

Bisimulation for Feller-Dynkin Processes

Linan Chen¹

*Department of Mathematics and Statistics
McGill University
Montreal, Canada*

Florence Clerc² Prakash Panangaden^{3,4}

*School of Computer Science
McGill University
Montreal, Canada*

Abstract

Bisimulation is a concept that captures behavioural equivalence. It has been studied extensively on nonprobabilistic systems and on discrete-time Markov processes and on so-called continuous-time Markov chains. In the latter, time is continuous but the evolution still proceeds in jumps. We propose two definitions of bisimulation on continuous-time stochastic processes where the evolution is a *flow* through time. We show that they are equivalent and we show that when restricted to discrete-time, our concept of bisimulation encompasses the standard discrete-time concept. The concept we introduce is not a straightforward generalization of discrete-time concepts.

Keywords: Stochastic Processes, Markov Processes, continuous time, bisimulation, diffusion.

1 Introduction

Bisimulation [24,26,28] is a fundamental concept in the theory of transition systems capturing a strong notion of behavioural equivalence. In particular, it is a notion stronger than that of trace equivalence. Bisimulation has been widely studied for discrete time systems where transitions happen as steps, both on discrete [23] and continuous state spaces [3,11,25]. In all these types of systems a crucial ingredient of the definition of bisimulation is the ability to talk about *the next step*. Thus, the general format of the definition of bisimulation is that one has some property that

¹ Email: linan.chen@mcgill.ca

² Email: florence.clerc@mail.mcgill.ca

³ Email: prakash@cs.mcgill.ca

⁴ This research has been supported by a grant from NSERC.

must hold “now” (in the states being compared) and then one says that the relation is preserved in the next step.

Some attempts have been made to talk about continuous-time [14], but even in what are called continuous-time Markov chains there is a discrete notion of time *step*; it is only that there is a real-valued duration associated with each state that makes such systems continuous time. They are often called “jump processes” in the mathematical literature, see, for example, [27,31], a phrase that better captures the true nature of such processes.

Outside of computer science, there is a vast range of systems that involve true continuous-time evolution: deterministic systems governed by differential equations and stochastic systems governed by “noisy” differential equations called stochastic differential equations. These have been extensively studied for over a century since the pioneering work of Einstein [16] on Brownian motion. In the computer science literature there have been studies of very special systems that feature continuous time: timed automata [2] and hybrid systems [1]. In these systems the time evolution is assumed to be piecewise constant (timed automata) or piecewise smooth (hybrid automata) and bisimulation is defined without recourse to talking about the next step. However, a general formalism that covers processes like diffusion is not available as far as we are aware.

In this work we aim at a general theory of bisimulation for stochastic systems with true continuous-time evolution. We focus on a class of systems called Feller-Dynkin processes for which a good mathematical theory exists. These systems are the most general version of Markov processes defined on continuous state spaces and with continuous time evolution. Such systems encompass Brownian motion and its many variants.

The obvious extension of previous definitions of bisimulation on discrete Markov processes or on jump processes fail to provide a meaningful notion of behavioural equivalence as we will illustrate later on. It is a mistake to think that one can get a reasonably good understanding of such systems by considering suitable “limits” of discrete-time systems. Intuitively, the notion of bisimulation is sensitive to small changes that are not captured when taking the limit. It is true that, for example, Brownian motion can be seen as arising as a limit, in the sense of convergence in distribution, of a discrete random walk as both the discrete time unit and the step size go to zero. However, entirely new phenomena occur with the trajectories of the Brownian motion which are not understandable through the limiting process at least not in any naive sense: the probability of being at any single state x at a given time t is zero, but the probability of hitting x before a given time s is strictly positive.

To avoid those issues, we work with the set of trajectories of the system. A number of possible ways had to be explored and in the end the particular version we present here turned out to have the desired properties: (a) corresponds to our intuition in a number of examples and (b) correctly specializes to the discrete-time case.

Section 2 explains the mathematical background on Feller-Dynkin processes and

Brownian motion. In section 3, we show why a naive extension of previous definition of bisimulation does not work and we propose a new definition of bisimulation as an equivalence relation that we illustrate on a number of examples. In section 4, we give an equivalent definition of bisimulation as a cospan of morphisms extending the previous notion of span of “zig-zag” morphisms in the discrete-time case. In section 5, we show that our definition of bisimulation is coherent with the previous definition of bisimulation in discrete time. Much remains to be done, of course, as we describe in the concluding section. We will omit most proofs in this version, however, a full version is available at [7].

2 Background on Feller-Dynkin processes

We assume that basic concepts like topology, measure theory and basic concepts of probability on continuous spaces are well known; see, for example [5,15,25].

The basic arena for the action is a probability space.

Definition 2.1 A *probability space* is a triple (S, \mathcal{F}, P) where S is a space (usually some kind of topological space), \mathcal{F} is a σ -algebra (usually its Borel algebra) and P is a probability measure on \mathcal{F} .

Given a measurable space (X, Σ) a (sub)-Markov kernel is a map $\tau : X \times \Sigma \rightarrow [0, 1]$ which is measurable in its first argument, i.e. $\tau(\cdot, A \in \Sigma) : X \rightarrow \mathbb{R}$ is measurable for any fixed A in Σ and for any fixed $x \in X$, $\tau(x, \cdot)$ is a (sub)probability measure. These kernels describe transition probability functions.

A crucial concept is that of a *filtration*. They will play a central role in the description of a process.

Definition 2.2 A *filtration* on a measurable space (Ω, \mathcal{F}) is a nondecreasing family $(\mathcal{F}_t)_{t \geq 0}$ of sub- σ -algebras of \mathcal{F} , i.e. $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \leq s < t < \infty$.

This concept is used to capture the idea that at time t what is “known” or “observed” about the process is encoded in the sub- σ -algebra \mathcal{F}_t .

Definition 2.3 A *stochastic process* is a collection of random variables $(X_t)_{0 \leq t < \infty}$ on a measurable space (Ω, \mathcal{F}) that take values in a second measurable space (S, \mathcal{S}) called the *state space*. We say that a stochastic process is *adapted* to a filtration $(\mathcal{F}_t)_{t \geq 0}$ if for each $t \geq 0$ we have X_t is \mathcal{F}_t -measurable.

Note that a stochastic process is always adapted to the filtration $(\mathcal{G}_t)_{t \geq 0}$, where for each $t \geq 0$, \mathcal{G}_t is defined as the σ -algebra generated by all the random variables $\{X_s | s \leq t\}$. The filtration $(\mathcal{G}_t)_{t \geq 0}$ is also referred to as the natural filtration associated to $(X_t)_{t \geq 0}$.

Before stating the definition of the continuous-time processes we will be interested in, let us first start by recalling the definition of their discrete-time counterparts.

Definition 2.4 A labelled Markov process (LMP) is a triple (X, Σ, τ) where (X, Σ) is a measurable space and τ is a Markov kernel.

We will quickly review the theory of continuous-time processes on continuous state space; much of this material is adapted from “Diffusions, Markov Processes and Martingales, Volume I” by Rogers and Williams [27] and we use their notations. Another useful source is “Functional analysis for probability and stochastic processes” by A. Bobrowski [6]. Let E be a locally compact Hausdorff space with countable base and let it be equipped with the Borel σ -algebra $\mathcal{E} = \mathcal{B}(E)$. E_∂ is the one-point compactification of E : $E_\partial = E \uplus \{\partial\}$. The physical picture is that the added state, ∂ , represents a point at infinity; we will view it as an absorbing state.

We say that a continuous real-valued function f on E “vanishes at infinity” if for every $\varepsilon > 0$ there is a compact subset $K \subset E$ such that $\forall x \in E \setminus K$ we have $|f(x)| \leq \varepsilon$. This space is a Banach space with the sup norm.

Definition 2.5 A *semigroup* of operators on any Banach space is a family of linear continuous (bounded) operators T_t indexed by $t \in \mathbf{R}^{\geq 0}$ such that

$$\forall s, t \geq 0, T_s \circ T_t = T_{s+t}$$

and

$$T_0 = I \quad (\text{the identity}).$$

The first equation above is called the semigroup property. The operators in a semigroup are continuous however there is a useful continuity property of the semigroup as a whole.

Definition 2.6 For X a Banach space, we say that a semigroup $T_t : X \rightarrow X$ is *strongly continuous* if

$$\forall x \in X, \lim_{t \downarrow 0} T_t x = x$$

which is equivalent to saying

$$\forall x \in X, \lim_{t \downarrow 0} \|T_t x - x\| \rightarrow 0.$$

Definition 2.7 A *Feller-Dynkin (FD) semigroup* is a strongly continuous semigroup $(\hat{P}_t)_{t \geq 0}$ of linear operators on $C_0(E)$ (the space of continuous functions on E which vanish at infinity) satisfying the additional condition:

$$\forall t \geq 0 \quad \forall f \in C_0(E), \text{ if } 0 \leq f \leq 1, \text{ then } 0 \leq \hat{P}_t f \leq 1$$

The following important proposition relates these FD semigroups with Markov processes which allows one to see the connection with more familiar probabilistic transition systems.

Proposition 2.8 *Given such an FD semigroup, it is possible to define a unique family of sub-Markov kernels $(P_t)_{t \geq 0} : E \times \mathcal{E} \rightarrow [0, 1]$ such that for all $t \geq 0$ and $f \in C_0(E)$,*

$$\hat{P}_t f(x) = \int f(y) P_t(x, dy).$$

A very important ingredient in the theory is the space of trajectories of a FD process (FD semigroup) as a probability space. This space does not appear explicitly in the study of labelled Markov processes but one does see it in the study of continuous-time Markov chains and jump processes.

Definition 2.9 We define a *trajectory* ω on E_∂ to be a *cadlag*⁵ function $[0, \infty) \rightarrow E_\partial$ such that if either $\omega(t-) := \lim_{s \leq t, s \rightarrow t} \omega(s) = \partial$ or $\omega(t) = \partial$ then $\forall u \geq t, \omega(u) = \partial$. We can extend ω to a map from $[0, \infty]$ to E_∂ by setting $\omega(\infty) = \partial$.

It is possible to associate to such an FD semigroup a *canonical FD process*. Let Ω be the set of trajectories $\omega : [0, \infty) \rightarrow E_\partial$.

Definition 2.10 The *canonical FD process* associated to the FD semigroup (\hat{P}_t) is

$$(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, (X_t)_{0 \leq t \leq \infty}, (\mathbb{P}^x)_{x \in E_\partial})$$

where

- $X_t(\omega) = \omega(t)$
- $\mathcal{G} = \sigma(X_s \mid 0 \leq s < \infty)$ ⁶, $\mathcal{G}_t = \sigma(X_s \mid 0 \leq s \leq t)$
- given any probability measure μ on E_∂ , by the Kolmogorov extension theorem, there exists a unique probability measure \mathbb{P}^μ on (Ω, \mathcal{G}) such that for all $n \in \mathbb{N}$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$ and x_0, x_1, \dots, x_n in E_∂ ,

$$\mathbb{P}^\mu(X_0 \in dx_0, X_{t_1} \in dx_1, \dots, X_{t_n} \in dx_n) = \mu(dx_0)P_{t_1}^{+\partial}(x_0, dx_1) \dots P_{t_n - t_{n-1}}^{+\partial}(x_{n-1}, dx_n)$$
⁷

where $P_t^{+\partial}$ is the Markov kernel extending the Markov kernel P_t to E_∂ by $P_t^{+\partial}(x, \{\partial\}) = 1 - P_t(x, E)$ and $P_t^{+\partial}(\partial, \{\partial\}) = 1$. We set $\mathbb{P}^x = \mathbb{P}^{\delta_x}$.

This is the version of the system that will be most useful for us. In order to bring it more in line with the kind of transition systems that have hitherto been studied in the computer science literature we introduce a finite set of atomic propositions AP and such a FD process is equipped with a function $obs : E \rightarrow 2^{AP}$. This function is extended to a function $obs : E_\partial \rightarrow 2^{AP} \uplus \{\partial\}$ by setting $obs(\partial) = \partial$.

Instead of following the dynamics of the system step by step as one does in a discrete system we have to study the behaviour of sets of trajectories. The crucial ingredient is the distribution \mathbb{P}^x which gives a measure on the space of trajectories for a system started at the point x .

2.1 Brownian motion as a FD process

Brownian motion is a stochastic process describing the irregular motion of a particle being buffeted by invisible molecules. Now its range of applicability extends far beyond its initial application [22]. The following definition is from [22].

⁵ By *cadlag* we mean right-continuous with left limits.

⁶ The σ -algebra \mathcal{G} is the same as the one induced by the Skorohod metric, see theorem 16.6 of [4]

⁷ The dx_i in this equation should be understood as infinitesimal volumes. This notation is standard in probabilities and should be understood by integrating it over measurable state sets C_i .

Definition 2.11 A standard one-dimensional Brownian motion is a Markov process adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$,

$$B = (W_t, \mathcal{F}_t), 0 \leq t < \infty$$

defined on a probability space (Ω, \mathcal{F}, P) with the properties

- (i) $W_0 = 0$ almost surely,
- (ii) for $0 \leq s < t$, $W_t - W_s$ is independent of \mathcal{F}_s and is normally distributed with mean 0 and variance $t - s$.

In this very special process, one can start at any place, there is an overall translation symmetry which makes calculations more tractable. In order to do any calculations we use the following fundamental formula: If the process is at x at time 0 then at time t the probability that it is in the (measurable) set D is given by

$$P_t(x, D) = \int_{y \in D} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right) dy.$$

The associated FD semigroup is the following: for $f \in C_0(\mathbb{R})$ and $x \in \mathbb{R}$,

$$\hat{P}_t(f)(x) = \int_y \frac{f(y)}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right) dy.$$

3 Bisimulation

The concept of bisimulation is fundamental and its history is well documented [28]. We recall the definition of bisimulation on continuous state spaces with discrete time steps [11,25], we call it a *DT-bisimulation* to emphasize that it pertains to discrete-time systems. We consider LMPs equipped with a family of atomic propositions AP where $A \in AP$ is interpreted on a specific LMP as a subset of the state space represented by its characteristic function χ_A .

Definition 3.1 Given an LMP $(X, \Sigma, \tau, (\chi_A)_{A \in AP})$, a *DT-bisimulation* R is an equivalence relation on X such that if xRy , then

- for all $A \in AP$, $\chi_A(x) = \chi_A(y)$
- for all R -closed sets $C \in \Sigma$ (i.e. if $z \in C$ and yRz , then $y \in C$), $\tau(x, C) = \tau(y, C)$.

3.1 Naive approach

The key idea of bisimulation is that “what can be observed now is the same” and bisimulation is preserved by the evolution. In order to capture this we need two conditions: the first captures what is immediately observable and the second captures the idea that the evolution preserves bisimulation.

Consider the naive extension of bisimulation in discrete time: let R be an equivalence relation on the state space E such that whenever $x R y$ ($x, y \in E$):

(initiation 1) $obs(x) = obs(y)$, and

(induction 1) for all R -closed sets C in \mathcal{E} , for all times t , $P_t(x, C) = P_t(y, C)$

Let us illustrate on an example why this definition is not enough.

We consider the case of Brownian motion on the reals where there is a single atomic proposition marking 0: $obs(0) = 1$ and $obs(x) = 0$ for $x \neq 0$. Intuitively, we would like that two states x and y are bisimilar if and only if $|x| = |y|$ as the only symmetry that this system has is point reflection with respect to 0.

However, the two conditions [\(initiation 1\)](#) and [\(induction 1\)](#) are not strong enough to enforce that this equivalence relation is the greatest bisimulation.

Let us define the equivalence

$$R = (\mathbb{R}^* \times \mathbb{R}^*) \cup \{(0, 0)\} \text{ where } \mathbb{R}^* = \mathbb{R} \setminus \{0\}.$$

This equivalence satisfies both conditions [\(induction 1\)](#) and [\(initiation 1\)](#). The last one follows directly from the definitions of R and obs .

For the induction condition, the only R -closed sets are \emptyset , $\{0\}$, \mathbb{R}^* and \mathbb{R} , and for any state $z \neq 0$ and time $t \geq 0$, $P_t(z, \emptyset) = P_t(z, \{0\}) = 0$ and $P_t(z, \mathbb{R}^*) = P_t(z, \mathbb{R}) = 1$.

For the proof that our notion of bisimulation yields the expected greatest bisimulation, we refer the reader to section [3.4](#).

3.2 Definition

As we have just shown, unlike in the discrete-time case we cannot just say that the “next step” preserves the relation. Therefore we have to talk about the trajectories; but then we need to choose the right condition on sets of trajectories.

Definition 3.2 An equivalence relation R on the state space E is a *bisimulation* if whenever xRy , the following conditions are satisfied:

(initiation 1) $obs(x) = obs(y)$, and

(induction 2) for all R -closed sets B in \mathcal{G} , $\mathbb{P}^x(B) = \mathbb{P}^y(B)$ where by R -closed, we mean that for all $\omega \in B$ if a trajectory ω' is such that for all time $t \geq 0$, $\omega(t)R\omega'(t)$, then $\omega' \in B$.

Clearly equality is trivially a bisimulation, and furthermore, by definition of \mathbb{P}^x , condition [\(induction 2\)](#) implies [\(induction 1\)](#).

We have chosen to give names to the conditions. The reason for choosing those names will become clear in section [3.3.2](#).

Remark 3.3 Usually, for discrete time, instead of a single kernel τ , a labelled Markov process is a family of Markov kernels indexed by a family of *actions*. These actions correspond to the environment or the user acting on the process. The second condition of bisimulation is then stated on the corresponding Markov kernels for all actions. It is possible to do the same for continuous-time. We can consider a family of FD processes indexed by a set of actions. Condition [\(induction 2\)](#) is then stated

for all these actions. The remaining of this paper can also be adapted to several actions in a similar way.

Lemma 3.4 *An equivalence relation R is a bisimulation if and only if whenever xRy , the following conditions are satisfied:*

(*initiation 2*) for all obs-closed sets B in \mathcal{G} , $\mathbb{P}^x(B) = \mathbb{P}^y(B)$ where by obs-closed, we mean that for all $\omega \in B$ if a trajectory ω' is such that $\text{obs} \circ \omega = \text{obs} \circ \omega'$, then $\omega' \in B$.

(*induction 2*) for all R -closed sets B in \mathcal{G} , $\mathbb{P}^x(B) = \mathbb{P}^y(B)$.

Definition 3.5 Two states are *bisimilar* if there is a bisimulation that relates them.

Proposition 3.6 *The relation “is bisimilar to” is the greatest bisimulation.*

We now consider several examples and give their greatest bisimulation. Proving that an equivalence is the greatest bisimulation follows the following outline: first proving that the equivalence satisfies conditions (*initiation 1*) and (*induction 2*) (and hence it is a bisimulation), and then using (*initiation 2*) to prove that it is the greatest bisimulation possible.

3.3 Basic examples

3.3.1 Deterministic Drift

Consider a deterministic drift on the real line \mathbb{R} with constant speed $a \in \mathbb{R}$ (i.e. in terms of random variable $X_t = X_0 + at$). We consider two cases: with 0 as the only distinguished point and with all the integers distinguished from the other points.

With zero distinguished:

Consider the case when there is a single atomic proposition called *obs*, and $\text{obs}(x) = 1$ if and only if $x = 0$.

Proposition 3.7 *Two states x and y are bisimilar if and only if either $ax > 0$ and $ay > 0$ or $x = y$.*

Proof. Without loss of generality, we can assume that $a > 0$ (case $a < 0$ works in a similar fashion and case $a = 0$ is trivial). Denote R the corresponding relation.

Consider x and y such that xRy . We have to consider two cases:

- If $\text{obs}(x) = 1$, this means that $x = 0$. The state 0 is only bisimilar to itself, which means that $y = 0$ and therefore $\text{obs}(x) = \text{obs}(y)$.
- If $\text{obs}(x) = 0$, this means that $x \neq 0$. The state 0 is only bisimilar to itself, which means that $y \neq 0$ and therefore $\text{obs}(x) = \text{obs}(y)$.

Consider a measurable set B . First, for any $z \in \mathbb{R}$, we denote ω_z the trajectory $\omega_z(t) = z + at$ and note that $\mathbb{P}^z(B) = \delta_B(\omega_z)$.

Consider an R -closed measurable set B . We want to show that $\mathbb{P}^x(B) = \mathbb{P}^y(B)$. For that, there are two cases to consider:

- either $x \leq 0$, in which case $x = y$ since xRy and therefore $\mathbb{P}^x(B) = \mathbb{P}^y(B)$,

- or $x > 0$, in which case $y > 0$. In that case, for all t , $\omega_x(t) > 0$ and $\omega_y(t) > 0$ and in particular for all time t , $\omega_x(t) R \omega_y(t)$. Since B is R -closed, $\omega_x \in B$ if and only if $\omega_y \in B$.

This concludes the second part of the proof.

We now prove that this is the greatest such bisimulation. We are using condition (initiation 2) for that:

Consider $x > 0$ and $y \leq 0$. For all $t \geq 0$, $\omega_x(t) \neq 0$, but $\omega_y(-y/a) = 0$. Define $B = \{\omega \mid \omega(-y/a) = 0\}$. This set is *obs*-closed but $\mathbb{P}^x(B) = 0$ and $\mathbb{P}^y(B) = 1$. These two states cannot be bisimilar.

Consider $x, y \leq 0$. Note that $\omega_x(t) \neq 0$ for all $t \neq -x/a$. Define $B = \{\omega \mid \omega(-y/a) = 0\}$. We have that $\mathbb{P}^y(B) = 1$ and the only way $\mathbb{P}^x(B) = 1$ is to have $-x/a = -y/a$, i.e. $x = y$. This concludes the proof. \square

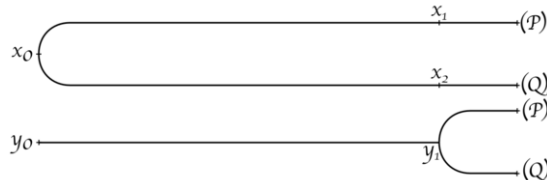
With all integers distinguished:

We consider the case when there is a single atomic proposition and $\text{obs}(x) = 1$ if and only if $x \in \mathbb{Z}$. In this case two states x and y are bisimilar if and only if $x - \lfloor x \rfloor = y - \lfloor y \rfloor$, i.e. $x - y \in \mathbb{Z}$. This is intuitive and it is worth seeing that our definition captures this.

3.3.2 Fork

One could think that since trajectories are already included in the initiation condition (initiation 2), the additional induction condition is not necessary. However, this example illustrates the crucial role of the induction condition in the definition of bisimulation. It is an extension of the standard “vending machine” example in discrete time to our continuous-time setting and it shows that even the condition (induction 1) are enough to discriminate between states that (initiation 2) cannot distinguish.

The full depiction of this example can be found in the full version. Consider the following state space:



There are two atomic propositions (denoted P and Q on the diagram), that are satisfied by the final state of some of the branches. The process is a drift at a constant speed to the right. When it reaches a fork, it moves to either branch with probability $1/2$ (and stops when he hits an atomic proposition).

The basic claim is that the states x_0 and y_0 cannot be bisimilar since states x_1, x_2, y_1 cannot be bisimilar either. This is where condition (induction 2) is really important since the two states x_0 and y_0 have similar traces as they both satisfy

the condition (initiation 2).

3.4 Examples based on Brownian motion

3.4.1 Standard Brownian Motion

With zero distinguished:

Consider the case when there is a single atomic proposition and $obs(x) = 1$ if and only if $x = 0$.

Proposition 3.8 *Two states x and y are bisimilar if and only if $|x| = |y|$.*

Proof. First let us prove that $R = \{(x, y) \mid |x| = |y|\}$ is a bisimulation.

Consider $x R y$, i.e. $|x| = |y|$. This means that $x = 0$ if and only if $y = 0$. In other terms, $obs(x) = 1$ if and only if $obs(y) = 1$ and hence $obs(x) = obs(y)$.

Let now B be an R -closed measurable set of trajectories. This means that $\omega \in B$ if and only if $-\omega \in B$ since $|\omega(t)| = |-\omega(t)|$ for all time $t \geq 0$. And therefore $\mathbb{P}^x(B) = \mathbb{P}^{-x}(-B) = \mathbb{P}^{-x}(B)$ where $-B := \{t \mapsto -\omega(t) \mid \omega \in B\}$.

There now remains to show that this is the greatest bisimulation. It can easily be seen that 0 and $x \neq 0$ since $obs(0) \neq obs(x)$.

Now Consider two different states x and y . We can define the set $B_t = \{\omega \mid \exists s < t \text{ } \omega(s) = 0\}$. This set is obs -closed. It can also be expressed as $B_t = T_0^{-1}([0, t))$ where T_0 is the hitting time for Brownian motion and we know that for any state z ,

$$\mathbb{P}^z(B_t) = \sqrt{\frac{2}{\pi}} \int_{|z|\sqrt{t}}^{\infty} e^{-s^2/2} ds$$

If $|x| \neq |y|$, it is impossible to have that $\mathbb{P}^x(B_t) = \mathbb{P}^y(B_t)$. This proves that no equivalence strictly bigger than R may satisfy (initiation 2). \square

With all integers distinguished:

Consider the case when there is a single atomic proposition and $obs(x) = 1$ if and only if $x \in \mathbb{Z}$.

Proposition 3.9 *Two states x and y are bisimilar if and only if $x - \lfloor x \rfloor = y - \lfloor y \rfloor$ or $\lceil y \rceil - y$.*

This is not trivial and requires a slightly delicate argument which we omit in this short version.

With an interval distinguished:

Consider the case when there is a single atomic proposition and $obs(x) = 1$ if and only if $x \in [-1, 1]$.

Proposition 3.10 *Two states x and y are bisimilar if and only if $|x| = |y|$.*

This is also a calculation that requires explicit properties of Brownian motion and the use of Laplace transforms; we describe it in the full paper.

3.4.2 Brownian motion with drift

Consider a Brownian process with drift: $W'_t = W_t + at$ (where W_t is the standard Brownian motion and $a > 0$, note that the case $a < 0$ is symmetric).

With zero distinguished:

Consider the case when there is a single atomic proposition and $obs(x) = 1$ if and only if $x = 0$.

Proposition 3.11 *Two states x and y are bisimilar if and only if $x = y$.*

Proof. As stated before, the equivalence where a state is only related to itself is a bisimulation.

Let us now show that this is the greatest bisimulation. Consider two different states x and y . Similarly to what we did for the standard Brownian motion, we can rule out the case where $x = 0$ (and $y \neq 0$) or $y = 0$ (and $x \neq 0$) by simply looking at the function obs .

We can define the set $B_t = \{\omega \mid \exists s < t \ \omega(s) = 0\}$. This set is obs -closed. It can also be expressed as $B_t = T_0^{-1}([0, t))$ where T_0 is the hitting time for Brownian motion and we know that for any state z ,

$$\mathbb{P}^z(T_0 \in ds) = \frac{|z|}{\sqrt{2\pi}s^3} \exp\left(-\frac{(z+as)^2}{2s}\right) ds$$

Since we have that for all t , $\mathbb{P}^x(B_t) = \mathbb{P}^y(B_t)$, then we also have that for all $s \geq 0$,

$$|x| \exp\left(-\frac{(x+as)^2}{2s}\right) = |y| \exp\left(-\frac{(y+as)^2}{2s}\right)$$

Since $x, y \neq 0$, we have that for all $s, t \geq 0$,

$$-\frac{(y+as)^2}{2s} + \frac{(x+as)^2}{2s} = -\frac{(y+at)^2}{2t} + \frac{(x+at)^2}{2t}$$

which is equivalent to $(s-t)y^2 = (s-t)x^2$. This means that in that case $|x| = |y|$. Going back to the original expression, we have that for all $s \geq 0$, $-(x+as)^2 = -(y+as)^2$ and therefore $2asx = 2asy$. Since $a \neq 0$, we get that $x = y$ in order to have (initiation 2). \square

With all integers distinguished:

Consider the case when there is a single atomic proposition and $obs(x) = 1$ if and only if $x \in \mathbb{Z}$.

Proposition 3.12 *Two states x and y are bisimilar if and only if $x - \lfloor x \rfloor = y - \lfloor y \rfloor$.*

This is quite a long proof but interesting as it shows many of the complexities of dealing with continuous time. The proof is in the full paper.

With an interval distinguished:

Consider the case when there is a single atomic proposition and $obs(x) = 1$ if and only if $x \in [-1, 1]$. Here also we have the intuitive result but it requires some work to prove it.

Proposition 3.13 *Two states x and y are bisimilar if and only if $x = y$.*

3.4.3 Brownian motion with absorbing wall

Another usual variation on Brownian motion is to add boundaries and to consider that the process does not move anymore or dies once it has hit a boundary. Since all our previous examples involved probability distributions (as opposed to subprobabilities), we will see the boundary as killing the process.

Absorption at 0:

Consider the case of Brownian motion with absorption at the origin and without any atomic proposition. The state space is $\mathbb{R}_{>0}$.

Proposition 3.14 *Two states x and y are bisimilar if and only if $x = y$.*

Proof. We know that equality is a bisimulation. Let us prove that it is the greatest.

For all $t \geq 0$, the set $B_t = \{\omega \mid obs \circ \omega(t) = \partial\}$ is obs -closed. Let us clarify the intuition behind that set B_t : it is the set of trajectories such that the process following one of these trajectories is dead at time t .

For all state $x \geq 0$,

$$\begin{aligned}\mathbb{P}_{abs}^x(B_t) &= \mathbb{P}^x(T_0 < t) \\ &= \frac{2}{\sqrt{2\pi}} \int_{x/\sqrt{t}}^{+\infty} e^{-z^2/2} dz\end{aligned}$$

The only way $\mathbb{P}_{abs}^x(B_t) = \mathbb{P}_{abs}^y(B_t)$ is therefore to have $x = y$. □

Absorption at 0 and b :

Consider the case of Brownian motion with absorption at the origin and at $b > 0$ and without any atomic proposition. The state space is therefore $(0, b)$.

Proposition 3.15 *Two states x and y are bisimilar if and only if $x = y$ or $x = b - y$.*

Absorption at 0 and $4b$ with atomic proposition at b :

In the full version of the paper, we also consider the case where the atomic proposition distinguishes a state that is in the middle of the state space. In the following case, we break that symmetry.

Consider the case of Brownian motion with absorption at the origin and at $4b > 0$, so the state space is $(0, 4b)$, and with a single atomic proposition such that $obs(b) = 1$ and $obs(x) = 0$ for $x \neq b$. The proof can be found in the full version.

Proposition 3.16 *Two states x and y are bisimilar if and only if $x = y$.*

4 Feller-Dynkin cospan

The concept of bisimulation that we have discussed so far is defined between states of a process. One often wants to compare different processes with different state spaces. For this one needs to use functions that relate the state spaces of different processes. One does want to preserve the relational character of bisimulation. In the coalgebra literature one uses spans of so-called “zigzag” morphisms. In previous work [9] on (discrete-time) Markov processes people have considered cospans as this leads to a smoother theory. Intuitively, the difference is whether one thinks of an equivalence relation as a set of ordered pairs or as a collection of equivalence classes.

4.1 Feller-Dynkin homomorphism

This definition of bisimulation can easily be adapted to states in different Markov processes by constructing the disjoint union of the Markov processes.

The disjoint union of two Markov processes is defined as such: given two FD processes $(E_j, \mathcal{E}_j, (\hat{P}_t^j), (P_t^j), \Omega_j, \mathcal{G}_j, (\mathbb{P}_j^x), obs_j)_{j=1,2}$, we write $i_1 : E_1 \rightarrow E_1 \uplus E_2$ and $i_2 : E_2 \rightarrow E_1 \uplus E_2$ for the two corresponding inclusions. The disjoint unions of the two FD processes is the process $(E_1 \uplus E_2, \mathcal{E}, (\hat{P}_t), (P_t), \Omega, \mathcal{G}, (\mathbb{P}^x), obs)$ where:

- the topology on $E_1 \uplus E_2$ is generated by the topologies on E_1 and E_2 : an open set of $E_1 \uplus E_2$ is $i_1(O_1) \cup i_2(O_2)$ where O_1 and O_2 are opens of E_1 and E_2 respectively,
- \mathcal{E} is the Borel-algebra generated by this topology. It can also be expressed as the σ -algebra generated by $\{i_1(C) \mid C \in \mathcal{E}_1\}$ and $\{i_2(C) \mid C \in \mathcal{E}_2\}$,
- for any state $x \in E_1 \uplus E_2$, any time $t \geq 0$ and any function in $C_0(E_1 \uplus E_2)$, we define the semigroup:

$$\hat{P}_t f(x) = \begin{cases} \hat{P}_t^1 f_1(x_1) & \text{if } x = i_1(x_1) \\ \hat{P}_t^2 f_2(x_2) & \text{if } x = i_2(x_2) \end{cases}$$

where $f_j : E_j \rightarrow \mathbb{R}$ is defined by $f_j(y) = f \circ i_j(y)$. The semigroup \hat{P}_t inherits the desired properties from \hat{P}_t^1 and \hat{P}_t^2 for it to be a FD semigroup,

- for any state $x \in E_1 \uplus E_2$, any time $t \geq 0$ and any measurable set $C \in \mathcal{E}$, the kernel can be made explicit as:

$$P_t(x, C) = \begin{cases} P_t^1(x_1, i_1^{-1}(C)) & \text{if } x = i_1(x_1) \\ P_t^2(x_2, i_2^{-1}(C)) & \text{if } x = i_2(x_2) \end{cases},$$

- for any state $x \in E_1 \uplus E_2$, we set

$$obs(x) = \begin{cases} obs_1(x_1) & \text{if } x = i_1(x_1) \\ obs_2(x_2) & \text{if } x = i_2(x_2) \end{cases},$$

- the set of trajectories on $(E_1 \uplus E_2)_\partial$ is denoted Ω . Note that a trajectory in Ω can switch between E_1 and E_2 . The set Ω is equipped with a σ -algebra \mathcal{G} as is

standard for FD processes. For a state x , we can give explicitly the probability distribution for $B \in \mathcal{G}$:

$$\mathbb{P}^x(B) = \begin{cases} \mathbb{P}_1^{x_1}(B_1) & \text{if } x = i_1(x_1) \\ \mathbb{P}_2^{x_2}(B_2) & \text{if } x = i_2(x_2) \end{cases}$$

where $B_j = \{\omega \in \Omega_j \mid i_j \circ \omega \in B\}$. Note that for any $x \in E_1 \uplus E_2$, for any measurable set $B \subset \{\omega \in \Omega \mid \exists t_1, t_2 \omega(t_1) \in i_1(E_1) \text{ and } \omega(t_2) \in i_2(E_2)\}$, $\mathbb{P}^x(B) = 0$.

We can also make explicit what a bisimulation is in that context (we will omit to mention the inclusions i_1 and i_2 to be readable):

Definition 4.1 Given two FD processes $(E_j, \mathcal{E}_j, (\hat{P}_t^j), (P_t^j), \Omega_j, \mathcal{G}_j, (\mathbb{P}_j^x), obs_j)_{j=1,2}$, a bisimulation between the two FDPs is an equivalence R on $E_1 \uplus E_2$ such that for all xRy ($x \in E_i, y \in E_j$),

(**initiation 1**) $obs_i(x) = obs_j(y)$, and

(**induction 2**) for all measurable R -closed sets B , $\mathbb{P}^x(B \cap \Omega_i) = \mathbb{P}^y(B \cap \Omega_j)$.

This condition can also be stated as follows. For all sets $B_1 \in \mathcal{G}_1$ and $B_2 \in \mathcal{G}$, $\mathbb{P}_i^x(B_i) = \mathbb{P}_j^y(B_j)$ if the two sets satisfy the following condition:

$$\forall \omega_k \in B_k \quad \forall \omega_l \in \Omega_l \quad (\forall t \geq 0 \quad \omega_k(t) R \omega_l(t)) \Rightarrow \omega_l \in B_l$$

In that formulation, $B_k = B \cap \Omega_k$ and the condition states that the set B is R -closed in terms of the sets B_1 and B_2 .

Note that $R \cap (E_j \times E_j)$ is a bisimulation on $(E_j, \mathcal{E}_j, (P_t^j), (\mathbb{P}_j^x))$. To proceed with our cospan idea we need a functional version of bisimulation; we call these Feller-Dynking homomorphisms or FD-homomorphisms for short.

Definition 4.2 A continuous function $f : E \rightarrow E'$ is called a *FD-homomorphism* if it satisfies the following conditions:

- $obs = obs' \circ f$,
- for all $x \in E$ and for all measurable sets $B' \subset \Omega'$, $\mathbb{P}^{f(x)}(B') = \mathbb{P}^x(B)$ where $B = \{\omega \in \Omega \mid f \circ \omega \in B'\}$.

Note that if f and g are FD-homomorphisms, then so is $g \circ f$.

Proposition 4.3 The equivalence relation R defined on $E \uplus E'$ as

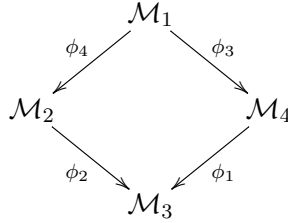
$$R = \{(x, y) \in E \times E \mid f(x) = f(y)\} \cup \{(x, y), (y, x) \mid f(x) = y\}$$

is a bisimulation on E .

The proof can be found in the full version. A nice consequence of this result is that the equivalence relation R defined on E as $R' = \{(x, y) \in E \times E \mid f(x) = f(y)\}$ is a bisimulation on E .

Here is an example with one atomic proposition. Let \mathcal{M}_1 be the standard Brownian motion on the real line with $obs_1(x) = 1$ if and only if $x \in \mathbb{Z}$. Let \mathcal{M}_2 be the reflected Brownian motion on $[0, 1]$ with $obs_2(x) = 1$ if and only if $x = 0$ or 1 . Let \mathcal{M}_3 be the reflected Brownian motion on $[0, \frac{1}{2}]$ with $obs_3(x) = 1$ if and only if $x = 0$. Let \mathcal{M}_4 be the standard Brownian motion on the circle of radius $\frac{1}{2\pi}$ (we will identify points on the circle with the angle wrt the vertical) with $obs_4(x) = 1$ if and only if $x = 0$.

We can define some natural mappings between these processes:



where

$$\begin{aligned}
 \phi_1 : [-\pi, \pi] &\rightarrow [0, 1/2] \\
 \theta &\mapsto |\theta|/2\pi \\
 \phi_2 : [0, 1] &\rightarrow [0, 1/2] \\
 x &\mapsto x && \text{if } x \leq 1/2 \\
 x &\mapsto 1 - x && \text{otherwise} \\
 \phi_3 : \mathbb{R} &\rightarrow [-\pi, \pi] \\
 x &\mapsto 2\pi|x - y| && \text{where } y \in \mathbb{Z} \text{ such that } |x - y| \leq 1/2 \\
 \phi_4 : \mathbb{R} &\rightarrow [0, 1] \\
 x &\mapsto x - 2n && \text{if } \exists n \in \mathbb{Z} \quad 2n \leq x < 2n + 1 \\
 x &\mapsto 2n + 2 - x && \text{if } \exists n \in \mathbb{Z} \quad 2n + 1 \leq x < 2n + 2
 \end{aligned}$$

Note that the condition in the definition of ϕ_3 means that y is the closest integer to x . All these mappings are FD-homomorphisms.

4.2 Definition of cospans

In this section, we are going to omit to mention actions since there are already many indices. However, it is simple to adapt this by adding actions whenever they are needed.

Definition 4.4 A *FD-cospan* is a cospan of FD-homomorphisms.

Theorem 4.5 The category with Feller-Dynkin processes as objects and FD-homomorphisms as morphisms has pushouts.

We omit this proof for the abstract.

This proves that FD-cospan corresponds to an equivalence relation. We have already showed how FD-homomorphisms (and hence FD-cospans) yield bisimula-

tions. The converse is also true as is stated next. The proof can be found in the full version.

Theorem 4.6 *For all bisimulations R , there exists (f, g) a FD-cospan such that*

- *for all $x \in E_1, y \in E_2$, $x R y$ if and only if $f(x) = g(y)$,*
- *for all $x, x' \in E_1$, $x R x'$ if and only if $f(x) = f(x')$, and*
- *for all $y, y' \in E_2$, $y R y'$ if and only if $g(y) = g(y')$.*

We thus have the following correspondance between bisimulation and FD-cospans:

Theorem 4.7 *Two states x and y are bisimilar if and only if there exists a FD-cospan (f, g) such that $f(x) = g(y)$.*

5 Comparison to discrete time bisimulation

The goal of this work is to extend the notion of bisimulation that exists in discrete time to a continuous-time setting. Therefore an important question is the following: do we get back the definition of bisimulation that existed in discrete time when we restrict Feller-Dynkin processes to (some kind of) discrete-time processes?

As we have seen, the case with a single action can be easily adapted to accomodate several actions. We will therefore for the sake of readability not mention actions in this section.

Given an LMP $(X, \Sigma, \tau, (\chi_A)_{A \in AP})$ (there is a single action), we can always view it as a FD process where transitions happen at every time unit. Since the process has to remain memoryless, a state of the FD process is a pair of a state in X and a time explaining how long it has been since the last transition. For trajectories to be cadlag, that time is in $[0, 1)$.

Formally, the state space of the FD procss is (E, \mathcal{E}) with $E = X \times [0, 1)$ and $\mathcal{E} = \Sigma \otimes \mathcal{B}([0, 1))$ and the corresponding kernel is: for all $x \in X$ and $C \in \mathcal{E}$, $t \geq 0$ and $s \in [0, 1)$, $P_t((x, s), C) = \tau_{\lfloor t+s \rfloor}(x, C')$ where $C' = \{z \mid (z, t + s - \lfloor t + s \rfloor) \in C\}$ and for $k \geq 1$,

$$\tau_0(x, C') = \delta_{C'}(x), \quad \tau_1(x, C') = \tau(x, C') \quad \text{and} \quad \tau_{k+1}(x, C') = \int_{y \in X} \tau(x, dy) \tau_k(y, C')$$

We also define $(obs(x, s))_i = \chi_{A_i}(x)$ (where $AP = \{A_1, A_2, \dots\}$).

A DT-bisimulation (see definition 3.1) yields a bisimulation on the corresponding FD process.

Proposition 5.1 *If the equivalence R is a DT-bisimulation, then the relation R' defined as*

$$R' = \{((x, s), (y, s)) \mid s \in [0, 1), x R y\}$$

is a bisimulation in the sense of FD processes.

Going from bisimulation in the sense of FD processes to DT-bisimulation is not as direct and requires some further notions.

Definition 5.2 An equivalence R on the state space of an LMP viewed as a FD process is *time-coherent* if for all x, y in the state space of the LMP and for all $0 \leq t < 1$,

$$(x, t)R(y, t) \Rightarrow \forall s \in [0, 1) (x, s)R(y, s)$$

Given any equivalence R on the state space of an LMP viewed as a FD process, we define its *time-coherent closure* (denoted $\text{time}(R)$) as the smallest time-coherent equivalence containing R .

Proposition 5.3 *If R is a bisimulation on an LMP viewed as a FD process, then so is $\text{time}(R)$.*

Theorem 5.4 *If the equivalence R is a time-coherent bisimulation, then the relation R' defined as*

$$R' = \{(x, y) \mid \exists t \in [0, 1) \text{ such that } ((x, t), (y, t)) \in R\}$$

is a DT-bisimulation.

These results can be summed up in the following theorem relating bisimulation and DT-bisimulation.

Theorem 5.5 *Two states x and y (in the LMP) are DT-bisimilar if and only if for all $t \in [0, 1)$, the states (x, t) and (y, t) (in the Feller-Dynkin process) are bisimilar.*

6 Conclusion

We have given two definitions of bisimulation, one as an equivalence relation and the other as a cospan of morphisms respecting the dynamics of the process. We have also studied many examples; the full version of this conference submission contains many more examples. It would be interesting to know if (and under what conditions) an equivalence satisfying conditions (initiation 2) and (induction 1) is a bisimulation.

However, there are many aspects to explore, as suggested by previous work on step-based systems. First of all is a quantitative description of bisimulation through the definition of some metrics on the state space. This was done in two ways.

- Either through the definition of a set of $[0, 1]$ -valued functions that can be viewed as experiments performed on the system (see [13,30]). We are hoping that this set could be obtained by looking at hitting times and occupation times.
- Or as a fixed-point of some operators on metrics (see [30]). Such a fixed-point metric was defined on jump processes in [19,20]. It seems that some of the details in that work are possibly incorrect, so we hope to fix those details and to adapt similar ideas to our framework.

Another important and interesting question is that of approximations (see [12,10,8]) of our Markov processes. Here we will undoubtedly face new subtleties as we will have to cope with both spatial and temporal limits.

Finally, a fundamental result in this area is the logical characterization of bisimulation [29,21] which was also extended to the probabilistic case [11]. We hope to be able to provide such a logic for continuous-time processes based on the set of $[0, 1]$ -valued functions used to obtain a bisimulation metric. A game interpretation of bisimulation could also be provided [18]. Perhaps some interesting insights could also come from nonstandard analysis [17] where there is also a notion of equivalence but one which is quite different from bisimulation. In that work the notion of adapted spaces is fundamental.

References

- [1] R. Alur, C. Courcoubetis, N. Halbwachs, T.A. Henzinger, P.-H. Ho, X. Nicollin, A. Olivero, J. Sifakis, and S. Yovine. The algorithmic analysis of hybrid systems. *Theoretical Computer Science*, 138:3–34, 1995.
- [2] R. Alur and D. Dill. A theory of timed automata. *Theoretical Computer Science*, 126:183–235, 1994.
- [3] R. Blute, J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. In *Proceedings of the Twelfth IEEE Symposium On Logic In Computer Science, Warsaw, Poland.*, 1997.
- [4] P. Billingsley. *Convergence of Probability Measures*. Wiley Interscience, 2nd edition, 1999.
- [5] Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.
- [6] Adam Bobrowski. *Functional analysis for probability and stochastic processes: an introduction*. Cambridge University Press, 2005.
- [7] Linan Chen, Florence Clerc, and Prakash Panangaden. Bisimulation for Feller-Dynkin Processes. *arXiv e-prints*, page arXiv:1904.00976, Apr 2019.
- [8] Philippe Chaput, Vincent Danos, Prakash Panangaden, and Gordon Plotkin. Approximating Markov processes by averaging. *J. ACM*, 61(1):5:1–5:45, January 2014.
- [9] Vincent Danos, Josée Desharnais, François Laviolette, and Prakash Panangaden. Bisimulation and cocongruence for probabilistic systems. *Information and Computation*, 204(4):503–523, 2006.
- [10] Vincent Danos, Josée Desharnais, and Prakash Panangaden. Conditional expectation and the approximation of labelled Markov processes. In Roberto Amadio and Denis Lugiez, editors, *CONCUR 2003 - Concurrency Theory*, volume 2761 of *Lecture Notes In Computer Science*, pages 477–491. Springer-Verlag, 2003.
- [11] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labeled Markov processes. *Information and Computation*, 179(2):163–193, Dec 2002.
- [12] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximating labeled Markov processes. *Information and Computation*, 184(1):160–200, July 2003.
- [13] Josée Desharnais, Vineet Gupta, Radhakrishnan Jagadeesan, and Prakash Panangaden. A metric for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, June 2004.
- [14] Josée Desharnais and Prakash Panangaden. Continuous stochastic logic characterizes bisimulation for continuous-time Markov processes. *Journal of Logic and Algebraic Programming*, 56:99–115, 2003. Special issue on Probabilistic Techniques for the Design and Analysis of Systems.
- [15] R. M. Dudley. *Real Analysis and Probability*. Wadsworth and Brookes/Cole, 1989.
- [16] A. Einstein. The theory of the brownian movement. *Ann. der Physik*, 17:549, 1905.
- [17] Sergio Fajardo and H Jerome Keisler. *Model theory of stochastic processes*, volume 14 of *Lecture Notes in Logic*. Cambridge University Press, 2017.
- [18] Nathanael Fijalkow, Bartek Klin, and Prakash Panangaden. The expressiveness of probabilistic modal logic revisited. In *Proceedings of the 44th International Colloquium on Automata Languages and Programming*, 2017.

- [19] Vineet Gupta, Radhakrishnan Jagadeesan, and Prakash Panangaden. Approximate reasoning for real-time probabilistic processes. In *The Quantitative Evaluation of Systems, First International Conference QEST04*, pages 304–313. IEEE Press, 2004.
- [20] Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Approximate reasoning for real-time probabilistic processes. *Logical Methods in Computer Science*, 2(1):paper 4, 2006.
- [21] Matthew Hennessy and Robin Milner. On observing nondeterminism and concurrency. In Jaco de Bakker and Jan van Leeuwen, editors, *Automata, Languages and Programming*, volume 85 of *Lecture Notes in Computer Science*, pages 299–309. Springer Berlin / Heidelberg, 1980.
- [22] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science and Business Media, 2012.
- [23] K. G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94:1–28, 1991.
- [24] R. Milner. *A Calculus for Communicating Systems*, volume 92 of *Lecture Notes in Computer Science*. Springer-Verlag, 1980.
- [25] Prakash Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.
- [26] D. Park. Concurrency and automata on infinite sequences. In *Proceedings of the 5th GI Conference on Theoretical Computer Science*, number 104 in *Lecture Notes In Computer Science*, pages 167–183. Springer-Verlag, 1981.
- [27] L. Chris G. Rogers and David Williams. *Diffusions, Markov processes and martingales: Volume 1. Foundations*. Cambridge university press, 2nd edition, 2000.
- [28] Davide Sangiorgi. On the origins of bisimulation and coinduction. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 31(4):15, 2009.
- [29] Johan van Benthem. *Modal correspondence theory*. PhD thesis, University of Amsterdam, 1976.
- [30] Franck van Breugel, Michael Mislove, Joel Ouaknine, and James Worrell. Domain theory, testing and simulation for labelled Markov processes. *Theoretical Computer Science*, 333(1-2):171–197, 2005.
- [31] W. Whitt. *An Introduction to Stochastic-Process Limits and their Applications to Queues*. Springer Series in Operations Research. Springer-Verlag, 2002.