be interesting to know whether one can weaken or even drop the (Souslin) assumptions for the main results in Chapters 4 and 5. Regarding this aspect, the following question arises: given a Souslin STS $\mathcal{T}$ and states $s_a$ and $s_b$ with $s_a \simeq s_b$, does it hold $s_a \simeq^{\mathrm{sou}} s_b$, i.e., does there exist a Souslin bisimulation $R$ for $\mathcal{T}$ with $\langle s_a, s_b \rangle \in R$? An analogous question is relevant for simulations. By the main result in Chapter 5, we have already identified a large subclass of Souslin STSs for which the latter question can be answered positively. An interesting question is whether results from Chapter 5 can be in particular extended for a subclass of STSs that goes beyond image-finite systems. Here, one may focus on, e.g., image-compact STSs in a first step. In this context, the article [Ter15] is also interesting where a not image-finite NLMP is presented whose bisimulation equivalence forms a Souslin set, however, no Borel set.

This thesis focused on exact notions for simulation and bisimulation. However, it would be interesting to extend our results for weak simulation and bisimulation, approximation notions thereof, as well as behavioural metrics. Regarding this issue, there already exists many literature focusing on subclasses of STSs. For instance, the recent work [BA17] shows that for approximate bisimilar states of a labelled Markov chain the induced finite-trace distributions are close with respect to the total variation distance (see also [HAV16]).

Besides approximate notions, there are also metrics for stochastic models to escape from the fragility of exact equivalences. For instance, the contributions for labelled Markov chains [CvW12], for concurrent labelled Markov chains [CGPX14], LMPs [Des99, DGJP03, Pan09], generalised semi-Markov processes [GJP04], as well as the abstract coalgebraic setting [BBKK14] yield an excellent starting point for the study of behavioural metrics for (Souslin) STSs (see also [DD09] for a survey).

Weak equivalences abstract from internal steps that are supposed to be not observable from an external agent, e.g., [SL94, BH97, PLS00, DGJP10, Cat05, CSKN05, DH13b]. Weak notions for simulations and bisimulations have been for instance already discussed in the context of trace-distribution relations of STSs [Cat05] (see also Section 4.7) as well as concerning a logical characterisation of the bisimulation equivalence of labelled concurrent Markov chains [DGJP10].

Our congruence results for the span-coupling composition operator focused on the relations induced by simulation and bisimulation. It is well known that trace-distribution equivalence is not a congruence with respect to (standard) composition for subclasses of STSs [Seg95, LSV07]. Regarding this issue, it seems plausible to combine ideas of the latter cited articles with the developed techniques in this thesis to extend the results concerning probabilistic contexts in [LSV07] for (Souslin) STSs.

In statistics there is the concept of copulas [Nel06], i.e., specific functions convenient for the specification of dependencies between real-valued random variables. Adapting and extending (statistical) techniques referring the analysis and synthesis of copulas to STSs and our developed compositional framework would be another interesting idea left as future work. As a first simple starting point, one may consider coupling-constraint functions that are induced by appropriate (parametric) families of copulas.

The presented compositional framework for STSs in this thesis yields a basis for our composition operator presented at [GBK16] covering a generic model for stochastic hybrid systems called stochastic hybrid motion automata. In this new model the progressing flow is recorded within states and the adaption of flows depends on executing commands rather than happening on arbitrary occasions. As a next step, one may extend this class of hybrid systems by allowing for stochastic flows, i.e., flows where stochastic choices can be made continuously over time (see also [BL04, BLB05, BL06]). The presented algebraic theory on behavioural relations then needs to be extended for this expressive class of stochastic hybrid systems.

# 9 Bibliography

[AAGT15] Manindra Agrawal, S. Akshay, Blaise Genest, and P. S. Thiagarajan. Approximate Verification of the Symbolic Dynamics of Markov Chains. *Journal of the ACM*, 62(1):2:1–2:34, 2015.

[AB06] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, 3 edition, 2006.

[ABC10] Alessandro Aldini, Marco Bernardo, and Flavio Corradini. *A Process Algebraic Approach to Software Architecture Design*. Springer, 2010.

[ACH+95] 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.

[AD94] Rajeev Alur and David L. Dill. A Theory of Timed Automata. *Theoretical Computer Science*, 126:183–235, 1994.

[AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser, 2005.

[AH97] 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.

[AH99] Rajeev Alur and Thomas A. Henzinger. Reactive Modules. *Formal Methods in System Design*, 15(1):7–48, 1999.

[AKLP10] Alessandro Abate, Joost-Pieter Katoen, John Lygeros, and Maria Prandini. Approximate Model Checking of Stochastic Hybrid Systems. *European Journal of Control*, 16(6):624–641, 2010.

[AKM11] Alessandro Abate, Joost-Pieter Katoen, and Alexandru Mereacre. Quantitative Automata Model Checking of Autonomous Stochastic Hybrid Systems. In *14th International Conference on Hybrid Systems: Computation and Control (HSCC)*, pages 83–92. ACM, 2011.

[APLS08] Alessandro Abate, Maria Prandini, John Lygeros, and Shankar Sastry. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11):2724–2734, 2008.

[Arv98] William Arveson. *An Invitation to C*-Algebras*. Graduate Texts in Mathematics. Springer, 1998.

[ASB+95] Adnan Aziz, Vigyan Singhal, Felice Balarin, Robert K. Brayton, and Alberto L. Sangiovanni-Vincentelli. It usually works: The temporal logic of stochastic systems. In *7th International Conference on Computer Aided Verification (CAV)*, pages 155–165. Springer, 1995.

[ASSB00] Adnan Aziz, Kumud Sanwal, Vigyan Singhal, and Robert Brayton. Verifying continuous time Markov chains. *ACM Transactions on Computational Logic*, 1(1):162–170, 2000.

[BA17] Gaoang Bian and Alessandro Abate. On the Relationship between Bisimulation and Trace Equivalence in an Approximate Probabilistic Context. In *Proceedings of the 20th International Conference on Foundations of Software Science and Computational Structures (FOSSACS)*, Lecture Notes in Computer Science. Springer, 2017. To appear.

[Bae05] J.C.M. Baeten. A brief history of process algebra. *Theoretical Computer Science*, 335(2–3):131–146, 2005.

[BBB+14] Nathalie Bertrand, Patricia Bouyer, Thomas Brihaye, Quentin Menet, Christel Baier, Marcus Größer, and Marcin Jurdziński. Stochastic Timed Automata. *Logical Methods in Computer Science*, 10(4:6), 2014.

[BBC06] Patricia Bouyer, Thomas Brihaye, and Fabrice Chevalier. Control in o-minimal Hybrid Systems. In *Proceedings of the 21st Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 367–378, 2006.

[BBCMar] 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).

[BBKK14] Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Behavioral Metrics via Functor Lifting. In *Proceedings of the 34th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS)*, volume 29 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 403–415. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014.

[BBLM14] 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.

[BBM98] Michael S. Branicky, Vivek S. Borkar, and Sanjoy K. Mitter. A unified framework for hybrid control: model and optimal control theory. *IEEE Transactions on Automatic Control*, 43(1):31–45, 1998.

[BCHK14] Udi Boker, Krishnendu Chatterjee, Thomas A. Henzinger, and Orna Kupferman. Temporal Specifications with Accumulative Values. *ACM Transactions on Computational Logic*, 15(4):27:1–27:25, 2014.

[Bd95] Andrea Bianco and Luca de Alfaro. Model checking of probabilistic and nondeterministic systems. In *15th International Conference on Foundation of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 499–513. Springer, 1995.

[BD04] 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.

[BDEP97] Richard Blute, Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for labelled Markov processes. In *Proceedings of the 12th Annual IEEE Symposium on Logic in Computer Science (LICS)*, pages 149–158, 1997.

[BDHK06] 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.

[BEG+15] Gilles Barthe, Thomas Espitau, Benjamin Grégoire, Justin Hsu, Léo Stefanesco, and Pierre-Yves Strub. Relational Reasoning via Probabilistic Coupling. In *Proceedings of the 20th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR)*, pages 387–401. Springer, 2015.

[BG02] Mario Bravetti and Roberto Gorrieri. The theory of interactive generalized semi-Markov processes. *Theoretical Computer Science*, 282(1):5–32, 2002.

[BGG+16a] Gilles Barthe, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. A Program Logic for Union Bounds. In *Proceedings of the 43rd International Colloquium on Automata, Languages and Programming (ICALP)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 107:1–107:15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016.

[BGG+16b] Gilles Barthe, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. Proving Differential Privacy via Probabilistic Couplings. In *Proceedings of the 31th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 749–758. ACM, 2016.

[BGHS17] Gilles Barthe, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. Coupling Proofs Are Probabilistic Product Programs. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL)*, pages 161–174. ACM, 2017.

[BH97] Chistel Baier and Holger Hermanns. Weak bisimulation for fully probabilistic processes. In *Proceedings of the 9th International Conference on Computer Aided Verification (CAV)*, pages 119–130. Springer, 1997.

[BHHK03] Christel Baier, Boudewijn Haverkort, Holger Hermanns, and Joost-Pieter Katoen. Model-checking algorithms for continuous-time Markov chains. *IEEE Transactions on Software Engineering*, 29(7), 2003.

[Bil95] Patrick Billingsley. *Probability and Measure*. Wiley-Interscience, 3 edition, 1995.

[Bil99] Patrick Billingsley. *Convergence of Probability Measures*. Wiley-Interscience, 2 edition, 1999.

[BK08] Christel Baier and Joost-Pieter Katoen. *Principles of Model Checking*. MIT Press, 2008.

[BKKW14] Christel Baier, Joachim Klein, Sascha Klüppelholz, and Sascha Wunderlich. Weight Monitoring with Linear Temporal Logic: Complexity and Decidability. In *Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pages 11:1–11:10. ACM, 2014.

[BL04] Manuela L. Bujorianu and John Lygeros. General stochastic hybrid systems: Modelling and optimal control. In *Proceedings of the 43rd International Conference on Decision and Control (IEEE CDC)*, pages 1872–1877, 2004.

[BL06] 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.

[BLB05] Manuela L. Bujorianu, John Lygeros, and Marius C. Bujorianu. Bisimulation for General Stochastic Hybrid Systems. In *8th International Conference on Hybrid Systems: Computation and Control (HSCC)*, pages 198–214. Springer, 2005.

[BM05] Thomas Brihaye and Christian Michaux. On the expressiveness and decidability of o-minimal hybrid systems. *Journal of Complexity*, 21(4):447–478, 2005.

[BNL13] Marco Bernardo, Rocco De Nicola, and Michele Loreti. A uniform framework for modeling nondeterministic, probabilistic, stochastic, or mixed processes and their behavioral equivalences. *Information and Computation*, 225:29–82, 2013.

[Bog07] Vladimir Igorevich Bogachev. *Measure Theory Volume*, volume 1 and 2. Springer, 2007.

[Bra04] 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.

[BS96] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Athena scientific optimization and computation series. Athena Scientific, 1996.

[BSA04] Mikhail Bernadsky, Raman Sharykin, and Rajeev Alur. Structured Modeling of Concurrent Stochastic Hybrid Systems. In *Joint International Conferences on Formal Modeling and Analysis of Timed Systems (FORMATS) and on Formal Techniques in Real-Time and Fault-Tolerant Systems (FTRTFT)*, pages 309–324. Springer, 2004.

[Cat05] Stefano Cattani. *Trace-based Process Algebras for Real-Time Probabilistic Systems*. PhD thesis, University of Birmingham, 2005.

[CGPX14] Konstantinos Chatzikokolakis, Daniel Gebler, Catuscia Palamidessi, and Lili Xu. Generalized Bisimulation Metrics. In *Proceedings of the 25th International Conference on Concurrency Theory (CONCUR)*, pages 32–46. Springer, 2014.

[Clo06] Lucia Cloth. *Model Checking Algorithms for Markov Reward Models*. PhD thesis, University of Twente, 2006.

[CMBC93] Giovanni Chiola, Marco Ajmone Marsan, Gianfranco Balbo, and Gianni Conte. Generalized stochastic Petri nets: a definition at the net level and its implications. *IEEE Transactions on Software Engineering*, 19(2):89–107, 1993.

[CSKN05] Stefano Cattani, Roberto Segala, Marta Z. Kwiatkowska, and Gethin Norman. Stochastic Transition Systems for Continuous State Spaces and Nondeterminism. 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.

[CvW12] Di Chen, Franck van Breugel, and James Worrell. On the Complexity of Computing Probabilistic Bisimilarity. In *Proceedings of the 15th International Conference on Foundations of Software Science and Computational Structures (FOSSACS)*, pages 437–451. Springer, 2012.

[D'A99] Pedro R. D'Argenio. *Algebras and Automata for Timed and Stochastic Systems*. PhD thesis, University of Twente, 1999.

[dAH01] Luca de Alfaro and Thomas A. Henzinger. Interface Automata. *SIGSOFT Software Engineering Notes*, 26(5):109–120, 2001.

[DD09] Yuxin Deng and Wenjie Du. The Kantorovich Metric in Computer Science: A Brief Survey. *Electronic Notes in Theoretical Computer Science*, 253(3):73–82, 2009.

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

[Den15] Yuxin Deng. *Semantics of Probabilistic Processes: An Operational Approach*. Springer-Verlag and Shanghai Jiao Tong University Press, 2015.

[DEP02] Josée Desharnais, Abbas Edalat, and Prakash Panangaden. Bisimulation for Labelled Markov Processes. *Information and Computation*, 179(2):163–193, 2002.

[Des99] Josée Desharnais. *Labelled Markov processes*. PhD thesis, McGill University, Montréal, 1999.

[DGJP03] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximating labelled Markov processes. *Information and Computation*, 184(1):160–200, 2003.

[DGJP10] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Weak bisimulation is sound and complete for pCTL*. *Information and Computation*, 208(2):203–219, 2010.

[DH13a] Yuxin Deng and Matthew Hennessy. Compositional reasoning for weighted Markov decision processes. *Science of Computer Programming*, 78(12):2537–2579, 2013.

[DH13b] Yuxin Deng and Matthew Hennessy. On the semantics of Markov automata. *Information and Computation*, 222:139–168, 2013.

[DHK99] Pedro R. D'Argenio, Holger Hermanns, and Joost-Pieter Katoen. On Generative Parallel Composition. *Electronic Notes in Theoretical Computer Science*, 22:30–54, 1999. PROBMIV'98, First International Workshop on Probabilistic Methods in Verification.

[DHR08] Laurent Doyen, Thomas A. Henzinger, and Jean-Francois Raskin. Equivalence of Labeled Markov Chains. *International Journal of Foundations of Computer Science*, 19(03):549–563, 2008.

[DIY79] Evgenii Borisovich Dynkin, Aleksandr Adolfovich Iushkevich, and Alexander Adolph Yushkevich. *Controlled Markov processes*. Grundlehren der mathematischen Wissenschaften. Springer, 1979.

[DK05] Pedro R. D'Argenio and Joost-Pieter Katoen. A theory of stochastic systems part I: Stochastic automata and part II: Process algebra. *Information and Computation*, 203(1):1–74, 2005.

[DLM16] Pedro R. D'Argenio, Matias David Lee, and Raúl E. Monti. Input/Output Stochastic Automata. In *Proceedings of the 14th International Conference on Formal Modeling and Analysis of Timed Systems (FORMATS)*, pages 53–68. Springer, 2016.

[DM88] Claude Dellacherie and Paul André Meyer. *Probabilities and potential C: potential theory for discrete and continuous semigroups*. North-Holland Mathematics Studies. Elsevier, 1988.

[DMS14] Laurent Doyen, Thierry Massart, and Mahsa Shirmohammadi. *Limit Synchronization in Markov Decision Processes*, pages 58–72. Springer, 2014.

[Dob07] Ernst-Erich Doberkat. *Stochastic Relations: Foundations for Markov Transition Systems*. Chapman & Hall/CRC Studies in Informatics Series. CRC Press, 2007.

[DP03] Josée Desharnais and Prakash Panangaden. Continuous stochastic logic characterizes bisimulation of continuous-time Markov processes. *The Journal of Logic and Algebraic Programming*, 56(1):99–115, 2003.

[DTW12] Pedro R. D'Argenio, Pedro Sánchez Terraf, and Nicolás Wolovick. Bisimulations for non-deterministic labelled Markov processes. *Mathematical Structures in Computer Science*, 22:43–68, 2012.

[DV90] Rocco De Nicola and Frits Vaandrager. Action versus state based logics for transition systems. In *Proceedings LITP Spring School on Theoretical Computer Science La Roche Posay (Semantics of Systems of Concurrent Processes)*, pages 407–419. Springer, 1990.

[EC82] E. Allen Emerson and Edmund M. Clarke. Using branching time temporal logic to synthesize synchronization skeletons. *Science of Computer Programming*, 2(3):241–266, 1982.

[Edw78] David Alan Edwards. On the existence of probability measures with given marginals. *Annales de l'institut Fourier*, 28(4):53–78, 1978.

[EG16] Lukas Esterle and Radu Grosu. Cyber-physical systems: challenge of the 21st century. *e & i Elektrotechnik und Informationstechnik*, 133(7):299–303, 2016.

[EGF15] Christian Ellen, Sebastian Gerwinn, and Martin Fränzle. Statistical model checking for stochastic hybrid systems involving nondeterminism over continuous domains. *International Journal on Software Tools for Technology Transfer*, 17(4):485–504, 2015.

[EHZ10] 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.

[EMA16] Sadegh Esmaeil Zadeh Soudjani, Rupak Majumdar, and Alessandro Abate. Safety Verification of Continuous-Space Pure Jump Markov Processes. In *22nd Conference on Tools and Algorithms for the Construction and Analysis of Systems* (*TACAS*), pages 147–163. Springer, 2016.

[FHH$^+$11] 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.

[FKP17] Nathanaël Fijalkow, Bartek Klin, and Prakash Panangaden. Expressiveness of probabilistic modal logics, revisited. In *Proceedings of the 44th International Colloquium on Automata, Languages and Programming* (*ICALP*), volume 80 of *Leibniz International Proceedings in Informatics* (*LIPIcs*), pages 105:1–105:12. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017.

[FKS16] Nathanaël Fijalkow, Stefan Kiefer, and Mahsa Shirmohammadi. Trace Refinement in Labelled Markov Decision Processes. In *Proceedings of the 19th International Conference on Foundations of Software Science and Computational Structures* (*FOSSACS*), pages 303–318. Springer, 2016.

[Fre01] David H. Fremlin. *Measure Theory Volume 1*. Torres-Fremlin, 2001.

[Fre08] Goran Frehse. PHAVer: algorithmic verification of hybrid systems past HyTech. *International Journal on Software Tools for Technology Transfer*, 10(3):263–279, 2008.

[FTE10] Martin Fränzle, Tino Teige, and Andreas Eggers. Satisfaction Meets Expectations. In *Proceedings of the 8th International Conference on Integrated Formal Methods (IFM)*, pages 168–182. Springer, 2010.

[FZ14] Yuan Feng and Lijun Zhang. When equivalence and bisimulation join forces in probabilistic automata. In *Proceedings of the 19th International Symposium on Formal Methods (FM)*, pages 247–262. Springer, 2014.

[Gao08] Su Gao. *Invariant Descriptive Set Theory*. Chapman & Hall/CRC Pure and Applied Mathematics. CRC Press, 2008.

[GB17] Daniel Gburek and Christel Baier. Stochastic transition systems: composition operators, simulations, and trace distributions. Submitted. Journal version of [GBK16], 2017.

[GB18] Daniel Gburek and Christel Baier. Bisimulations, logics, and trace distributions for stochastic systems with rewards. Accepted for publication in the proceedings of the 16th International Conference on Hybrid Systems: Computation and Control (HSCC), 2018.

[GBK16] Daniel Gburek, Christel Baier, and Sascha Klüppelholz. Composition of stochastic transition systems based on spans and couplings. In *Proceedings of the 43rd International Colloquium on Automata, Languages and Programming (ICALP)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 102:1–102:15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016.

[Gir82] Michéle Giry. A categorical approach to probability theory. In *Proceedings of an International Conference Held at Carleton University, Ottawa, 1981*, volume 915 of *Lecture notes in Mathematics*, pages 68–85. Springer, 1982.

[GJP04] Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximate Reasoning for Real-Time Probabilistic Processes. *International Conference on Quantitative Evaluation of Systems*, 0:304–313, 2004.

[GJP06] Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximate reasoning for real-time probabilistic processes. *Logical Methods in Computer Science*, 2(1), 2006.

[GJS94] Alessandro Giacalone, Chi-Chang Jou, and Scott A. Smolka. *Algebraic Reasoning for Probabilistic Concurrent Systems*. Department of Computer Science, SUNY, 1994.

[Gly89] Peter W. Glynn. A GSMP formalism for discrete event systems. In *Proceedings of the IEEE*, pages 14–23, 1989.

[GSS95] Rob J. van Glabbeek, Scott A. Smolka, and Bernhard Steffen. Reactive, Generative, and Stratified Models of Probabilistic Processes. *Information and Computation*, 121(1):59–80, 1995.

[Hah13] Ernst Moritz Hahn. *Model checking stochastic hybrid systems*. PhD thesis, Universität des Saarlandes, 2013.

[HAV16] Sofie Haesaert, Alessandro Abate, and Paul M. J. Van den Hoff. Verification of General Markov Decision Processes by Approximate Similarity Relations and Policy Refinement. In *Proceedings of the 13th International Conference on Quantitative Evaluation of Systems (QEST)*, pages 227–243. Springer, 2016.

[HCH+02] Boudewijn Haverkort, Lucia Cloth, Holger Hermanns, Joost-Pieter Katoen, and Christel Baier. Model checking performability properties. In *Proceedings of the 3rd International Conference on Dependable Systems and Networks (DSN)*, pages 103–112, 2002.

[Hen96] 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.

[Hen12] Matthew Hennessy. Exploring probabilistic bisimulations, part I. *Formal Aspects of Computing*, 24(4):749–768, 2012.

[Her02] Holger Hermanns. *Interactive Markov Chains: And the Quest for Quantified Quality*. Springer, 2002.

[HH13] Ernst Moritz Hahn and Holger Hermanns. Rewarding probabilistic hybrid automata. In *Proceedings of the 16th International conference on Hybrid systems: computation and control (HSCC)*, pages 313–322. ACM, 2013.

[HH15] Arnd Hartmanns and Holger Hermanns. In the quantitative automata zoo. *Science of Computer Programming*, 112:3–23, 2015.

[HHHK12] 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.

[HHWt97] Thomas A. Henzinger, Pei-Hsin Ho, and Howard Wong-toi. HyTech: A Model Checker for Hybrid Systems. *Software Tools for Technology Transfer*, 1:460–463, 1997.

[Hil96] Jane Hillston. *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.

[HJ90] Hans Hansson and Bengt Jonsson. A calculus for communicating systems with time and probabilities. In *Proceedings of the 11th IEEE Symposium on Real-Time Systems (RTS)*, pages 278–287, 1990.

[HJ94] Hans Hansson and Bengt Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing*, 6(5):512–535, 1994.

[HJS07] Ichiro Hasuo, Bart Jacobs, and Ana Sokolova. Generic Trace Semantics via Coinduction. *Logical Methods in Computer Science*, 3(4), 2007.

[HKK14] 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.

[HM80] Matthew Hennessy and Robin Milner. On Observing Nondeterminism and Concurrency. In *7th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 299–309. Springer, 1980.

[HM85] Matthew Hennessy and Robin Milner. Algebraic Laws for Nondeterminism and Concurrency. *Journal of the ACM*, 32(1):137–161, 1985.

[Hoa85] Charles Antony Richard Hoare. *Communicating Sequential Processes*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1985.

[HPV81] C. Himmelberg, T. Parthasarathy, and F. Van Vleck. On measurable relation. *Fundamenta Mathematicae*, 111(2):161–167, 1981.

[Iof79] A. D. Ioffe. Single-Valued Representation of Set-Valued Mappings. *Transactions of the American Mathematical Society*, 252:133–145, 1979.

[Jon90] Claire Jones. *Probabilistic Non-Determinism*. PhD thesis, University of Edinburgh, 1990.

[JS09] Bart Jacobs and Ana Sokolova. Traces, Executions and Schedulers, Coalgebraically. In *Proceedings of the 3th International Conference on Algebra and Coalgebra in Computer Science (CALCO)*, pages 206–220. Springer, 2009.

[JSS15] Bart Jacobs, Alexandra Silva, and Ana Sokolova. Trace semantics via determinization. *Journal of Computer and System Sciences*, 81(5):859–879, 2015.

[Kal02] Olav Kallenberg. *Foundations of Modern Probability*. Springer, 2002.

[Kan95] Akihiro Kanamori. The Emergence of Descriptive Set Theory. In *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, pages 241–262. Springer, 1995.

[Kec95] Alexander S. Kechris. *Classical Descriptive Set Theory*, volume 156 of *Graduate Texts in Mathematics*. Springer, 1995.

[Kel84] Hans G. Kellerer. Duality theorems for marginal problems. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(4):399–432, 1984.

[KK12] 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.

[KKO77] T. Kamae, U. Krengel, and G. L. O'Brien. Stochastic Inequalities on Partially Ordered Spaces. *The Annals of Probability*, 5(6):899–912, 1977.

[KNP11] Marta Kwiatkowska, Gethin Norman, and David Parker. PRISM 4.0: Verification of Probabilistic Real-time Systems. In *Proceedings of the 23rd International Conference on Computer Aided Verification (CAV)*, volume 6806 of *Lecture Notes in Computer Science*, pages 585–591. Springer, 2011.

[KNSS02] Marta Kwiatkowska, Gethin Norman, Roberto Segala, and Jeremy Sproston. Automatic verification of real-time systems with discrete probability distributions. *Theoretical Computer Science*, 282(1):101–150, 2002.

[Kop96] Peter William Kopke. *The theory of rectangular hybrid automata*. PhD thesis, Cornell University, New York, 1996.

[KVAK10] Vijay Anand Korthikanti, Mahesh Viswanathan, Gul Agha, and YoungMin Kwon. Reasoning about MDPs as Transformers of Probability Distributions. In *Proceedings of the 7th International Conference on Quantitative Evaluation of Systems (QEST)*, pages 199–208, 2010.

[Leb05] Henri Lebesgue. Sur les fonctions représentables analytiquement. *Journal de Mathématiques Pures et Appliquées*, 1:139–216, 1905.

[Les10] Lasse Leskelä. Stochastic Relations of Random Variables and Processes. *Journal of Theoretical Probability*, 23(2):523–546, 2010.

[LHK01] Gabriel G. Infante López, Holger Hermanns, and Joost-Pieter Katoen. Beyond Memoryless Distributions: Model Checking Semi-Markov Chains. In *1st Joint International Workshop on Process Algebra and Probabilistic Methods - Performance Modelling and Verification (PAPM-PROBMIV)*, pages 57–70. Springer, 2001.

[Lin92] Torgny Lindvall. *Lectures on the coupling method*. Wiley series in probability and mathematical statistics. Wiley, 1992.

[Lin99] Torgny Lindvall. On Strassen's Theorem on Stochastic Domination. *Electronic Communications in Probability*, 4:51–59, 1999.

[Lov12a] László Lovász. Coupling measure concentrated on a given set. Unpublished notes on [Lov12b], 2012.

[Lov12b] László Lovász. *Large Networks and Graph Limits*. American Mathematical Society colloquium publications. American Mathematical Society, 2012.

[Low91] Gavin Lowe. *Probabilities and Priorities in Timed CSP*. PhD thesis, University of Oxford, 1991.

[LPS00] Gerardo Lafferriere, George J. Pappas, and Shankar Sastry. O-Minimal Hybrid Systems. *Mathematics of Control, Signals, and Systems*, 13(1):1–21, 2000.

[LPW09] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. Providence, R.I. American Mathematical Society, 2009.

[LS89] Nancy A. Lynch and Eugene W. Stark. A proof of the Kahn principle for input/output automata. *Information and Computation*, 82(1):81–92, 1989.

[LS91] Kim G. Larsen and Arne Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, 1991.

[LSV03] Nancy Lynch, Roberto Segala, and Frits Vaandrager. Hybrid I/O automata. *Information and Computation*, 185(1):105–157, 2003.

[LSV07] Nancy Lynch, Roberto Segala, and Frits Vaandrager. Observing Branching Structure through Probabilistic Contexts. *SIAM Journal on Computing*, 37(4):977–1013, 2007.

[LT87] Nancy A. Lynch and Mark R. Tuttle. Hierarchical Correctness Proofs for Distributed Algorithms. In *Proceedings of the 6th Annual ACM Symposium on Principles of Distributed Computing (PODC)*, pages 137–151. ACM, 1987.

[Mil82] Robin Milner. *A Calculus of Communicating Systems*. Springer, 1982.

[ML07] 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.

[MM04] Annabelle McIver and Carroll Morgan. *Abstraction, Refinement And Proof For Probabilistic Systems (Monographs in Computer Science)*. Springer, 2004.

[MS00] A. Maitra and W. Sudderth. Randomized strategies and terminal distributions. In *Game theory, optimal stopping, probability and statistics*, volume 35 of *Lecture Notes–Monograph Series*, pages 39–52. Institute of Mathematical Statistics, 2000.

[NDN+16] Johanna Nellen, Kai Driessen, Martin Neuhäußer, Erika Ábrahám, and Benedikt Wolters. Two CEGAR-based approaches for the safety verification of PLC-controlled plants. *Information Systems Frontiers*, 18(5):927–952, 2016.

[Nel06] Roger B. Nelsen. *An introduction to Copulas*. Springer Series in Statistics. Springer, 2006.

[NK07] Martin R. Neuhäußer and Joost-Pieter Katoen. Bisimulation and Logical Preservation for Continuous-Time Markov Decision Processes. In *Proceedings of the 18th International Conference on Concurrency Theory (CONCUR)*, pages 412–427. Springer, 2007.

[Oxt71] John C. Oxtoby. *Measure and Category: A Survey of the Analogies between Topological and Measure Spaces*, volume 2 of *Graduate Texts in Mathematics*. Springer, 1971.

[Pan09] Prakash Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.

[Pla08] André Platzer. Differential Dynamic Logic for Hybrid Systems. *Journal of Automated Reasoning*, 41(2):143–189, 2008.

[Pla10] André Platzer. Differential-algebraic Dynamic Logic for Differential-algebraic Programs. *Journal of Logic and Computation*, 20(1):309–352, 2010.

[Pla12] André Platzer. A Complete Axiomatization of Quantified Differential Dynamic Logic for Distributed Hybrid Systems. *Logical Methods in Computer Science*, 8(4):1–44, 2012.

[Pla15] André Platzer. Differential Game Logic. *ACM Transactions on Computational Logic*, 17(1):1:1–1:51, 2015.

[PLS00] Anna Philippou, Insup Lee, and Oleg Sokolsky. Weak Bisimulation for Probabilistic Systems. In *Proceedings of the 11th International Conference on Concurrency Theory (CONCUR)*, pages 334–349. Springer, 2000.

[Pnu77] Amir Pnueli. The temporal logic of programs. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science (SFCS)*, pages 46–57, 1977.

[Put94] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 1 edition, 1994.

[Rie78] Ulrich Rieder. Measurable selection theorems for optimization problems. *manuscripta mathematica*, 24(1):115–131, 1978.

[SA13] Sadegh Esmaeil Zadeh Soudjani and Alessandro Abate. Adaptive and Sequential Gridding Procedures for the Abstraction and Verification of Stochastic Processes. *SIAM Journal on Applied Dynamical Systems*, 12(2):921–956, 2013.

[Sch08] Klaus D. Schmidt. *Maß und Wahrscheinlichkeit*. Springer, 2008.

[Seg95] Roberto Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, Massachusetts Institute of Technology, 1995.

[SGA15] Sadegh Esmaeil Zadeh Soudjani, Caspar Gevaerts, and Alessandro Abate. FAUST $^2$: Formal Abstractions of Uncountable-STate STochastic Processes. In *Proceedings of the 21th Conference on Tools and Algorithms for the Construction and Analysis of Systems* (*TACAS*), volume 9035 of *Lecture Notes in Computer Science*, pages 272–286. sv, 2015.

[Ska93] Heinz J. Skala. The Existence of Probability Measures with Given Marginals. *The Annals of Probability*, 21(1):136–142, 1993.

[SL94] Roberto Segala and Nancy Lynch. Probabilistic simulations for probabilistic processes. In *Proceedings of the 11th International Conference on Concurrency Theory* (*CONCUR*), pages 481–496. Springer, 1994.

[Spr01] Jeremy Sproston. *Model Checking for Probabilistic Timed and Hybrid Systems*. PhD thesis, University of Birmingham, 2001.

[Spr11] Jeremy Sproston. Discrete-Time Verification and Control for Probabilistic Rectangular Hybrid Automata. In *Proceedings of the 8th International Conference on Quantitative Evaluation of Systems* (*QEST*), pages 79–88, 2011.

[Spr15] Jeremy Sproston. Verification and Control of Probabilistic Rectangular Hybrid Automata. In *Proceedings of the 13th International Conference on Formal Modeling and Analysis of Timed Systems* (*FORMATS*), pages 1–9. Springer, 2015.

[Sri08] Shashi M. Srivastava. *A Course on Borel Sets*, volume 180 of *Graduate texts in mathematics*. Springer, 2008.

[SS11] Alexandra Silva and Ana Sokolova. Sound and Complete Axiomatization of Trace Semantics for Probabilistic Systems. *Electronic Notes in Theoretical Computer Science*, 276:291–311, 2011.

[Str65] Volker Strassen. The Existence of Probability Measures with Given Marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, 1965.

[Str05] Stefan Nicolaas Strubbe. *Compositional modelling of stochastic hybrid systems*. PhD thesis, University of Twente, 2005.

[Swa96] J.M. Swart. A conditional product measure theorem. *Statistics & Probability Letters*, 28(2):131 – 135, 1996.

[Ter15] Pedro Sánchez Terraf. Bisimilarity is not Borel. *Mathematical Structures in Computer Science*, pages 1–20, 2015.

[Tin07] Simone Tini. Notes on Generative Probabilistic Bisimulation. *Electronic Notes in Theoretical Computer Science*, 175(1):77–88, 2007.

[TMKA13] Ilya Tkachev, Alexandru Mereacre, Joost-Pieter Katoen, and Alessandro Abate. Quantitative Automata-based Controller Synthesis for Non-autonomous Stochastic Hybrid Systems. In *Proceedings of the 16th International Conference on Hybrid Systems: Computation and Control (HSCC)*, pages 293–302. ACM, 2013.

[TMKA16] Ilya Tkachev, Alexandru Mereacre, Joost-Pieter Katoen, and Alessandro Abate. Quantitative model-checking of controlled discrete-time Markov processes. *Information and Computation*, 2016.

[Tof94] Chris Tofts. Processes with probabilities, priority and time. *Formal Aspects of Computing*, 6(5):536–564, 1994.

[vB76] Johan van Benthem. *Modal correspondence theory*. PhD thesis, University of Amsterdam, 1976.

[vG90] Rob J. van Glabbeek. The Linear Time - Branching Time Spectrum. In *Proceedings of the 1st International Conference on Concurrency Theory (CONCUR)*, pages 278–297. Springer, 1990.

[vG93] Rob J. van Glabbeek. The Linear Time - Branching Time Spectrum II. In *Proceedings of the 4th International Conference on Concurrency Theory (CONCUR)*, pages 66–81. Springer, 1993.

[Vil09] Cédric Villani. *Optimal transport: old and new*. Grundlehren der mathematischen Wissenschaften. Springer, 2009.

[Wag80] Daniel H. Wagner. Survey of measurable selection theorems: An update. pages 176–219. Springer, 1980.

[Whi80] Ward Whitt. Continuity of Generalized Semi-Markov Processes. *Mathematics of Operations Research*, 5(4):494–501, 1980.

[WJ06] Nicolás Wolovick and Sven Johr. A Characterization of Meaningful Schedulers for Continuous-Time Markov Decision Processes. In *Proceedings of the*

*4th International Conference on Formal Modeling and Analysis of Timed Systems* (*FORMATS*), volume 4202 of *Lecture Notes in Computer Science*, pages 352–367. Springer, 2006.

[Wol12] Nicolás Wolovick. *Continuous Probability and Nondeterminism in Labeled Transition Systems*. PhD thesis, FaMAF - Universidad Nacional de Córdoba, 2012.

[WSS97] Sue-hwey Wu, Scott A. Smolka, and Eugene W. Stark. Composition and behaviors of probabilistic I/O automata. *Theoretical Computer Science*, 176(1–2):1–38, 1997.

[Yi91] Wang Yi. CCS + Time = An Interleaving Model for Real Time Systems. In *Proceedings of the 18th International Colloquium on Automata, Languages and Programming* (*ICALP*), volume 510 of *Lecture Notes in Computer Science*, pages 217–228. Springer, 1991.

[Yi94] Wang Yi. Algebraic Reasoning for Real-Time Probabilistic Processes with Uncertain Information. In *Proceedings of the 3rd International Symposium Organized Jointly with the Working Group Provably Correct Systems on Formal Techniques in Real-Time and Fault-Tolerant Systems* (*ProCoS*), pages 680–693. Springer, 1994.

[YJZ17] Pengfei Yang, David N. Jansen, and Lijun Zhang. Distribution-Based Bisimulation for Labelled Markov Processes. In *Proceedings of the 15th International Conference on Formal Modeling and Analysis of Timed Systems* (*FORMATS*), pages 170–186. Springer, 2017.

[YL92] Wang Yi and Kim Guldstrand Larsen. Testing Probabilistic and Nondeterministic Processes. In *Proceedings of the 12th International Symposium on Protocol Specification, Testing and Verification* (*PSTV*), pages 47–61. North-Holland Publishing Co., 1992.

[ZBC12] Paolo Zuliani, Christel Baier, and Edmund M. Clarke. Rare-event Verification for Stochastic Hybrid Systems. In *Proceedings of the 15th International Conference on Hybrid Systems: Computation and Control* (*HSCC*), pages 217–226. ACM, 2012.

[ZEM+14] Majid Zamani, Peyman Mohajerin Esfahani, Rupak Majumdar, Alessandro Abate, and John Lygeros. Symbolic Control of Stochastic Systems via Approximately Bisimilar Finite Abstractions. *IEEE Transactions on Automatic Control*, 59(12):3135–3150, 2014.

[ZSR+12] Lijun Zhang, Zhikun She, Stefan Ratschan, Holger Hermanns, and Ernst Moritz Hahn. Safety Verification for Probabilistic Hybrid Systems. *European Journal of Control*, 18(6):572–587, 2012.