

PAPER

Behavioural equivalences for continuous-time Markov processes

Linan Chen¹, Florence Clerc^{*2,3} , and Prakash Panangaden^{2,3}

¹Department of Mathematics and Statistics, McGill University, Montreal, Canada, ²School of Computer Science, McGill University DEEL Québec, Montreal, Canada, ³School of Computer Science, McGill University Montreal Institute of Learning Algorithms (MILA), Montreal, Canada

*Corresponding author. Email: florence.clerc@mail.mcgill.ca

(Received 27 September 2021; revised 7 November 2022; accepted 19 February 2023; first published online 30 March 2023)

Abstract

Bisimulation is a concept that captures behavioural equivalence of states in a variety of types of transition systems. It has been widely studied in a discrete-time setting. The core of this work is to generalise the discrete-time picture to continuous time by providing a notion of behavioural equivalence for continuous-time Markov processes. In Chen *et al.* [(2019). *Electronic Notes in Theoretical Computer Science* 347 45–63.], we proposed two equivalent definitions of bisimulation for continuous-time stochastic processes where the evolution is a *flow* through time: the first one as an equivalence relation and the second one as a cospan of morphisms. In Chen *et al.* [(2020). *Electronic Notes in Theoretical Computer Science.*], we developed the theory further: we introduced different concepts that correspond to different behavioural equivalences and compared them to bisimulation. In particular, we studied the relation between bisimulation and symmetry groups of the dynamics. We also provided a game interpretation for two of the behavioural equivalences. The present work unifies the cited conference presentations and gives detailed proofs.

Keywords: Stochastic processes; Markov processes; continuous time; bisimulation; diffusion.

1. Introduction

Bisimulation Milner (1980); Park (1981); Sangiorgi (2009) is a fundamental concept in the theory of transition systems capturing a strong notion of behavioural equivalence. The extension to probabilistic systems is due to Larsen and Skou (1991); henceforth, we will simply say “bisimulation” instead of “probabilistic bisimulation”. Bisimulation has been studied for discrete-time systems where transitions happen as steps, both on discrete Larsen and Skou (1991) and continuous state spaces Blute (1997); Desharnais *et al.* (2002); Panangaden (2009). 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 must hold “now” (in the states being compared, this is an initiation condition), and then, one says that the relation is preserved in the *next* step (this is an induction condition).

There is a vast range of systems that involve 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 (1905) on Brownian motion.

This work focuses on suggesting a notion of bisimulation to stochastic systems with true continuous-time evolution based on previous works on discrete-time systems. Some work had previously been done in what are called continuous-time systems Desharnais and Panangaden (2003), 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, e.g., Rogers and Williams (2000); Whitt (2002)), a phrase that better captures the true nature of such processes.

We focused on a class of systems called Feller-Dynkin processes for which a good mathematical theory exists. These systems are Markov processes defined on continuous state spaces and with continuous-time evolution. Such systems encompass Brownian motion and its many variants.

A first thought might be to take the well-known theory of discrete-time processes, like random walks for example and to develop a theory of continuous time by taking some kind of limit of step sizes going to zero. However, it is not enough to take limits of what happens in discrete time. Strange phenomena occur when shifting from discrete steps to continuous time. It is necessary to talk about trajectories instead.

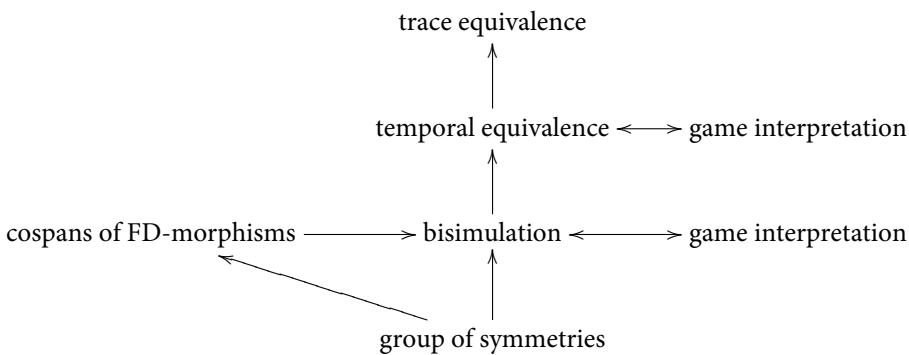
We explored several notions of behavioural equivalence for such continuous-time processes. The strongest notion is one that captures the symmetries of the system: a group of symmetries is a set of bijections that leave the dynamics of the system unchanged.

Bisimulation as a discrete-time notion has two conditions: an initiation one and a (co)induction one. Depending on whether we extend the initiation or the induction condition using trajectories, we get two notions of behavioural equivalences in continuous time: temporal equivalence and bisimulation, respectively. Temporal equivalence can be summed up as trace equivalence with some additional step-like constraints. A bisimulation is a temporal equivalence, but it is still an open question whether these notions are equivalent. We have shown that trace equivalence is a strictly weaker notion than temporal equivalence (and hence the other notions).

It is possible to view discrete-time systems as continuous-time systems by artificially adding a timer and having transitions occur every unit of time. This has allowed us to compare our notions to the previous definition of bisimulation for discrete-time systems. This shows that temporal equivalence indeed generalises bisimulation to continuous-time systems.

Finally, we give two game interpretations: one for bisimulation and one for temporal equivalence. They closely mirror the one provided in Fijalkow *et al.* (2017). The game for bisimulation also emphasises the importance of trajectories for the study of behavioural equivalences in continuous time.

The relations between these different behavioural equivalences can be displayed as follows:



This paper is organised as follows. In Section 2, we quickly go over mathematical prerequisites and bisimulation in discrete time. In Section 3, we go over the definition of Feller-Dynkin

processes and an example of such process: Brownian motion. We also provide some tools that will be useful later on when studying examples: hitting times. In Section 4, we define the various behavioural equivalences that we have come up with in this work and we explain the relations between them. We then compare them to discrete-time bisimulation. Those continuous-time notions of behavioural equivalence are illustrated in Section 5 with various examples. In Section 6, we go over a more categorical approach to behavioural equivalences. Finally, we provide a game interpretation for bisimulation and temporal equivalence in Section 7.

2. Background

2.1 Mathematical background

We assume the reader to be familiar with the basic notions of topology (topology, continuous function, compact, locally compact, compactification, σ -compact, Hausdorff), Polish spaces, Banach spaces, measure theory (σ -algebra, measurable functions, measure, probability space). For a detailed introduction, we refer the reader to Dudley (1989); Panangaden (2009); Billingsley (2008).

2.2 Discrete-time systems

Markov processes are stochastic processes where what “happens next” depends only on the current state and not the full history of how that state was reached. They are used in many fields as statistical models of real-world problems and have therefore been extensively studied.

In some cases, the environment or a user may influence how things will evolve, for instance, by pressing a key on a keyboard. This is modelled by using a countable set *Act* of *actions* that determine at each step which stochastic process is used.

Definition 1. A labelled Markov process (LMP) is a tuple $(S, \Sigma, (\tau_a)_{a \in \text{Act}})$, where (S, Σ) is a measurable space called the state space and for all $a \in \text{Act}$, $\tau_a : S \times \Sigma \rightarrow [0, 1]$ is a Markov kernel, that is

- for all states x in S , $\tau_a(x, \cdot)$ is a subprobability measure, and
- for all C in Σ , $\tau_a(\cdot, C)$ is Σ -measurable.

We will also assume that an LMP is equipped with a measurable function $\text{obs} : S \rightarrow 2^{AP}$, where AP denotes a set of *atomic propositions* and 2^{AP} is equipped with the discrete σ -algebra. The function obs is useful to isolate areas of the state space. For instance, if the state space is a set of temperatures, there could be atomic propositions “freezing” and “heat wave” isolating, respectively, the subsets of the state space $(-273, 0)$ and $(30, +\infty)$.

Definition 2. An equivalence R on S is a bisimulation if for every states $x, y \in S$, if $x R y$, then the two following conditions are satisfied:

- $\text{obs}(x) = \text{obs}(y)$,
- for every measurable and R -closed set C (meaning that for every $z R z'$, $z \in C$ if and only if $z' \in C$) and every action a , $\tau_a(x, C) = \tau_a(y, C)$

Two states are bisimilar if there exists a bisimulation that relates them.

There is a greatest bisimulation which corresponds to the equivalence “are bisimilar”. This greatest bisimulation is called “bisimilarity”.

Example 3. Random walk is a very standard example. The most basic version of it takes place on the set of signed integers \mathbb{Z} , and the intuition is that at each step, the process goes either to the left or to the right in an unbiased manner. More formally, the corresponding Markov kernel is the following for all integers n, m :

$$\tau(n, m) = \begin{cases} \frac{1}{2} & \text{if } m = n - 1 \text{ or } n + 1, \\ 0 & \text{otherwise.} \end{cases}$$

The state space \mathbb{Z} can be equipped with the function $obs(x) = 0$ if $x = 0$ and $obs(x) = 1$ if $x \neq 0$. In this case, two states n and m are bisimilar if and only if $|n| = |m|$.

In Desharnais *et al.* (2002), a more categorical approach to bisimulation is also provided.

Definition 4. Given two LMPs $\mathcal{S} = (S, \Sigma, (\tau_a)_{a \in Act})$ and $\mathcal{S}' = (S', \Sigma', (\tau'_a)_{a \in Act})$, a function $f : \mathcal{S} \rightarrow \mathcal{S}'$ is a zigzag if $f : (S, \Sigma) \rightarrow (S', \Sigma')$ is surjective and measurable and for all action a , state $s \in S$ and $B' \in \Sigma'$, $\tau_a(s, f^{-1}(B')) = \tau'_a(f(s), B')$.

It is quite easy to see that for a state s and a zigzag f , the states s and $f(s)$ are bisimilar. In Desharnais *et al.* (2002), the authors compare bisimulation to spans of zigzags and showed that the two notions coincided if the state space is analytic.

It was shown that the notions induced respectively by spans of zigzags and by cospans of zigzags were equivalent on analytic Borel spaces Doberkat (2005) and on separable universally measurable spaces (isomorphic to a universally measurable subset of a separable completely metrisable space) Pahl and Sánchez Terraf (2020).

3. Continuous-Time Systems

Before looking into similarities of behaviours of continuous-time Markov processes, one needs to have a definition of what a continuous-time Markov process is. We decided that the best trade-off for this study was to restrict to Feller-Dynkin processes. Such processes have some regularity on their trajectories which gives us tools to handle them. Moreover, this family of processes is general enough to encompass most systems that might be of interest.

3.1 Definition of Feller-Dynkin processes

We will quickly review the theory of continuous-time processes on continuous state space; much of this material is adapted from Rogers and Williams (2000), and we use their notations. Another useful source is Bobrowski (2005).

Definition 5. A filtration on a measurable space (X, Σ) is a time-indexed, non-decreasing family $(\mathcal{F}_t)_{t \geq 0}$ of sub- σ -algebras of Σ , i.e. $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \Sigma$ 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 6. A stochastic process is a collection of random variables $(Y_t)_{0 \leq t < \infty}$ on a measurable space (Ω, Σ_Ω) that take values in a second measurable space (X, Σ_X) called the state space. More explicitly, we have a time-indexed family of random variables $Y_t : (\Omega, \Sigma_\Omega) \rightarrow (X, \Sigma_X)$ that are all $(\Sigma_\Omega, \Sigma_X)$ -measurable (i.e. for every set $B \in \Sigma_X$, $Y_t^{-1}(B) \in \Sigma_\Omega$). Given a filtration $(\mathcal{F}_t)_{t \geq 0}$ on (Ω, Σ_Ω) , the stochastic process $(Y_t)_{t \geq 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ if for each $t \geq 0$, Y_t is \mathcal{F}_t -measurable.

Define the filtration $(\mathcal{G}_t)_{t \geq 0}$ as follows: for every $t \geq 0$, the σ -algebra \mathcal{G}_t is the one generated by all the random variables $\{Y_s | s \leq t\}$, that is

$$\mathcal{G}_t = \sigma(\{Y_s^{-1}(A) \mid s \leq t \text{ and } A \in \Sigma_X\})$$

This filtration $(\mathcal{G}_t)_{t \geq 0}$ is also referred to as the *natural filtration* associated with the stochastic process $(Y_t)_{t \geq 0}$. A stochastic process is always adapted to its natural filtration.

Let E be a locally compact, Hausdorff space with countable base. These standard topological hypotheses allow us to perform the one-point compactification of the set E by adding the absorbing state ∂ . We write E_∂ for the corresponding set: $E_\partial = E \uplus \{\partial\}$. We also equip the set E with its Borel σ -algebra $\mathcal{B}(E)$ that we will denote \mathcal{E} . The previous topological hypotheses also imply that E is σ -compact and Polish.

Definition 7. A semigroup of operators on any Banach space X is a family of linear continuous (bounded) operators $\mathcal{P}_t : X \rightarrow X$ indexed by $t \in \mathbb{R}_{\geq 0}$ such that

$$\forall s, t \geq 0, \mathcal{P}_s \circ \mathcal{P}_t = \mathcal{P}_{s+t}$$

and

$$\mathcal{P}_0 = I \quad (\text{the identity}).$$

The first equation above is called the semigroup property. The operators in a semigroup are continuous operators. Moreover, there is a useful continuity property with respect to time t of the semigroup as a whole.

Definition 8. For X a Banach space, we say that a semigroup $\mathcal{P}_t : X \rightarrow X$ is strongly continuous if

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

Example 9. Let us give a simple example on the simple Banach space $\mathbb{R} \rightarrow \mathbb{R}$. For instance, $\mathcal{P}_t f = f + t$ is a strongly continuous semigroup modelling some deterministic drift at a constant rate. A counterexample would be $\mathcal{P}_t f = 5t^2 + f$ (modelling an example of a free fall): this family of operators does not satisfy $\mathcal{P}_s \circ \mathcal{P}_t = \mathcal{P}_{s+t}$.

What the semigroup property expresses is that we do not need to understand the past (what happens before time t) in order to compute the future (what happens after some additional time s , so at time $t + s$) as long as we know the present (at time t). However, it is simple to have such a semigroup describe a free fall as long as we are working on $\mathbb{R}^2 \rightarrow \mathbb{R}$, as physics teaches us. The first coordinate represents the position and the second the velocity. In this case, we have that

$$(\mathcal{P}_t f)(x, v) = f(5t^2 + vt + x, 10t + v)$$

We say that a continuous real-valued function f on E “vanishes at infinity” if for every $\varepsilon > 0$ there is a compact subset $K \subseteq E$ such that for every $x \in E \setminus K$, we have $|f(x)| \leq \varepsilon$. To give an intuition, if E is the real line, this means that $\lim_{x \rightarrow \pm\infty} f(x) = 0$. The space $C_0(E)$ of continuous real-valued functions that vanish at infinity is a Banach space with the sup norm.

Definition 10. A Feller-Dynkin (FD) semigroup is a strongly continuous semigroup $(\hat{P}_t)_{t \geq 0}$ of linear operators on $C_0(E)$ satisfying the additional condition:

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

Under these conditions, strong continuity is equivalent to the apparently weaker condition (see Lemma III. 6.7 in Rogers and Williams (2000) for the proof):

$$\forall f \in C_0(E) \quad \forall x \in E, \quad \lim_{t \downarrow 0} (\hat{P}_t f)(x) = f(x)$$

The following important proposition relates these FD-semigroups with Markov kernels which allows one to see the connection with more familiar probabilistic transition systems. We provide a detailed proof of the following proposition in Appendix A.

Proposition 11. *Given an FD-semigroup $(\hat{P}_t)_{t \geq 0}$ on $C_0(E)$, 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 12. *A function $\omega : [0, \infty) \rightarrow E_\partial$ is called cadlag¹ if for every $t \geq 0$,*

$$\lim_{s > t, s \rightarrow t} \omega(s) = \omega(t) \quad \text{and} \quad \omega(t-) := \lim_{s < t, s \rightarrow t} \omega(s) \text{ exists}$$

We define a trajectory ω on E_∂ to be a cadlag function $[0, \infty) \rightarrow E_\partial$ such that if either $\omega(t-) := \lim_{s < t, s \rightarrow t} \omega(s) = \partial$ 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 $\omega(\infty) = \partial$.

As an intuition, a cadlag function ω is an ‘‘almost continuous’’ function with some jumping in a reasonable fashion.

The intuition behind the additional condition of a trajectory is that once a trajectory has reached the terminal state ∂ , it stays in that state, and similarly if a trajectory ‘‘should’’ have reached the terminal state ∂ , then there cannot be a jump to avoid ∂ and once it is in ∂ , it stays there.

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

Definition 13. *The canonical FD-process associated with the FD-semigroup $(\hat{P}_t)_{t \geq 0}$ is*

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

where

- the random variable $X_t : \Omega \rightarrow E_\partial$ is defined as $X_t(\omega) = \omega(t)$
- $\mathcal{G} = \sigma(X_s \mid 0 \leq s < \infty)$ ², $\mathcal{G}_t = \sigma(X_s \mid 0 \leq s \leq t)$
- given any probability measure μ on E_∂ , by the Daniell-Kolmogorov 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_∂ ,

$$\mathbb{P}^\mu(X_0 \in dx_0, X_{t_1} \in dx_1, \dots, X_{t_n} \in dx_n) = \mu(dx_0) P_{t_1}^{+\partial}(x_0, dx_1) \dots P_{t_n - t_{n-1}}^{+\partial}(x_{n-1}, dx_n)$$
³

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

The distribution \mathbb{P}^x is a measure on the space of trajectories for a system started at the point x .

Remark 14. Let us give an intuition for this definition based on discrete-time Markov chain. There are several ways to describe a Markov chain. One is to define a Markov kernel on the state space. By nesting Markov kernels, we can then compute the probability of jumping from a certain state to a certain set in a certain amount of steps. However, one may also directly define the step-indexed family of random variables describing where the process is at each step.

Here, the Markov kernels correspond to the family of subkernels $(P_t)_{t \geq 0}$ and the step-indexed family of random variables corresponds to the time-indexed family $(X_t)_{t \geq 0}$.

We have constructed here one process which we called the canonical FD-process. Note that there could be other tuples

$$(\Omega', \mathcal{G}', (\mathcal{G}'_t)_{t \geq 0}, (Y_t)_{0 \leq t < \infty}, (\mathbb{Q}^x)_{x \in E_\partial})$$

satisfying the same conditions except that the space Ω' would not be the space of trajectories.

In order to bring the FD-processes more in line with the kind of transition systems that have hitherto been studied in the computer science literature, we introduce a countable set of atomic propositions AP and such an FD-process is equipped with a measurable function $obs : E \rightarrow 2^{AP}$. This function is extended to a function $obs : E_\partial \rightarrow 2^{AP} \uplus \{\partial\}$ by setting $obs(\partial) = \partial$.

We further request that the function obs satisfies the following hypothesis: for each atomic proposition A , the atomic proposition partitions the state space into two subsets: one where A is satisfied ($obs^{-1}(S_A)$ where $S_A = \{a : AP \rightarrow \{0, 1\} \mid a(A) = 1\}$) and another one where it is not satisfied ($obs^{-1}(S_A)$ where $S_A = \{a : AP \rightarrow \{0, 1\} \mid a(A) = 0\}$). We assume that one of these spaces is open (and the other is closed). Note that it can be that for one atomic proposition A_0 , the set where it is satisfied is open and for a different atomic proposition A_1 , the set where it is satisfied is closed. This hypothesis will be useful in Section 6.

3.2 A classic example of FD-process: Brownian motion

Brownian motion is a stochastic process first introduced to describe the irregular motion of a particle being buffeted by invisible molecules. Now its range of applicability extends far beyond its initial application Karatzas and Shreve (2012): it is used in finance, in physics, in biology or as a way to model noise to mention only a few domains.

We refer the reader to Karatzas and Shreve (2012) for the complete definition. A standard one-dimensional Brownian motion is a Markov process on the real line \mathbb{R} that is invariant under reflexion around any real value and under translation by any real value.

The following fundamental formula is useful in computations: 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 $x \in \mathbb{R}$,

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

Remark 15. A useful intuition that one can have about Brownian motion is that it is the limit of random walk when the time between jumps and the distance between states is taken to 0.

There are many variants of Brownian motions that we will use and describe in detail in further examples. For instance, a drift may be added. Another classic variant is when “walls” are added: when the process hits a wall, it can either vanish (boundary with absorption) or bounce back (boundary with reflection). These correspond to the variants of random walk that we discussed in Example 3.

3.3 Hitting times and stopping times

We will use the notions of hitting times and stopping times later. We will follow the definitions of Karatzas and Shreve (2012).

Throughout this section, we assume that $(X_t)_{t \geq 0}$ is a stochastic process on (Ω, \mathcal{F}) such that $(X_t)_{t \geq 0}$ takes values in a state space $(E, \mathcal{B}(E))$, has right-continuous paths (for every time t , $\omega(t) = \lim_{s \rightarrow t, s > t} \omega(s)$) and is adapted to a filtration $(\mathcal{F}_t)_{t \geq 0}$.

Definition 16. Given a measurable set $C \in \mathcal{B}(E)$, the hitting time is the random time

$$T_C(\omega) = \inf\{t \geq 0 \mid X_t(\omega) \in C\}.$$

Intuitively, $T_C(\omega)$ corresponds to the first time when the trajectory ω “touches” the set C . The verb “touches” in previous sentence is important: this time is obtained as an infimum, not a minimum.

Definition 17. A random time T is a stopping time of the filtration $(\mathcal{F}_t)_{t \geq 0}$ if for every time t , the event $\{T \leq t\}$ belongs to the σ -field \mathcal{F}_t . A random time T is an optional time of the filtration $(\mathcal{F}_t)_{t \geq 0}$ if for every time t , $\{T < t\} \in \mathcal{F}_t$.

Recall that the intuition behind the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ is that the information about the process up to time t is stored in the σ -algebra \mathcal{F}_t . So intuitively, the random time T is a stopping time (resp. an optional time) if and only if there is enough information gathered at time t to decide if the event described by the random T is less (\leq and $<$, respectively) than t . We provide examples and counterexamples to this definition in the remaining of this section.

A stopping time is also an optional time but the converse may not be true. The next lemma may provide some additional intuition and some examples behind those notions.

Lemma 18. If the set C is open, T_C is an optional time.

If the set C is closed and the paths of the process X are continuous, then T_C is a stopping time.

Let us give an example of a random time that is neither a stopping time nor an optional time. Given a continuous trajectory $\omega : [0, +\infty) \rightarrow \mathbb{R}$, we define

$$T_{\max}(\omega) = \inf\{t \mid \omega(t) = \max \omega\}$$

So T_{\max} is the time that ω reaches its maximum value. In order to know if $T_{\max}(\omega) < t$, one would need to know the maximum value attained by the trajectory ω . More generally, any such random time that requires to know about the future behaviour of a trajectory is neither a stopping time nor an optional time.

Remark 19. For Brownian motion, we write T_x instead of $T_{\{x\}}$ for $x \in \mathbb{R}$.

4. Behavioural Equivalences for Continuous-Time Systems

4.1 Naive definition

A naive extension of the discrete-time definition of bisimulation where the induction condition would only require that steps of time t (for every such $t \geq 0$) preserve the relation does not work in continuous time.

Indeed, consider Brownian motion on the real line with a single atomic proposition such that $obs(x) = 0$ if and only if $x = 0$. The equivalence $x R y$ if and only if $x = y = 0$ or $xy \neq 0$ would satisfy the hypothesis of this naive extension of bisimulation. This would mean that any two non-zero

states, regardless of their distance to 0 would be considered equivalent. This does not extend the intuition that we should get from random walk (see example 3) nicely into continuous time.

As it turns out, the problem lies with considering single time-steps since for every state $z \neq 0$ and every time $t \geq 0$, $P_t(z, \{0\}) = 0$. Had we considered instead probabilities of reaching the state 0 over an interval of time ($\mathbb{P}^z(T_0 < t)$ for instance), we would not have had the same equivalence.

This is why we deal with trajectories instead of steps.

4.2 Closedness for sets of trajectories

Throughout the remaining of Section 4, we fix an FD-process as in Section 3.1.

Since the conditions for being a behavioural equivalence have to be stated for trajectories, we have to extend the notion of R -closedness of a set of states to a set of trajectories. This is what is done in the following definition.

Definition 20. *Given an equivalence R on E extended to E_∂ by setting $\partial R \partial$, a set B of trajectories is time- R -closed if for every trajectories ω and ω' such that for every time $t \geq 0$, $\omega(t) R \omega'(t)$, $\omega \in B$ if and only if $\omega' \in B$.*

Definition 21. *A set B of trajectories is called time-obs-closed if the following three conditions are satisfied:*

- *it is measurable (cf Definition 13 where we defined the σ -algebra on the space of trajectories of a canonical FD-process),*
- *for every trajectories ω and ω' such that $obs \circ \omega = obs \circ \omega'$, $\omega \in B$ if and only if $\omega' \in B$,*
- *define the map $\Pi : \Omega \rightarrow (2^{AP})^{\mathbb{N}}$ by $\Pi(\omega)(n) = obs(\omega(n))$ (here n is a time, hence $\omega(n)$ is a state in E), the set $\Pi(B)$ is measurable. Recall that 2^{AP} is equipped with the discrete σ -algebra. The σ -algebra on $(2^{AP})^{\mathbb{N}}$ is generated by the family of sets $(S_{k,A})_{k \in \mathbb{N}, A \subset 2^{AP}}$ where $S_{k,A}$ is defined as $\{f : \mathbb{N} \rightarrow 2^{AP} \mid f(k) \in A\}$.*

Remark 22. The last condition with the map Π may look unnatural; it was originally introduced as a technical trick to complete the correspondence with discrete time. However, it can be motivated as follows: the function obs determines what an external observer may view of the state space and hence the position of the process. The intuition behind the map Π is hence “given a trajectory ω , what does an outside user see of the state of the system at fixed times (every unit of time)?”.

That last condition deals with the question “how often should someone be allowed to observe the system?” and answers that an outside observer should at least be allowed to probe the process once every unit of time.

4.3 Bisimulation

There are two conditions (initiation and induction conditions) that can be modified to account for trajectories. Depending on which one is adapted, we get either temporal equivalence or bisimulation. In Chen *et al.* (2019), we introduced only the notion of “bisimulation” where the induction condition is adapted to trajectories.

Definition 23. *A bisimulation is an equivalence relation R on E such that for all $x, y \in E$, if $x R y$, then*

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

(induction) for all measurable time- R -closed sets B , $\mathbb{P}^x(B) = \mathbb{P}^y(B)$.

Remark 24. Even though condition (induction) is called an induction, it really is a coinduction condition.

Proposition 25. *There is a greatest bisimulation called bisimilarity. It is the union of all bisimulations.*

4.4 Temporal equivalence

We later came up with other notions of behavioural equivalences, and in Chen *et al.* (2020), we proposed three other notions of behavioural equivalences. It is quite likely that there are other interesting notions to be studied in future works.

Definition 26. *A temporal equivalence is an equivalence relation R on E such that for all $x, y \in E$, if $x R y$, then*

- (initiation) *for all time-obs-closed sets B , $\mathbb{P}^x(B) = \mathbb{P}^y(B)$, and*
- (induction) *for all measurable R -closed sets C , for all times t , $P_t(x, C) = P_t(y, C)$.*

Proposition 27. *There is a greatest temporal equivalence which is the union of all temporal equivalences.*

Definition 28. *Two states that are related by a temporal equivalence are called temporally equivalent.*

4.5 Trace equivalence

Another well-known concept is that of trace equivalence which corresponds to the initiation condition of temporal equivalence. Temporal equivalence can thus be viewed as trace equivalence which additionally accounts for step-like branching.

Definition 29. *Two states are trace equivalent if and only if for all time-obs-closed sets B , $\mathbb{P}^x(B) = \mathbb{P}^y(B)$.*

We will later see in details how it relates to the standard notion of trace equivalence for discrete-time processes (see Section 4.8).

4.6 Relation between these equivalences

We are going to show that the weakest equivalence is trace equivalence, and the strongest is bisimulation as hinted.

Theorem 30. *A bisimulation is also a temporal equivalence. If two states are temporally equivalent, then they are trace equivalent.*

Proof. Let R be a bisimulation and consider two states x and y such that $x R y$.

Consider a time-obs-closed set B . Then, it is also time- R -closed. Using the induction condition of bisimulation, we get that $\mathbb{P}^x(B) = \mathbb{P}^y(B)$.

Consider a measurable R -closed set C and a time t . The set $X_t^{-1}(C) = \{\omega \mid \omega(t) \in C\}$ is measurable and time- R -closed. We can then apply the induction condition, and we get

$$P_t(x, C) = \mathbb{P}^x(X_t^{-1}(C)) = \mathbb{P}^y(X_t^{-1}(C)) = P_t(y, C)$$

This concludes the proof that R is a temporal equivalence.

The second part of the lemma follows directly from the initiation condition of a temporal equivalence: this is precisely trace equivalence. \square

We provide in Section 5.1.2 an example where the greatest temporal equivalence is strictly greater than trace equivalence. It is still an open question as to whether bisimulation and temporal equivalence are the same notions or if they coincide only for a class of processes. We refer the reader to Section 5 for examples of these equivalences.

4.7 Symmetries of systems

An interesting notion that appears when looking at examples is that of symmetries of a system. It was the third notion introduced in Chen *et al.* (2020).

Given a function $h : E_\partial \rightarrow E_\partial$, we define $h_* : \Omega \rightarrow \Omega$ by $h_*(\omega) = h \circ \omega$.

A group of symmetries is a set of bijections on the state space that respect the dynamics of the FD-process.

Definition 31. A group of symmetries is a group (closed under inverse and composition) \mathcal{H} of homeomorphisms on the state space E extended to E_∂ by setting $h(\partial) = \partial$ for every $h \in \mathcal{H}$ such that

- for all $h \in \mathcal{H}$, $obs \circ h = obs$, and
- for all $x \in E_\partial$, for all $f \in \mathcal{H}$ and for all measurable sets B such that for all $h \in \mathcal{H}$, $h_*(B) = B$,

$$\mathbb{P}^x(B) = \mathbb{P}^{f(x)}(B).$$

Definition 32. Given a group of symmetries \mathcal{H} on E , we denote $R_{\mathcal{H}}$ the equivalence defined on E as follows: $x R_{\mathcal{H}} y$ if and only if there exists $h \in \mathcal{H}$ such that $h(x) = y$. The fact that \mathcal{H} is a group guarantees that $R_{\mathcal{H}}$ is an equivalence.

One of the requirements for being a group of symmetries is to be closed under inverse and composition. This condition is useful for getting an equivalence on the state space; however, it is usually easier (if possible) to view a group of symmetries as generated by a set of homeomorphisms.

Lemma 33. Consider a set of homeomorphisms \mathcal{H}_{gen} on the state space E and define \mathcal{H} as the closure under inverse and composition of the set \mathcal{H}_{gen} . Assume that the set \mathcal{H}_{gen} satisfies the following conditions:

- for all $f \in \mathcal{H}_{gen}$, $obs \circ f = obs$, and
- for all measurable sets B such that for all $f \in \mathcal{H}_{gen}$, $f_*(B) = B$, for all $x \in E_\partial$ and for all $g \in \mathcal{H}_{gen}$, $\mathbb{P}^x(B) = \mathbb{P}^{g(x)}(B)$.

Then the set \mathcal{H} is a group of symmetries.

Theorem 34. Given a group of symmetries \mathcal{H} , the equivalence $R_{\mathcal{H}}$ is a bisimulation.

The proof of Theorem 34 requires two additional lemmas. We will omit the proof of the first one since it is straightforward.

Lemma 35. Consider a time- $R_{\mathcal{H}}$ -closed set B . Then for every $f \in \mathcal{H}$, $f_*(B) \subseteq B$.

Lemma 36. Given a group of symmetries \mathcal{H} , if a set B is time- $R_{\mathcal{H}}$ -closed, then for every $h \in \mathcal{H}$, $h_*(B) = B$.

Proof. First, using lemma 35, $h_*(B) \subseteq B$.

To prove the converse implication, consider $\omega \in B$ and define ω' as $\omega' = (h^{-1})_*(\omega)$. Note that for all times t , $h(\omega'(t)) = \omega(t)$, that is $\omega(t) R_{\mathcal{H}} \omega'(t)$.

Since B is time- $R_{\mathcal{H}}$ -closed, $\omega' \in B$. We have defined $\omega' = (h^{-1})_*(\omega) = h^{-1} \circ \omega$. A direct consequence is that $\omega = h \circ \omega' = h_*(\omega')$, and therefore, $\omega \in h_*(B)$. \square

Proof of Theorem 34. Consider two equivalent states $x R_{\mathcal{H}} y$, that is there exists $h \in \mathcal{H}$ such that $h(x) = y$.

First, $obs(x) = obs \circ h(x) = obs(y)$.

Second, let us consider a measurable, time- $R_{\mathcal{H}}$ -closed B . Using Lemma 36, we know that for every $f \in \mathcal{H}$, $f_*(B) = B$, and hence, $\mathbb{P}^x(B) = \mathbb{P}^{h(x)}(B) = \mathbb{P}^y(B)$, which concludes the proof. \square

Remark 37. Given a group of symmetries \mathcal{H} , a set C of states is $R_{\mathcal{H}}$ -closed if and only if for every $h \in \mathcal{H}$, $h(C) = C$. It is tempting to find a nice characterisation of time- $R_{\mathcal{H}}$ -closed sets too; however in the case of trajectories, this is much more complicated. Lemma 36 showed that if a set B is time- $R_{\mathcal{H}}$ -closed, then for every $h \in \mathcal{H}$, $h_*(B) = B$ (which we used in the proofs of later examples) but this is no longer an equivalence.

4.8 Comparison to discrete time

Our work aims at extending the notion of bisimulation from discrete time to continuous time. It is possible to view a discrete-time process as a continuous-time one as explained below. It is important to check that our notions encompass the pre-existing notion of bisimulation in discrete time.

It is common in discrete time to consider several actions. Everything that was exposed for continuous time can easily be adapted to accommodate several actions. However, we will not mention actions in this section for the sake of readability.

Consider an LMP (S, Σ, τ, obs) where Σ is the Borel-algebra generated by a topology on S . We also assume that this LMP has at most finitely many atomic propositions. We can always view it as an FD-process where transitions happen at every unit of time. Since the corresponding continuous-time process has to satisfy the Markov property, it cannot keep track of how long it has spent in a state of the LMP. This is why we need to include a “timer” into states. A state in the corresponding FD-process is thus a pair of a state in S and a time explaining how long it has been since the last transition. That time is in $[0, 1)$, the right open bound is in order for trajectories to be cadlag.

Formally, the state space of the FD-process is (E, \mathcal{E}) where the space is defined as $E = S \times [0, 1)$ and is equipped with the product topology and the corresponding Borel σ -algebra $\mathcal{E} = \Sigma \otimes \mathcal{B}([0, 1))$. The corresponding kernel is defined for all $x \in S$ and $C \in \mathcal{E}$, $t \geq 0$ and $s \in [0, 1)$ as $P_t((x, s), C) = \tau_{\lfloor t+s \rfloor}(x, C')$ where $C' = \{z \mid (z, t + s - \lfloor t + s \rfloor) \in C\}$ and for $k \geq 1$,

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

We also define $obs(x, s) = obs(x)$.

This gives us a way to view an LMP as a continuous-time process and thus to compare the definition of bisimulation on the original LMP to the behavioural equivalences for the corresponding continuous-time process. In order to make clear when we talk about discrete-time bisimulation

and in order to avoid confusion for the reader, we will write “DT-bisimulation” for bisimulation in discrete time as was defined in Section 2.2.

Lemma 38. *If two states x and y of an LMP (S, Σ, τ, obs) are DT-bisimilar, then for all $n \geq 1$, for all measurable R -closed (where R is the greatest DT-bisimulation) sets A_1, \dots, A_n ,*

$$\begin{aligned} & \int_{x_1 \in A_1} \dots \int_{x_n \in A_n} \tau(x, dx_1)\tau(x_1, dx_2)\dots\tau(x_{n-1}, dx_n) \\ &= \int_{x_1 \in A_1} \dots \int_{x_n \in A_n} \tau(y, dx_1)\tau(x_1, dx_2)\dots\tau(x_{n-1}, dx_n) \end{aligned}$$

Proof. Denote $\pi_R : S \rightarrow S/R$ the corresponding quotient. We equip the quotient space S/R with the largest σ -algebra which makes the map π_R measurable. We can also define the function

$$\bar{\tau}(\pi_R(x), A) = \tau(x, \pi_R^{-1}(A))$$

Note that the choice of x (within an R -class) does not change the right term since R is a DT-bisimulation and $\pi_R^{-1}(A)$ is an R -closed set: you can replace x by x' as long as $x R x'$. A sequence of changes of variables yields:

$$\begin{aligned} & \int_{x_1 \in A_1} \dots \int_{x_n \in A_n} \tau(x, dx_1)\tau(x_1, dx_2)\dots\tau(x_{n-1}, dx_n) \\ &= \int_{x_1 \in A_1} \dots \int_{x_{n-1} \in A_{n-1}} \int_{y_n \in A_n/R} \tau(x, dx_1)\tau(x_1, dx_2)\dots\tau(x_{n-2}, dx_{n-1})\tau(x_{n-1}, \pi_R^{-1}(dy_n)) \\ &= \int_{x_1 \in A_1} \dots \int_{x_{n-1} \in A_{n-1}} \int_{y_n \in A_n/R} \tau(x, dx_1)\tau(x_1, dx_2)\dots\tau(x_{n-2}, dx_{n-1})\bar{\tau}(\pi_R(x_{n-1}), dy_n) \\ &= \int_{x_1 \in A_1} \dots \int_{x_{n-2} \in A_{n-2}} \int_{y_{n-1} \in A_{n-1}/R} \int_{y_n \in A_n/R} \tau(x, dx_1)\tau(x_1, dx_2)\dots\tau(x_{n-2}, \pi_R^{-1}(dy_{n-1}))\bar{\tau}(y_{n-1}, dy_n) \\ &= \int_{x_1 \in A_1} \dots \int_{x_{n-2} \in A_{n-2}} \int_{y_{n-1} \in A_{n-1}/R} \int_{y_n \in A_n/R} \tau(x, dx_1)\tau(x_1, dx_2)\dots\bar{\tau}(\pi_R(x_{n-2}), dy_{n-1})\bar{\tau}(y_{n-1}, dy_n) \\ &= \int_{y_1 \in A_1/R} \dots \int_{y_n \in A_n/R} \tau(x, \pi_R^{-1}(dy_1))\bar{\tau}(y_1, dy_2)\dots\bar{\tau}(y_{n-1}, dy_n) \end{aligned}$$

And since the two measures $\tau(x, \pi_R^{-1}(\cdot))$ and $\tau(y, \pi_R^{-1}(\cdot))$ are equal as R is a DT-bisimulation, this concludes the proof. □

It is possible to define the notion of trajectories in the LMP and that of trace equivalence just as we did in the case of FD-processes. A *trajectory* is a function $\omega : \mathbb{N} \rightarrow S \cup \{\partial\}$ such that if $\omega(n) = \partial$, then for every $k \geq n$, $\omega(k) = \partial$. Similarly to what was done in Section 3.1, for a state x we can define the probability distribution \mathbb{P}^x on the set of trajectories that extends the finite-time distributions using the Daniell-Kolmogorov theorem, and this allows us to define trace equivalence: two states x and y are trace equivalent if for every time-*obs*-closed sets B , $\mathbb{P}^x(B) = \mathbb{P}^y(B)$.

An important remark has to be made here. Traditionally, trace equivalence for discrete time is much simpler than our definition of trace equivalence: two states x and y are *DT-trace equivalent* if for every C_0, \dots, C_n measurable sets such that for every i , for every z, z' such that $obs(z) = obs(z')$, $z \in C_i$ if and only if $z' \in C_i$,

$$\mathbb{P}^x(X_0 \in C_0, \dots, X_n \in C_n) = \mathbb{P}^y(X_0 \in C_0, \dots, X_n \in C_n)$$

Lemma 38 shows that any two DT-bisimilar states are DT-trace equivalent. Our definition of trace equivalence is stronger than this: it requires that \mathbb{P}^x and \mathbb{P}^y agree for all time-*obs*-closed sets. The sets $\{X_0 \in C_0, \dots, X_n \in C_n\}$ are only examples of such time-*obs*-closed sets.

Lemma 39. *In an LMP with finitely many atomic propositions, any two states x and y that are DT-bisimilar are trace equivalent.*

Proof. We need to prove that for any time-obs-closed set $B \subset \mathbb{N} \rightarrow S \uplus \{\partial\}$, $\mathbb{P}^x(B) = \mathbb{P}^y(B)$.

Recall the definition of Π from the definition of time-obs-closed (see definition 21): for a trajectory $\omega : \mathbb{N} \rightarrow S \uplus \{\partial\}$, $\Pi(\omega)(n) = \text{obs}(\omega(n))$. Now note that $B = \Pi^{-1}(\Pi(B))$. The first inclusion ($B \subset \Pi^{-1}(\Pi(B))$) always holds. For the converse direction, consider $\omega \in \Pi^{-1}(\Pi(B))$, that is there exists $\omega' \in B$ such that $\Pi(\omega) = \Pi(\omega')$. This means that for every n , $\text{obs}(\omega(n)) = \Pi(\omega)(n) = \Pi(\omega')(n) = \text{obs}(\omega'(n))$ and using the fact that B is time-obs-closed, we obtain that $\omega \in B$ which proves the desired equality.

Consider a finite cylinder on $\mathbb{N} \rightarrow 2^{AP}$, that is a set of the form:

$$\{\delta : \mathbb{N} \rightarrow 2^{AP} \mid \forall i \in \llbracket 0, n \rrbracket \delta(i) \in A_i\}$$

where n is an integer and for every i , A_i is a subset of 2^{AP} . Since there are only finitely many atomic propositions, the set 2^{AP} is finite and the sets A_i are measurable. Lemma 38 shows that for any finite cylinder C on $\mathbb{N} \rightarrow 2^{AP}$,

$$\mathbb{P}^x(\Pi^{-1}(C)) = \mathbb{P}^y(\Pi^{-1}(C))$$

The finite cylinders form a generating π -system of the σ -algebra on $\mathbb{N} \rightarrow 2^{AP}$ which means that for every measurable subset A of $\mathbb{N} \rightarrow 2^{AP}$,

$$\mathbb{P}^x(\Pi^{-1}(A)) = \mathbb{P}^y(\Pi^{-1}(A))$$

We can now look at our set B .

$$\mathbb{P}^x(B) = \mathbb{P}^x(\Pi^{-1}(\Pi(B))) = \mathbb{P}^y(\Pi^{-1}(\Pi(B))) = \mathbb{P}^y(B)$$

The first and third equations hold since $B = \Pi^{-1}(\Pi(B))$, and the second equation holds since $\Pi(B)$ is measurable and for every measurable subset A of $\mathbb{N} \rightarrow 2^{AP}$,

$$\mathbb{P}^x(\Pi^{-1}(A)) = \mathbb{P}^y(\Pi^{-1}(A))$$

□

Remark 40. It may look convoluted to use the map Π and move to the space $\mathbb{N} \rightarrow 2^{AP}$ instead of $\mathbb{N} \rightarrow S \uplus \{\partial\}$. However, this is a necessity due to the fact that we do not know what the σ -algebra of time-obs-closed sets look like. We know that it contains the σ -algebra generated by the sets $\Pi^{-1}(C)$ (for all C finite cylinders on $\mathbb{N} \rightarrow 2^{AP}$), but this inclusion is not an equality.

Proposition 41. *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 temporal equivalence.

Proof. Consider $(x, s) R' (y, s)$, $t \geq 0$ and a measurable and R' -closed set C . By definition of P_t , $P_t((x, s), C) = \tau_{\lfloor t+s \rfloor}(x, C')$ where $C' = \{z \mid (z, s') \in C\}$ with $s' = t + s - \lfloor t + s \rfloor$ (and similarly for y).

The set C' is R -closed. Indeed, consider two states $z \in C'$ and $z' \in S$ such that $z R z'$. These conditions imply that $(z, s') \in C$ and $(z, s') R' (z', s')$. Since the set C is R' -closed, $(z', s') \in C$ and hence by definition of the set C' , $z' \in C'$.

Since $(x, s) R' (y, s)$, we also have that $x R y$. By lemma 38, we have that $\tau_{\lfloor t+s \rfloor}(x, C') = \tau_{\lfloor t+s \rfloor}(y, C')$. This allows us to conclude that $P_t((x, s), C) = P_t(y, s), C)$.

The initiation condition (trace equivalence) is a direct consequence of Lemma 39. □

Proposition 42. *If the equivalence R is a temporal equivalence, then the relation R' defined as the transitive closure of the relation*

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

is a DT-bisimulation.

Proof. The relation R' is indeed an equivalence.

Let us consider $x R' y$, that is there exists $(x_i)_{i=1,\dots,n}, (t_i)_{i=1,\dots,n-1}$ and $(t'_i)_{i=1,\dots,n-1}$ such that $x = x_1, y = x_n$ and for every $1 \leq i \leq n - 1, (x_i, t_i) R (x_{i+1}, t'_i)$.

The fact that $obs(x) = obs(y)$ is a direct consequence of the initiation condition of a temporal equivalence.

Now, consider a measurable (in Σ) and R' -closed set C' . Define the set $C = \{(z, s) \mid z \in C', s \in [0, 1)\}$. It is measurable (in \mathcal{E}) and R -closed. Since R is a temporal equivalence, for every $i, P_1((x_i, t_i), C) = P_1((x_{i+1}, t'_i), C)$ for every $i \leq n$. Additionally, note that for every $z \in E$ and every $s \in [0, 1), P_1((z, s), C) = \tau(z, C')$. This proves that $\tau(x, C') = \tau(y, C')$. \square

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

Theorem 43. *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 temporally equivalent.*

5. Examples

In this work, examples are important in order to get an intuition. The examples we provide can be divided into three categories: (mostly) deterministic drifts, Brownian motion (and its variants) and Poisson process.

Most of these examples follow the same type of proofs: first computing the trace equivalence, then providing a group of symmetries and showing that the equivalence generated by this group of symmetries corresponds to the trace equivalence.

We will only detail a few of these examples. All computations can be found in Clerc (2021).

5.1 Basic examples

5.1.1 Deterministic drift

Consider a deterministic drift on the real line \mathbb{R} with constant speed $v \in \mathbb{R}_{>0}$ and a single atomic proposition. We study two cases: the first one with 0 as the only distinguished point and the second one with all the integers distinguished from the other points. In both cases, trace equivalence and greatest bisimulation are the same.

atomic proposition distinguishes...	trace equivalence/ bisimulation: $x R y$ if and only if	group of symmetries
zero	$x = y$ or $x, y > 0$	generated by $\{f_{x,y} \mid x, y > 0\}$
integers	$x - \lfloor x \rfloor = y - \lfloor y \rfloor$	$\{t_k \mid k \in \mathbb{Z}\}$ generated by $\{t_1\}$

where we define for every $x \in \mathbb{R}$ and every $k \in \mathbb{Z}, t_k(x) = x + k$ and for $x, y \in \mathbb{R}_{>0}$ the function $f_{x,y} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f_{x,y}(z) = \begin{cases} z & \text{if } z \leq 0 \\ \frac{y}{x}z & \text{if } 0 \leq z \leq x \\ z - x + y & \text{if } z > x \end{cases}$$

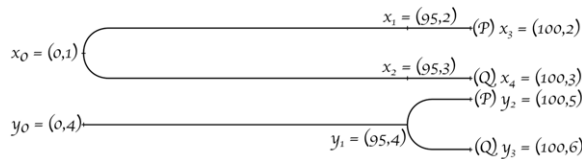
These functions are essentially identity functions with a rescaled portion in the middle to connect positive reals.

5.1.2 Fork

The following example is of particular interest because it shows that temporal equivalence and trace equivalence are not equivalent notions. It emphasises the importance of the induction conditions in the definitions of temporal equivalence and of bisimulation. It is an extension of a standard example in discrete time (see Section 4.1, p.86 of Milner (1989)) to our continuous-time setting. The process is a deterministic drift at constant speed with a single probabilistic fork (the process then goes to either branch with equal probability). We compare the case where the fork is at the start with the case where the fork happens later. There are two atomic propositions P and Q which enable the process to tell the difference between the ends of each fork.

One can find in Clerc (2021) a detailed and tedious description of the process and of the proofs. We will focus here on intuition.

A state is a pair (x, b) where x represents the time to be reached from the “origins” x_0 and y_0 and b represents the different “branches” of the state space. The state space with its two atomic propositions P and Q is the following:



This process is a deterministic drift at constant speed towards the right. When the process encounters a fork (in x_0 or in y_1), it randomly picks between the upper and lower branch and then continues as a deterministic drift to the right.

Lemma 44. Given $s \neq t \in [0, 100]$ and $i, j \in \llbracket 1, 6 \rrbracket$, the states (s, i) and (t, j) cannot be trace equivalent.

Proposition 45. Two states $x \neq y$ are trace equivalent if and only if

- $\{x, y\} = \{x_0, y_0\}$, or
- $\{x, y\} = \{(t, 2), (t, 5)\}$ for some $t \in (95, 100]$, or
- $\{x, y\} = \{(t, 3), (t, 6)\}$ for some $t \in (95, 100]$.

Proof. First, let us show that x_0 and y_0 are trace equivalent. There are two possible trajectories from x_0 : ω_2 and ω_3 going up (branch 2) and down (branch 3), respectively. Similarly, there are two trajectories from y_0 : ω_5 and ω_6 going first though branch 4 (from y_0 to y_1) and then going respectively up (branch 5) and down (branch 6). If a set B is time-obs-closed, ω_2 (resp. ω_3) is in the set B if and only if ω_5 (resp. ω_6) is in the set B . Furthermore,

$$\mathbb{P}^{x_0}(B) = \frac{1}{2}\delta_B(\omega_2) + \frac{1}{2}\delta_B(\omega_3) \text{ and } \mathbb{P}^{y_0}(B) = \frac{1}{2}\delta_B(\omega_5) + \frac{1}{2}\delta_B(\omega_6)$$

This shows that x_0 and y_0 are trace equivalent.

It is easy to see that two states satisfying condition 2 or 3 are trace equivalent.

In order to show that no other non-equal states can be trace equivalent, note that lemma 44 shows that it is enough to show that $(t, 2), (t, 3)$ and $(t, 4)$ are not trace equivalent. It is enough to consider the set of trajectories seeing the atomic proposition P after letting time go for 100:

$$B_P = \{\omega \mid \text{obs}(\omega(100)) = (1, 0)\}.$$

This set is time-*obs*-closed, but we have that $\mathbb{P}^{(t,2)}(B_P) = 1, \mathbb{P}^{(t,3)}(B_P) = 0$ and $\mathbb{P}^{(t,4)}(B_P) = 1/2$ □

Proposition 46. *The states x_0 and y_0 are not temporally equivalent (and hence they are not bisimilar either).*

Proof. First, the states x_1, x_2 and y_1 cannot be temporally equivalent. Indeed, consider the set of trajectories $B = \{\omega \mid \text{obs}(\omega(5)) = (1, 0)\}$. This set is time-*obs*-closed but we have that $\mathbb{P}^{x_1}(B) = 1, \mathbb{P}^{x_2}(B) = 0$ and $\mathbb{P}^{y_1}(B) = 1/2$.

Using Lemma 44, this shows that the states x_1, x_2 and y_1 can only be temporally equivalent to themselves.

Third, we can consider the set $\{x_1\}$. We have just shown that for any temporal equivalence R , the set $\{x_1\}$ is R -closed. But then, $P_t(x_0, \{x_1\}) = 1/2$ whereas $P_t(y_0, \{x_1\}) = 0$. □

This bisimulation is generated by the group of symmetries $\{f, g, id\}$ where f permutes branches 2 and 5 and g similarly permutes branches 3 and 6.

5.2 Examples based on Brownian motion

All these examples are variants of Brownian motion with either a single or no (first two cases of absorbing wall) atomic propositions.

5.2.1 Standard Brownian motion

atomic proposition distinguishes...	trace equivalence/ bisimulation: $x R y$ if and only if	group of symmetries
zero	$ x = y $	$\{s, id\}$
integers	$x - \lfloor x \rfloor = y - \lfloor y \rfloor$ or $\lceil y \rceil - y$	$\{id, t_k, s_k \mid k \in \mathbb{Z}\}$ generated by $\{s, t_1\}$
interval $[-1, 1]$	$ x = y $	$\{s, id\}$

where for every $x \in \mathbb{R}, s(x) = -x, s_k(x) = k - x$ and $t_k(x) = x + k$ for every $k \in \mathbb{Z}$.

With all integers distinguished: Let us consider the case when there is a single atomic proposition and $\text{obs}(x) = 1$ if and only if $x \in \mathbb{Z}$.

Proposition 47. *The set $\{id, t_k, s_k \mid k \in \mathbb{Z}\}$ is a group of symmetries.*

Proof. For every $k \in \mathbb{Z}$, the functions s_k and t_k are indeed homeomorphisms and $\text{obs} \circ t_k = \text{obs}$ for every $k \in \mathbb{Z}$ (and similarly for s_k).

Consider a state x and a measurable set B such that $(s_k)_*(B) = B$ and $(t_k)_*(B) = B$ for every $k \in \mathbb{Z}$. Brownian motion is invariant under translation which means that

$$\mathbb{P}^x(B) = \mathbb{P}^{t_k(x)}((t_k)_*(B))$$

Using the additional constraint that $(t_k)_*(B) = B$, we obtain the desired result. For s_k , note that $s_k = t_k \circ s$. Brownian motion is invariant under symmetry which means that

$$\mathbb{P}^x(B) = \mathbb{P}^{s(x)}(s_*(B)) = \mathbb{P}^{s(x)}(B)$$

and therefore

$$\mathbb{P}^x(B) = \mathbb{P}^{s_k(x)}(B)$$

□

Proposition 48. *Two states x and y are trace equivalent if and only if $x - \lfloor x \rfloor = y - \lfloor y \rfloor$ or $\lceil y \rceil - y$.*

Proof. The equivalence generated by the group of symmetries $\{id, t_k, s_k \mid k \in \mathbb{Z}\}$ is

$$\{(x, y \mid x - \lfloor x \rfloor = y - \lfloor y \rfloor \text{ or } \lceil y \rceil - y)\}$$

And therefore any two states x and y such that $x - \lfloor x \rfloor = y - \lfloor y \rfloor$ or $\lceil y \rceil - y$ are trace equivalent.

Let us show that no other states are trace equivalent. First note that if $x \in \mathbb{Z}$ and $y \notin \mathbb{Z}$, they cannot be trace equivalent. Let us now consider two trace equivalent states x and y that are not in \mathbb{Z} . Consider the sets $B_t = \{\omega \mid \exists s \in [0, t) \omega(s) \in \mathbb{Z}\}$. These sets are time-obs-closed:

- If $\omega \in B_t$ (i.e. there exists a time $s < t$ such that $\omega(s) \in \mathbb{Z}$) and ω' is such that for every time u , $obs(\omega(u)) = obs(\omega'(u))$, then in particular $\omega'(s) \in \mathbb{Z}$ and $\omega' \in B_t$.
- The sets B_t can also be expressed as:

$$B_t = \bigcup_{n \in \mathbb{Z}} \{\omega \mid T_n(\omega) < t\} = \bigcup_{n \in \mathbb{Z}} T_n^{-1}([0, t))$$

where T_n is the hitting time of $\{n\}$. From Lemma 18, we get that $T_n^{-1}([0, t))$ is measurable and hence that the sets B_t are measurable.

- We can now check the final condition: $\Pi(B_t) = \{0, 1\}^{\mathbb{N}}$ which is measurable.

Let us compute $\mathbb{P}^z(B_t)$ for any $z \in \mathbb{R}$. Since the distribution of Brownian motion is invariant under translation with respect to its starting point, we have that for every $z \in \mathbb{R}$ and $t > 0$,

$$\mathbb{P}^z(B_t) = \mathbb{P}^z(T_{\lceil z \rceil} \wedge T_{\lfloor z \rfloor} < t) = \mathbb{P}^{z - \lfloor z \rfloor}(T_0 \wedge T_1 < t).$$

Therefore, it is sufficient to study the distribution of $T_0 \wedge T_1$ under \mathbb{P}^c for every $c \in [0, 1)$. We claim that, for any pair $c, c' \in [0, 1)$, $T_0 \wedge T_1$ has identical distribution under \mathbb{P}^c and $\mathbb{P}^{c'}$ if and only if $|c - \frac{1}{2}| = |c' - \frac{1}{2}|$, that is, either $c = c'$ or $c = 1 - c'$. To see this, we consider the Laplace transform of $T_0 \wedge T_1$ under \mathbb{P}^c and use the formula 1.3.0.1 of Borodin and Salminen (1997) to write

$$\mathbb{E}^c \left[e^{-\lambda(T_0 \wedge T_1)} \right] = \frac{\cosh \left(\left(c - \frac{1}{2} \right) \sqrt{2\lambda} \right)}{\cosh \left(\frac{1}{2} \sqrt{2\lambda} \right)} \text{ for every } \lambda > 0.$$

Then, $T_0 \wedge T_1$ has identical distribution under \mathbb{P}^c and $\mathbb{P}^{c'}$ if and only if

$$\mathbb{E}^c \left[e^{-\lambda(T_0 \wedge T_1)} \right] = \mathbb{E}^{c'} \left[e^{-\lambda(T_0 \wedge T_1)} \right] \text{ for every } \lambda > 0,$$

which is equivalent to $|c - \frac{1}{2}| = |c' - \frac{1}{2}|$.

Thus, we have proven that for every pair $x, y \in \mathbb{R}$,

$$\mathbb{P}^x (B_t) = \mathbb{P}^y (B_t) \text{ for every } t > 0$$

if and only $x - \lfloor x \rfloor = y - \lfloor y \rfloor$ or $x - \lfloor x \rfloor = \lceil y \rceil - y$. □

The equivalence generated by this group of symmetry is trace equivalence which shows that trace equivalence, the greatest bisimulation and the greatest temporal equivalence are identical.

5.2.2 *Brownian motion with drift*

Let us consider a Brownian process with an additional drift: $W_t^{(v)} = W_t + vt$ (where W_t is the standard Brownian motion and $v > 0$).

atomic proposition distinguishes...	trace equivalence/ bisimulation: $x R y$ if and only if	group of symmetries
zero	$x = y$	{id}
integers	$x - \lfloor x \rfloor = y - \lfloor y \rfloor$	{id, $t_k \mid k \in \mathbb{Z}$ } generated by $\{t_1\}$
interval $[-1, 1]$	$x = y$	{id}

where for every $x \in \mathbb{R}$ and every $k \in \mathbb{Z}$, $t_k(x) = x + k$.

With zero distinguished: Let us consider the case when there is a single atomic proposition and $obs(x) = 1$ if and only if $x = 0$.

Proposition 49. *Two states are trace equivalent if and only if they are equal.*

Proof. The only thing to show is that two different states cannot be trace equivalent.

Define the set

$$B_t := \{\omega \mid \exists s < t \ \omega(s) = 0\} = \{T_0 < t\}$$

Similarly to what was done in the proof of proposition 48, this set is time-obs-closed.

Therefore, it is sufficient to study the distribution of T_0 under \mathbb{P}^z for every $z \in \mathbb{R}$. We claim that, for any pair x, y , T_0 has identical distribution under \mathbb{P}^x and \mathbb{P}^y if and only if $x = y$. To see this, we consider the Laplace transform of T_0 under \mathbb{P}^z for an arbitrary state $z \in \mathbb{R}$ and use the formula 2.2.0.1 of Borodin and Salminen (1997) to write

$$\mathbb{E}^z \left[e^{-\lambda T_0} \right] = \exp \left(-zv - |z| \sqrt{2\lambda + v^2} \right) \text{ for every } \lambda > 0.$$

Then, T_0 has identical distribution under \mathbb{P}^x and \mathbb{P}^y if and only if

$$\mathbb{E}^x \left[e^{-\lambda T_0} \right] = \mathbb{E}^y \left[e^{-\lambda T_0} \right] \text{ for every } \lambda > 0,$$

which is equivalent to

$$xv + |x| \sqrt{2\lambda + v^2} = yv + |y| \sqrt{2\lambda + v^2} \text{ for every } \lambda > 0,$$

which ultimately means that $x = y$. □

This shows that there is a single group of symmetries possible: {id} and that no two different states can be bisimilar or temporally equivalent.

5.2.3 Brownian motion with absorbing wall

Another common variation of Brownian motion is to add boundaries and to consider that the process does not move anymore or dies once it has hit a boundary. The boundary is called an “absorbing wall”.

The standard Brownian motion on the real line is denoted W_t and Ω is its space of trajectories. Assume the boundaries are at 0 and z . We introduce two hitting times T_0 and T_z (which are stopping times thanks to Lemma 18). We can now define the stopping time $T = T_0 \wedge T_z$. Finally, we obtain the Brownian motion with absorbing walls at 0 and z : for any ω in Ω

$$A_t(\omega) = \begin{cases} W_t(\omega) & \text{if } t < T(\omega) \\ \partial & \text{otherwise} \end{cases}$$

Note that here Ω is the space of trajectories for the original Brownian motion and serves as source of randomness for the new process.

Let us clarify why we have chosen this to be the state space; essentially the reason is to ensure that the trajectories are cadlag. The boundaries are not included in the state space, *that is* the process with absorbing walls at 0 and z has $(0, z)$ as its state space. This is forced by the requirement that trajectories are cadlag. To see this, consider what would happen if the wall was included in the state space. We would have to modify the behaviour of the process at 0 and z . But what would these modified trajectories look like? They would look like a trajectory of a standard Brownian motion: continuous until it reaches either state 0 or z (assume it happens at time s) and then the trajectory would jump to the state ∂ , that is the trajectory would be continuous on $[0, s]$ and then perform a jump and be continuous (constant even) on $(s, +\infty]$. This is not a cadlag trajectory because it is continuous on the “wrong side”.

absorbing boundary at	state space	atomic proposition distinguishes	trace equivalence/ bisimulation: $x R y$ if and only if	group of symmetries
0	$\mathbb{R}_{>0}$	-	$x = y$	{id}
0 and $b > 0$	$(0, b)$	-	$x = y$ or $x = b - y$	{id, s_b }
0 and $2b > 0$	$(0, 2b)$	b	$x = y$ or $x = 2b - y$	{id, s_{2b} }
0 and $4b > 0$	$(0, 4b)$	b	$x = y$	{id}

where $s_k(x) = k - x$ for every $x \in \mathbb{R}$ and $k = b$ or $2b$.

5.3 Poisson process

This is an example that we considered in Chen *et al.* (2020). A Poisson process models a continuous-time process in which a discrete variable is incremented. A standard example is the arrival of customers in a queue. Let us give some notations: a Poisson process is a non-decreasing process $(N_t)_{t \geq 0}$ onto the set of natural numbers \mathbb{N} , a discrete space. We define the set Ω of trajectories as usual on the state space. The probability distribution on the set Ω is defined as

$$\mathbb{P}^k(N_t = n) = \frac{(\lambda t)^{n-k}}{(n-k)!} e^{-\lambda t} \quad \text{for } n \geq k$$

We are going to study two cases. In the first case, we are able to test if there is an even or odd number of customers that have arrived. In the second case, we are able to test if there have been more customers in total than a critical value.

Testing evenness of number of customers: There is a single atomic proposition on the state space: $obs(k) = 1$ if and only if k is even.

Proposition 50. *Two states x and y are bisimilar (resp. temporally equivalent, trace equivalent) if and only if $x \equiv y \pmod 2$.*

Proof. Let us write R for the corresponding equivalence.

First, it is indeed a bisimulation. Consider $y = x + 2n$ where $n \in \mathbb{N}$ (note that $obs(x) = obs(y)$) and B a measurable, time- R -closed set.

For a measurable set B' , $\mathbb{P}^x(B') = \mathbb{P}^{x+2n}(B' + 2n)$, where $B' + 2n = \{t \mapsto \omega(t) + 2n \mid \omega \in B'\}$. In particular, $\mathbb{P}^x(B) = \mathbb{P}^y(B + 2n)$. Since B is time- R -closed, $B + 2n \subset B$, so $\mathbb{P}^x(B) \leq \mathbb{P}^y(B)$.

For the reverse direction, let M be the set of non-decreasing trajectories. Note that M is measurable since its complement is measurable:

$$\begin{aligned} M^c &= \{f \mid \exists t < s \text{ such that } f(t) > f(s)\} \\ &= \{f \mid \exists t, s \in \mathbb{Q} \text{ such that } t < s \text{ and } f(t) > f(s)\} \\ &= \bigcup_{t \in \mathbb{Q}} \bigcup_{s \in \mathbb{Q}, s > t} \{f \mid f(t) > f(s)\} \\ &= \bigcup_{t \in \mathbb{Q}} \bigcup_{s \in \mathbb{Q}, s > t} \bigcup_{n \in \mathbb{N}} \bigcup_{m=0}^{n-1} \{f \mid f(t) = m \text{ and } f(s) = n\}. \end{aligned}$$

Write $B_0 = B \cap M \cap \{\omega \mid \omega(0) = y\}$. Note that the process can only realise non-decreasing trajectories, therefore $\mathbb{P}^y(B) = \mathbb{P}^y(B_0)$. Define $B_1 = \{t \mapsto \omega(t) - 2n \mid \omega \in B_0\}$. Note that for every $\omega' \in B_1$ and $s \geq 0$, $\omega'(s) \in \mathbb{N}$ since for every $\omega \in B_0$ and $t \geq 0$, $\omega(t) \geq \omega(0) = x + 2n$. Since $B_1 + 2n = B_0$, we have that $\mathbb{P}^y(B_0) = \mathbb{P}^x(B_1)$. Furthermore, $B_1 \subset B$ since B is time- R -closed. Putting all this together: $\mathbb{P}^y(B) = \mathbb{P}^y(B_0) = \mathbb{P}^x(B_1) \leq \mathbb{P}^x(B)$.

This concludes the proof that R is a bisimulation. Now, notice that $x R y$ if and only if $obs(x) = obs(y)$. Since this equivalence is weaker than trace equivalence which is itself weaker than bisimulation, we have that R is the trace equivalence and the greatest bisimulation (and similarly the greatest temporal equivalence). □

Remark 51. This situation may look a lot like the deterministic or Brownian drift with parity as the atomic proposition. However, there is one key difference here: we are preventing translations to be bijections on the state space by only allowing non-negative numbers: those translations are not surjective. Therefore, the set of translations by an even number cannot be a group of symmetries. Proving that there is no greater group of symmetries than $\{id\}$ is not as trivial as it may look.

Testing for a critical value: Fix $m \in \mathbb{N}_{\geq 0}$, we define the function obs by $obs(x) = 1$ if and only if $x \geq m$.

Proposition 52. *Two states $x \neq y$ are bisimilar (resp. temporally equivalent, trace equivalent) if and only if $x, y \geq m$.*

Proof. Denote

$$R = \{(x, x) \mid x < m\} \cup \{(x, y) \mid x, y \geq m\}$$

Let us show that it is a bisimulation. Consider $x R y$ and assume $x \neq y$. This means that $x, y \geq m$ and hence $obs(x) = obs(y)$.

Now, also consider a measurable time- R -closed set B . Define $B' = B \cap M \cap \{\omega \mid \omega(0) \geq m\}$ where M is the set of non-decreasing functions. Similarly to previous example, M is measurable. Note that the process can only realise non-decreasing trajectories, therefore $\mathbb{P}^y(B) = \mathbb{P}^y(B')$ (and similarly for x).

There are now two cases to consider:

- If B' is empty, $\mathbb{P}^y(B) = \mathbb{P}^x(B) = 0$.
- Otherwise, there exists $\omega' \in B'$. Note that for every time $t \geq 0$, $\omega'(t) \geq \omega'(0) \geq m$ since ω' is non-decreasing.

We have that $B' = \{\omega \mid \omega(0) \geq m\} \cap M$. This means that for every $z \geq m$, $\mathbb{P}^z(B') = \mathbb{P}^z(\{\omega \mid \omega(0) \geq m\} \cap M) = 1$. And in particular, this shows that $\mathbb{P}^x(B) = \mathbb{P}^y(B) = 1$.

To prove that it is the greatest bisimulation, we show that it corresponds to trace equivalence. This proof can be adapted to show that $x, y \geq m$ are trace equivalent.

Clearly if $x < m$ and $y \geq m$, then x and y cannot be trace equivalent since the set $\{\omega \mid \text{obs}(\omega(0)) = 0\}$ is time-*obs*-closed, but $\mathbb{P}^x(\{\omega \mid \text{obs}(\omega(0)) = 0\}) = 1$ and $\mathbb{P}^y(\{\omega \mid \text{obs}(\omega(0)) = 0\}) = 0$.

Consider the case when $x \neq y$ are both less than m . Consider a time $t > 0$ and define $B_t = \{\omega \mid \omega(t) \geq m\}$. This set is time-*obs*-closed: the first two conditions are straightforward and $\Pi(B_t) = \{0, 1\}^{\mathbb{N}}$ if $t \notin \mathbb{N}$ and $\Pi(B_t) = \{\delta : \mathbb{N} \rightarrow \{0, 1\} \mid \delta(t) = 1\}$ if $t \in \mathbb{N}$. Both are measurable.

For $k < m$,

$$\mathbb{P}^k(B_t) = \sum_{n \geq m} \mathbb{P}^k(N_t = n) = e^{-\lambda t} \sum_{n \geq m-k} \frac{(\lambda t)^n}{n!}$$

This allows us to conclude that if $x \neq y$, $\mathbb{P}^x(B_t) \neq \mathbb{P}^y(B_t)$. □

There are several observations that can be made. First, in all examples, greatest temporal equivalence and bisimulation matched. We do not know if this is the case for all FD-processes or a subfamily of FD-processes. In particular, all the processes studied here are quite well-behaved and simple: Brownian motion has continuous trajectories (and not cadlag) for instance. Another interesting observation is that group of symmetries corresponds to intuition when handling a problem. Finally, notice how hitting times played a huge role in these proofs.

6. FD-Cospans: A Categorical Approach to Bisimulation

In this section, we explore a more categorical way of looking at bisimulation. We extend the notion of zigzag from discrete time. The core idea is that an equivalence can be viewed as a span or cospan of morphisms. This work was published in Chen *et al.* (2019).

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 Danos *et al.* (2006) 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. Spans and cospans of zigzag give rise to equivalent notions for analytic spaces, but it has been shown that it is not the case in general Sánchez Terraf (2011).

The intuition behind FD-homomorphisms is that they are quotients by bisimulations. However, topological properties do not behave well once they are quotiented. For instance, the quotient of a Hausdorff or locally compact space need not be Hausdorff or locally compact.

This section is where the additional hypothesis on *obs* (end of Section 3.1) is going to be vital: for each atomic proposition A , the atomic proposition partitions the state space into two subsets: one where A is satisfied and another one where it is not satisfied. We assume that one of these spaces is open (and the other is closed).

6.1 Feller-Dynkin homomorphism

The definition of bisimulation can easily be adapted to states in different Markov processes by constructing the disjoint union of the Markov processes in the following manner. First, one constructs the disjoint union of the two state spaces as topological spaces. It is then possible to extend the semigroups on the two state spaces to a semigroup on the disjoint union of the state spaces. The construction of the time-indexed family of Markov kernels, of the space of trajectories and of and of the space-indexed family of probabilities on the space of trajectories is the same as the one described in Section 3.1.

In that context, a bisimulation is defined in the natural way: If $x R y$ where $x \in E_i$ and $y \in E_j$ (i and j can be either 1 or 2 depending on which state space x and y are on), then:

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

(induction) for all measurable time- 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 $B_i \in \mathcal{G}_i$ and $B'_j \in \mathcal{G}_j$, $\mathbb{P}_i^x(B_i) = \mathbb{P}_j^y(B'_j)$ if the sets B_i and B'_j satisfy the following conditions: for all trajectories ω in $B_i \cup B'_j$,

$$\forall \omega' \in \Omega_i (\forall t \geq 0 \omega(t) R \omega'(t)) \Rightarrow \omega' \in B_i$$

and

$$\forall \omega' \in \Omega_j (\forall t \geq 0 \omega(t) R \omega'(t)) \Rightarrow \omega' \in B'_j$$

To go from the first statement of the induction condition to the other, write $B_i = B \cap \Omega_i$ and $B'_j = B \cap \Omega_j$. To go from the second statement to the first statement is slightly trickier to write down since B may contain trajectories on both E_1 and E_2 (switching state space). However, those trajectories have probability zero.

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

Definition 53. A continuous open surjective⁴ function $f : E \rightarrow E'$ is called a Feller-Dynkin homomorphism (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 $g \circ f$. This notion of FD-homomorphisms is designed to capture some aspects of bisimulation which is demonstrated in the following proposition.

Proposition 54. Given an FD-homomorphism f , the equivalence relation R defined on $E \uplus E'$ as

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

is a bisimulation on $E \uplus E'$.

Proof. Consider x and y such that $x R y$. We are going to assume that $f(x) = f(y)$, and we will be treating the case $xRf(x)$ at the same time.

First note that $obs(x) = obs' \circ f(x)$ since $obs = obs' \circ f$. Since $f(x) = f(y)$, we have that $obs(x) = obs' \circ f(x) = obs' \circ f(y) = obs(y)$. This gives us (initiation) for both cases.

Second, let us check the induction condition (induction). Write $\hat{\Omega}$ and $\hat{\mathcal{G}}$ for the set of trajectories and its σ -algebra on $E \uplus E'$. Consider a measurable time- R -closed set $\hat{B} \in \hat{\mathcal{G}}$.

Define the two sets

$$B' = \hat{B} \cap \Omega'$$

$$B = \{\omega \in \Omega \mid f \circ \omega \in B'\}$$

As f is an FD-homomorphism, we have that $\mathbb{P}^{f(x)}(B') = \mathbb{P}^x(B)$.

Let us show that $B = \hat{B} \cap \Omega$.

- Consider $\omega \in B$, that is $f \circ \omega \in \hat{B} \cap \Omega'$. By definition of the set B , $\omega \in \Omega$. Furthermore, $f \circ \omega \in \hat{B}$. By definition of R , we have that for all $t \geq 0$, $\omega(t) R (f \circ \omega(t))$. Since the set \hat{B} is time- R -closed and $f \circ \omega \in \hat{B}$, we have that $\omega \in \hat{B}$ which proves the first inclusion.
- Consider $\omega \in \hat{B} \cap \Omega$. The trajectory $f \circ \omega$ is well-defined and is in Ω' since f is continuous. Similarly to what was done for the first inclusion, we get that $f \circ \omega \in \hat{B}$ since $\omega \in \hat{B}$ and \hat{B} is R -closed. This proves that $\omega \in B$

We get that $\mathbb{P}^{f(x)}(\hat{B} \cap \Omega') = \mathbb{P}^x(\hat{B} \cap \Omega)$.

Since $f(x) = f(y)$, we also get that $\mathbb{P}^x(\hat{B} \cap \Omega) = \mathbb{P}^y(\hat{B} \cap \Omega)$. □

A nice consequence of this result is that the equivalence relation R defined on E as $R = \{(x, y) \in E \times E \mid f(x) = f(y)\} = \ker(f)$ is a bisimulation on E .

Proposition 55. *Given an FD-process on a state space E and a group of symmetries \mathcal{H} for that FD-process, then there exists an FD-process on a state space E' and an FD-homomorphism $\pi : E \rightarrow E'$ such that given two states $x, y \in E$, $\pi(x) = \pi(y)$ if and only if $x R y$ (where R is the bisimulation generated by the group of symmetries \mathcal{H}).*

Proof. Define the set $E' = E/R$ and the quotient map $\pi : E \rightarrow E/R$. We can equip the set E/R with the largest topology that makes the map π continuous.

Let us show that the map π is open: for any open set U in E ,

$$\begin{aligned} \pi^{-1}\pi(U) &= \{x \mid \exists y \in U \pi(x) = \pi(y)\} \\ &= \{x \mid \exists y \in U x R y\} \\ &= \{x \mid \exists y \in U \exists h \in \mathcal{H} x = h(y)\} \\ &= \bigcup_{h \in \mathcal{H}} h(U) \end{aligned}$$

Every symmetry $h \in \mathcal{H}$ is open since they are homeomorphisms; hence, $h(U)$ are all open sets. This means that the set $\pi^{-1}\pi(U)$ is open and hence the map π is open.

Let us show that $R \subset E \times E$ is closed. Consider a sequence (x_n, y_n) in R that converges to (x, y) . In particular, for every $n \in \mathbb{N}$, $\pi(x_n) = \pi(y_n)$. Furthermore, $\lim_n \pi(x_n) = \pi(x)$ and $\lim_n \pi(y_n) = \pi(y)$ since π is continuous. By uniqueness of limit, $\pi(x) = \pi(y)$ and hence $(x, y) \in R$.

Since the quotient map is open and the equivalence R is closed, the space E/R is Hausdorff.

Since the map π is open and the space E is locally compact, the space E/R is also locally compact.

Let us now clarify what the FD-process is on that state space. First, we define for $x \in E/R$

$$obs'(x) = obs(y) \quad \text{where } \pi(y) = x$$

This is indeed well-defined as whenever $y R z$ for $y, z \in E$, then there exists $h \in \mathcal{H}$ such that $h(y) = z$ and hence $obs(y) = obs(h(y)) = obs(z)$. Note that the function $obs' : E' \rightarrow 2^{AP}$ is measurable (the proof is similar to the proof of Lemma 71). Furthermore, $obs = obs' \circ \pi$ by definition of π .

Let us denote Ω' the set of trajectories on the one-point compactification of $E/R: E'_\partial$. This defines a time-indexed family of random variables $(X'_t)_{t \geq 0}$ by $X'_t(\omega) = \omega(t)$ for every trajectory $\omega \in \Omega'$ and every time $t \geq 0$. The σ -algebra \mathcal{G}' on the set of trajectories Ω' is defined as usual for FD-processes:

$$\mathcal{G}' = \sigma\{(X'_s)^{-1}(C') \mid s \geq 0, C' \in \mathcal{E}'\}$$

We define the following family of probabilities on the set Ω' : for $x \in E/R, B' \in \mathcal{G}'$,

$$\mathbb{Q}^x(B') = \mathbb{P}^y(B) \quad \text{if } x = \pi(y) \text{ with } B = \{\omega \in \Omega' \mid \pi \circ \omega \in B'\}$$

There are two things to show in order to check that \mathbb{Q}^x is well-defined:

- The set B is measurable. The proof is done by induction on the structure of B' .
 - First, if $B' = (X'_s)^{-1}(C') = \{\omega \in \Omega' \mid \omega(s) \in C'\}$ for $C' \in \mathcal{E}'$, then $B = (X_s)^{-1}(\pi^{-1}(C'))$. And since π is continuous, we know that $\pi^{-1}(C') \in \mathcal{E}$ and hence $B \in \mathcal{G}$.
 - If $B' = (A')^C$ with $A' \in \mathcal{G}'$ and $A = \{\omega \in \Omega \mid \pi \circ \omega \in A'\} \in \mathcal{G}$, then $B = \Omega \setminus A$. And since $A \in \mathcal{G}$, we get that $B \in \mathcal{G}$.
 - If $B' = \bigcup_{i \in \mathbb{N}} A'_i$ where for every $i \in \mathbb{N}, A'_i \in \mathcal{G}'$ and $A_i = \{\omega \in \Omega \mid \pi \circ \omega \in A'_i\} \in \mathcal{G}$, then $B = \bigcup_{i \in \mathbb{N}} A_i$. And since $A_i \in \mathcal{G}$ for every $i \in \mathbb{N}$, we get that $B \in \mathcal{G}$.
- If y, z are in E and such that $\pi(y) = \pi(z)$, then $\mathbb{P}^y(B) = \mathbb{P}^z(B)$. Indeed, having $\pi(y) = \pi(z)$ means that $y R z$ where R is the bisimulation generated by the group of symmetries \mathcal{H} . It is therefore enough to show that the set B is time- R -closed. Consider two trajectories ω and ω' such that at every time $t \geq 0, \omega(t) R \omega'(t)$, that is $\pi(\omega(t)) = \pi(\omega'(t))$. This means that $\pi \circ \omega = \pi \circ \omega'$ and hence by definition of $B, \omega \in B$ if and only if $\omega' \in B$.

By definition of \mathbb{Q} , we know that for all $x \in E$ and for all measurable sets $B' \subset \Omega', \mathbb{Q}^{\pi(x)}(B') = \mathbb{P}^x(B)$ where $B := \{\omega \in \Omega \mid \pi \circ \omega \in B'\}$. We have already shown that the map π was surjective, continuous and open and that $obs = obs' \circ \pi$. Which shows that π is an FD-homomorphism.

However, we still have to check that this family $(\mathbb{Q}^y)_{y \in E/R}$ corresponds to an FD-process. That family of probabilities $(\mathbb{Q}^x)_{x \in E/R}$ on Ω' yields a Markov kernel on E/R :

$$\text{for } x \in E/R, t \geq 0 \text{ and } C \in \mathcal{E}' \quad P'_t(x, C) = \mathbb{Q}^x((X'_t)^{-1}(C'))$$

which in turns yields an FD-semigroup:

$$\text{for } f \in C_0(E') \text{ and } x \in E/R \quad (\hat{P}'_t f)(x) = \int f dP'_t(x, -)$$

In order to see that $(\hat{P}'_t)_{t \geq 0}$ forms an FD-semigroup, it is convenient to note that for any z such that $x = \pi(z)$

$$\hat{P}'_t f(x) = \int_{E/R} f dP'_t(x, -) = \int_E f \circ \pi dP_t(z, -) = [\hat{P}_t(f \circ \pi)](z)$$

This concludes the proof. □

Example 56. Here is an example with one atomic proposition. We consider the standard Brownian motion and three of its variations.

- (1) First define $\mathcal{M}_{standard}$ to be the standard Brownian motion on the real line with $obs_{standard}(x) = 1$ if and only if $x \in \mathbb{Z}$.
- (2) We write $\mathcal{M}_{refl,1}$ for the reflected Brownian motion on $[0, 1]$ with $obs_{refl,1}(x) = 1$ if and only if $x = 0$ or 1 .
- (3) We write $\mathcal{M}_{refl,1/2}$ for the reflected Brownian motion on $[0, \frac{1}{2}]$ with $obs_{refl,1/2}(x) = 1$ if and only if $x = 0$.
- (4) Finally, we write \mathcal{M}_{circ} for the standard Brownian motion on the circle of radius $\frac{1}{2\pi}$ and of perimeter 1. We will identify points on the circle with the angle wrt the vertical. The atomic proposition is given by $obs_{circ}(x) = 1$ if and only if $x = 0$.

Let us detail how the Markov kernel is obtained for $\mathcal{M}_{refl,1}$. Write $(P_t)_{t \geq 0}$ for the Markov kernel for the standard Brownian motion on \mathbb{R} . We define the Markov kernel $(Q_t)_{t \geq 0}$ for $\mathcal{M}_{refl,1}$ for every $x \in [0, 1]$ and C measurable subset of $[0, 1]$:

$$Q_t(x, C) = P_t(x, C')$$

where

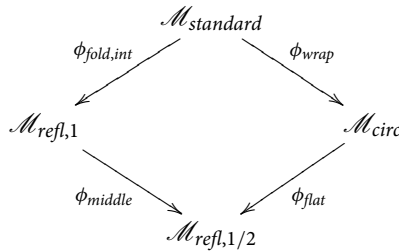
$$C' = \bigcup_{k \in \mathbb{Z}} (C + 2k) \uplus \bigcup_{k \in \mathbb{Z}} (\bar{C} + 1 + 2k)$$

with $C + k = \{x + k \mid x \in C\}$ is the translation by k of the original set C and $\bar{C} = \{1 - x \mid x \in C\}$ is the symmetry of the original set C around $1/2$. The intuition for that variation is that we “fold” the real line at each integer. The atomic proposition is therefore held at both extremities since they correspond to the integers of the real line once folded.

We will only provide intuition for the remaining two variations. First, the reflexive variation on $[0, 1/2]$ is obtained from the standard Brownian motion by folding the real line at each integer and at each half-integer (i.e. $k + 1/2$ where $k \in \mathbb{Z}$). The atomic proposition therefore only holds at 0. Indeed, all integers of the original real line get folded into 0.

Second \mathcal{M}_{circ} is obtained from the standard Brownian motion by “wrapping” the real line around a circle of perimeter 1. All the integers of the initial real line are thus mapped to a single point on the circle (which we decide to be 0).

We can define some natural mappings between these processes:



where the upper two morphisms correspond to the intuition provided before:

$$\begin{aligned} \phi_{wrap} : \mathbb{R} &\rightarrow [0, 2\pi) \\ x &\mapsto 2\pi(x - \lfloor x \rfloor) \end{aligned}$$

$$\begin{aligned} \text{and } \phi_{fold,int} : \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto x - 2n \text{ with } n \in \mathbb{Z} \text{ such that } 2n \leq x < 2n + 1 \\ x &\mapsto 2n + 2 - x \text{ with } n \in \mathbb{Z} \text{ such that } 2n + 1 \leq x < 2n + 2 \end{aligned}$$

Now, the remaining two morphisms intuitively make sense. How do we obtain the reflexive Brownian motion on $[0, 1/2]$ from a reflexive Brownian motion on $[0, 1]$? We “fold” the interval

$[0, 1]$ at its middle point $1/2$:

$$\begin{aligned} \phi_{middle} : [0, 1] &\rightarrow [0, 1/2] \\ x &\mapsto x \text{ if } x \leq 1/2 \\ x &\mapsto 1 - x \text{ otherwise} \end{aligned}$$

In order to obtain the reflexive variation on $[0, 1/2]$ from the circle, intuitively we “flatten” the circle by identifying the two halves that meet at 0 (the point with the atomic proposition):

$$\begin{aligned} \phi_{flat} : [0, 2\pi) &\rightarrow [0, 1/2] \\ \theta &\mapsto \theta/2\pi \text{ if } \theta \leq \pi \\ \theta &\mapsto (2\pi - \theta)/2\pi \text{ if } \theta \geq \pi \end{aligned}$$

All these morphisms correspond to the construction of one variation of Brownian motion from another one. It is therefore easy to deduce that all these morphisms are FD-homomorphisms.

6.2 Definition of cospans

It is time to retrieve the relational aspect of bisimulation in this functorial framework. For discrete time, the initial definition used spans of zigzags Desharnais *et al.* (2002). This is equivalent to considering cospans of zigzags in analytic spaces, but this is not necessarily the case for non-analytic spaces Sánchez Terraf (2011).

Definition 57. An FD-cospan is a cospan of FD-homomorphisms.

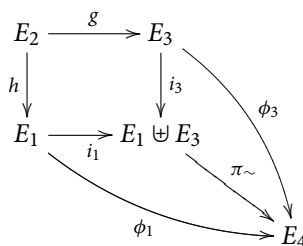
In order to show that FD-cospans behave like an equivalence, it is a necessity to show that they are somehow transitive. Namely, if we have two cospans $S \xrightarrow{f} E_1 \xleftarrow{h} E_2$ and $E_2 \xrightarrow{g} E_3 \xleftarrow{k'} S'$, can we construct an FD-cospan relating S and S' ? The following theorem shows that we can.

Theorem 58. The category with FD-processes as objects and FD-homomorphisms as morphisms has pushouts.

The proof is already quite long, and the proofs of some sublemmas can be found in Appendix B.

Proof. Consider three FD-processes $(E_j, \mathcal{E}_j, (\hat{P}_t^j), (P_t^j), \Omega_j, \mathcal{G}_j, (\mathbb{P}_j^x), obs_j)_{j=1,2,3}$ and two FD-homomorphisms $h : E_2 \rightarrow E_1$ and $g : E_2 \rightarrow E_3$.

There are two inclusions $i_1 : E_1 \rightarrow E_1 \uplus E_3$ and $i_3 : E_3 \rightarrow E_1 \uplus E_3$. Define the equivalence relation \sim on $E_1 \uplus E_3$ as the smallest equivalence such that for all $x_1 \in E_1$ and $x_3 \in E_3$, $i_1(x_1) \sim i_3(x_3)$ if there exists $z \in E_2$, such that $x_1 = h(z)$, and $x_3 = g(z)$. Define $E_4 = E_1 \uplus E_3 / \sim$ with its corresponding quotient $\pi_{\sim} : E_1 \uplus E_3 \rightarrow E_4$ and the two maps $\phi_3 = \pi_{\sim} \circ i_3$ and $\phi_1 = \pi_{\sim} \circ i_1$. This corresponds to the pushout in *Set*.



We equip this set with the largest topology that makes π_{\sim} continuous (where the topology on $E_1 \uplus E_3$ is the topology inherited from the inclusions). This means that a set $O \subset E_4$ is open if and only if $\pi_{\sim}^{-1}(O) \subset (E_1 \uplus E_3)$ is open. This corresponds to the pushout in *Top*.

It is worth noting that this topological space $(E_1 \uplus E_3)/\sim$ is bijective to the space E_2/\approx where \approx is defined on E_2 as the smallest equivalence such that if $g(z) = g(z')$ or $h(z) = h(z')$, then $z \approx z'$ (see Lemma 68). We will write $\pi_{\approx} : E_2 \rightarrow E_4 = (E_1 \uplus E_3)/\sim$ for the map $E_2 \xrightarrow{h} E_1 \xrightarrow{i_1} E_1 \uplus E_3 \xrightarrow{h} E_4$. This is equal to the map $E_2 \xrightarrow{g} E_3 \xrightarrow{i_3} E_1 \uplus E_3 \xrightarrow{h} E_4$. It is a quotient map and thus surjective.

Lemma 69 shows that the space E_4 equipped with its topology and Borel-algebra is locally compact and Hausdorff.

Note that both maps ϕ_1 and ϕ_3 are surjective, continuous and open (see Lemma 70)

We define obs_4 as such: $obs_4(x_4) = obs_2(x_2)$ where $x_4 = \pi_{\approx}(x_2)$. This is well-defined since if $z \approx z'$, then $obs_2(z) = obs_2(z')$. Indeed, that would mean that there is a sequence $z = z_0, z_1, \dots, z_n = z'$ where z_i and z_{i+1} have same image by either g or h . Since g and h are FD-homomorphisms, that means that $obs_2(z_i) = obs_2(z_{i+1})$. It is equivalent to defining obs_4 as:

$$obs_4(x_4) = \begin{cases} obs_1(x_1) & \text{if } x_4 = \phi_1(x_1) \\ obs_3(x_3) & \text{if } x_4 = \phi_3(x_3) \end{cases}$$

Lemma 71 shows that this map is indeed measurable.

Finally, let us define the FD-process on E_4 . Let Ω_4 be the set of trajectories on E_4 . The random variable X^4 is straightforward to define: $X_t^4(\omega) = \omega(t)$ for every trajectory $\omega \in \Omega_4$ and time $t \geq 0$. The set of trajectories Ω_4 is equipped with a σ -algebra \mathcal{G}_4 defined in the standard way for FD-processes:

$$\mathcal{G}_4 = \sigma(X_s^4 \mid 0 \leq s < \infty) = \sigma(\{(X_s^4)^{-1}(C) \mid 0 \leq s < \infty, C \in \mathcal{E}_4\})$$

We are now equipped to define the FD-process on E_4 . Consider $x_4 \in E_4$ and $B_4 \in \mathcal{G}_4$. We define

$$\mathbb{P}_4^{x_4}(B_4) = \mathbb{P}_2^{x_2}(B_2)$$

where $x_2 \in E_2$ and $B_2 \subset \Omega_2$ are such that $x_4 = \pi_{\approx}(x_2)$ and $B_2 = \{\omega \in \Omega_2 \mid \pi_{\approx} \circ \omega \in B_4\}$. We show in Lemma 72 that this is indeed well-defined and that

$$\mathbb{P}_4^{x_4}(B_4) = \mathbb{P}_2^{x_2}(B_2) = \mathbb{P}_3^{x_3}(B_3) = \mathbb{P}_1^{x_1}(B_1)$$

where $x_1 \in E_1$ is such that $\phi_1(x_1) = x_4$, $x_3 \in E_3$ is such that $\phi_3(x_3) = x_4$ and $B_1 = \{\omega \in \Omega_1 \mid \phi_1 \circ \omega \in B_4\}$ and $B_3 = \{\omega \in \Omega_3 \mid \phi_3 \circ \omega \in B_4\}$.

Let us now check that this is indeed an FD-process. We define first the corresponding kernel for $t \geq 0, x \in E_4$ and $C \in \mathcal{E}_4$,

$$P_t^4(x, C) = \mathbb{P}_4^x(\{\omega \mid \omega(t) \in C\})$$

For $x_2 \in E_2$ such that $\pi_{\approx}(x_2) = x$, we have that $P_t^4(x, C) = P_t^2(x_2, \pi_{\approx}^{-1}(C))$. We can now define the operator

$$\hat{P}_t^4 f(x) = \int_{y \in E_4} f(y) P_t^4(x, dy)$$

Using the corresponding equality for the kernel and a change of measure, we get that if $x = \pi_{\approx}(x_2)$, then $\hat{P}_t^4 f(x) = \hat{P}_t^2(f \circ \pi_{\approx})(x_2)$. The fact that $(\hat{P}_t^4)_{t \geq 0}$ is a semigroup follows from $(\hat{P}_t^2)_{t \geq 0}$ being itself a semigroup.

By construction, ϕ_1 and ϕ_3 are indeed FD-homomorphisms.

Let us now check the universal property. Assume there is an FD-process $(E_5, \mathcal{E}_5, (P_t^5)_t)$ and two FD-homomorphisms $\psi_1 : E_1 \rightarrow E_5$ and $\psi_3 : E_3 \rightarrow E_5$ such that $\psi_1 \circ h = \psi_3 \circ g$. Since E_4 is the corresponding pushout in *Top*, there exists a continuous morphism $\gamma : E_4 \rightarrow E_5$ such that $\psi_j = \gamma \circ \phi_j$ (for $j = 1, 2$). Let us show that it is an FD-homomorphism.

First, we want to show that $obs_4 = obs_5 \circ \gamma$. Let $x_4 \in E_4$, there are two cases: $x_4 = \phi_1(x_1)$ or $x_4 = \phi_3(x_3)$. Consider the first case $x_4 = \phi_1(x_1)$ (the other case is similar).

$$\begin{aligned} obs_5 \circ \gamma(x_4) &= obs_5 \circ \gamma \circ \phi_1(x_1) \\ &= obs_5 \circ \psi_1(x_1) \\ &= obs_1(x_1) \quad \text{since } \psi_1 \text{ is an FD-homomorphism} \\ &= obs_4(\phi_1(x_1)) \quad \text{since } \phi_1 \text{ is an FD-homomorphism} \\ &= obs_4(x_4) \end{aligned}$$

Second, we show that for $x_4 \in E_4$ and $B_5 \in \mathcal{G}_5$, $\mathbb{P}_5^{\gamma(x_4)}(B_5) = \mathbb{P}_4^{x_4}(B_4)$ with $B_4 = \{\omega \in \Omega_4 \mid \gamma \circ \omega \in B_5\}$. Consider the first case $x_4 = \phi_1(x_1)$ (the other case $x_4 = \phi_3(x_3)$ is similar). First, as ψ_1 is an FD-homomorphism,

$$\begin{aligned} \mathbb{P}_5^{\gamma(x_4)}(B_5) &= \mathbb{P}_5^{\gamma(\phi_1(x_1))}(B_5) = \mathbb{P}_5^{\psi_1(x_1)}(B_5) \\ &= \mathbb{P}_1^{x_1}(B_1) \quad \text{where } B_1 = \{\omega \in \Omega_1 \mid \psi_1 \circ \omega \in B_5\} \end{aligned}$$

Second, as ϕ_1 is an FD-homomorphism,

$$\mathbb{P}_4^{x_4}(B_4) = \mathbb{P}_4^{\phi_1(x_1)}(B_4) = \mathbb{P}_1^{x_1}(B'_1)$$

where $B'_1 = \{\omega \in \Omega_1 \mid \phi_1 \circ \omega \in B_4\}$. Finally, we have that

$$\begin{aligned} B'_1 &= \{\omega \in \Omega_1 \mid \phi_1 \circ \omega \in B_4\} \\ &= \{\omega \in \Omega_1 \mid \gamma \circ \phi_1 \circ \omega \in B_5\} = B_1 \end{aligned}$$

which concludes the proof. □

What this theorem shows is that FD-cospans behave like equivalence relations. Indeed, this means that if we have an FD-cospan (f_1, g_1) between FD-processes \mathcal{M}_1 and \mathcal{M}_2 and another FD-cospan (f_2, g_2) between FD-processes \mathcal{M}_2 and \mathcal{M}_3 , then there exists an FD-cospan (f_3, g_3) between the FD-processes \mathcal{M}_1 and \mathcal{M}_3 .

We have already showed how FD-homomorphisms yield bisimulations in Proposition 54. Now, consider an FD-cospan (f, g) with $f : E_1 \rightarrow E_3$ and $g : E_2 \rightarrow E_3$ inducing the bisimulations R_f on $E_1 \uplus E_3$ and R_g on $E_2 \uplus E_3$ respectively. We can define a bisimulation R on $E_1 \uplus E_2$ as

$$R = \{(x, y), (y, x) \mid \exists z \in E_3 \text{ such that } (x, z) \in R_f \text{ and } (y, z) \in R_g\}$$

Conversely, given a group of symmetries, it is also possible to define an FD-cospan corresponding to its corresponding bisimulation as is stated next:

Theorem 59. *Consider two FD-processes on the state spaces E_1 and E_2 respectively. Let \mathcal{H} be a group of symmetries of the FD-process on $E_1 \uplus E_2$ such that*

$$\text{for all } x_1 \in E_1 \text{ there exists } x_2 \in E_2 \text{ such that } i_1(x_1) R i_2(x_2)$$

and

$$\text{for all } x_2 \in E_2 \text{ there exists } x_1 \in E_1 \text{ such that } i_1(x_1) R i_2(x_2)$$

where R is the bisimulation generated by the group of symmetries \mathcal{H} on $E_1 \uplus E_2$. Then, there exists (f, g) an FD-cospan such that

- for all $x_1 \in i_1(E_1), x_2 \in i_2(E_2), x_1 R x_2$ if and only if $f(x_1) = g(x_2)$,
- for all $x_1, y_1 \in i_1(E_1), x_1 R y_1$ if and only if $f(x_1) = f(y_1)$, and
- for all $x_2, y_2 \in i_2(E_2), x_2 R y_2$ if and only if $g(x_2) = g(y_2)$.

Proof. We use the same notations as before for the FD-processes on E_1 and E_2 .

Define the set $E = (E_1 \uplus E_2)/R$. There are two inclusions $i_1 : E_1 \rightarrow E_1 \uplus E_2$, $i_2 : E_2 \rightarrow E_1 \uplus E_2$ and a quotient $\pi : E_1 \uplus E_2 \rightarrow E$. Using Proposition 55, we know that π is an FD-homomorphism and we know what the FD-process $(E, \mathcal{E}, \Omega, \mathcal{G}, obs, (\mathcal{P}^y)_{y \in E})$ on E is.

We define $f = \pi \circ i_1 : E_1 \rightarrow E$ and $g = \pi \circ i_2 : E_2 \rightarrow E$.

We only have to prove that f is indeed an FD-homomorphism.

First note that f is continuous as both i_1 and π are. Similarly, f is relatively open (open on its image set) since i_1 is relatively open and π is open. It is also surjective: consider $y \in E$, then there exists $x \in E_1 \uplus E_2$ such that $\pi(x) = y$. Then there are two possibilities:

- either $x = i_1(x_1)$ where $x_1 \in E_1$, in which case $f(x_1) = y$
- or $x = i_2(x_2)$ where $x_2 \in E_2$. But then, we know that there exists $x_1 \in E_1$ such that $i_1(x_1) R i_2(x_2)$, that is such that $\pi \circ i_1(x_1) = \pi \circ i_2(x_2) = y$

There are two remaining conditions to check for f to be indeed an FD-homomorphism: that $obs_1 \circ f = obs$ and for $x_1 \in E_1$ and $B \in \mathcal{G}$, $\mathbb{P}^{f(x_1)}(B) = \mathbb{P}_1^{x_1}(B)$ where $B = \{\omega \in \Omega_1 \mid f \circ \omega \in B'\}$. These two conditions directly follow from the fact that π itself satisfies these conditions. □

We thus have the following correspondence between bisimulation, groups of symmetries and FD-cospan:

Theorem 60. *If there exists a group of symmetries \mathcal{H} such that two states are related by that group of symmetries, then there exists an FD-cospan (f, g) such that $f(x) = g(y)$.*

If there exists an FD-cospan (f, g) such that $f(x) = g(y)$, then the two states x, y are bisimilar.

7. Game Interpretation

Fijalkow *et al.* (2017) and Clerc *et al.* (2019) introduced a game for discrete-time processes. We have adapted that game to the continuous-time setting. We have published those games in Chen *et al.* (2020). We will omit the proofs in this section but they can be found in Clerc (2021). However, it is especially interesting to note that the game interpretation of bisimulation emphasises once again the role of trajectories in that concept whereas the game interpretation of temporal equivalence resembles that in discrete time very closely.

7.1 Game interpretation of bisimulation

Definition 61. *Two trajectories ω and ω' are time-bisimilar if for all times $t \geq 0$, $\omega(t)$ and $\omega'(t)$ are bisimilar.*

Lemma 62. *Two states x and y are bisimilar if and only if the trajectories ω_x and ω_y are time-bisimilar where ω_z is the trajectory defined by $\omega_z(t) = z$ for all times $t \geq 0$ for a given state z .*

We define the following game. Duplicator's plays are pairs of trajectories that he claims are time-bisimilar. Spoiler is trying to prove him wrong.

- Given two trajectories ω and ω' , Spoiler chooses $t \geq 0$ and $B \neq \emptyset \in \mathcal{G}$ such that $\mathbb{P}^{\omega(t)}(B) \neq \mathbb{P}^{\omega'(t)}(B)$
- Duplicator answers by choosing $\omega_0 \in B$ and $\omega_1 \notin B$ such that $obs \circ \omega_0 = obs \circ \omega_1$ and the game continues from (ω_0, ω_1)

A player who cannot make a move at any point loses. Duplicator wins if the game goes on forever. The only way for Spoiler to win is to choose a time-*obs*-closed set.

Theorem 63. *Two trajectories ω and ω' are time-bisimilar if and only if Duplicator has a winning strategy from (ω, ω') .*

Corollary 64. *Two states x and y are bisimilar if and only if Duplicator has a winning strategy from (ω_x, ω_y) .*

7.2 Game interpretation of temporal equivalence

We define the following game. Duplicator's plays are pairs of states that he claims are temporally equivalent. Spoiler is trying to prove him wrong.

- Given two states x and y , Spoiler chooses $t \geq 0$ and $C \neq \emptyset \in \mathcal{E}$ such that $P_t(x, C) \neq P_t(y, C)$.
- Duplicator answers by choosing $x_1 \in C$ and $y_1 \notin C$ that are trace equivalent and the game continues from (x_1, y_1)

A player who cannot make a move at any point loses. Duplicator wins if the game goes on forever. The only way for Spoiler to win is to choose a set that is closed under trace equivalence. Duplicator's only valid moves are pairs of trace equivalent states.

Theorem 65. *Two states x and y are temporally equivalent if and only if Duplicator has a winning strategy from (x, y) .*

8. Conclusion

This work has been about extending the notion of behavioural equivalence (bisimulation) that existed in discrete time into continuous time. We have shown that several notions could actually be defined and have given other characterisations of them. However, what this work has emphasised is that there are measurability issues that make it much harder to work with continuous time: for instance, hitting times that were key in all the computations of our examples need not be measurable in general. There are numerous exciting further directions nonetheless, one of which being metrics and approximations.

Acknowledgements. Linan Chen was supported by NSERC.

Florence Clerc was supported by NSERC and IVADO.

Prakash Panagaden is grateful for the support of Huawei through a grant to the University of Edinburgh which allowed him to visit the School of Informatics

Notes

1 cadlag stands for the French "continu à droite, limite à gauche"

2 The σ -algebra \mathcal{G} is the same as the one induced by the Skorohod metric, see theorem 16.6 of Billingsley (1999)

3 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 .

4 Recall that f is continuous if and only if for every open set U' in E' , $f^{-1}(U')$ is open in E and that f is an open map if and only if for every open set U in E , $f(U)$ is open in E' .

References

Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd edition, Wiley Interscience.

Billingsley, P. (2008). *Probability and Measure*, John Wiley & Sons.

Blute, R., Desharnais, J., Edalat, A. and Panangaden, P. (1997). Bisimulation for labelled Markov processes. In *Proceedings of the Twelfth IEEE Symposium on Logic in Computer Science, Warsaw, Poland*.

- Bobrowski, A. (2005). *Functional Analysis for Probability and Stochastic Processes: An Introduction*. Cambridge University Press.
- Borodin, A. N. and Salminen, P. (1997). Handbook of Brownian motion-facts and formulae. *Journal of the Royal Statistical Society-Series A Statistics in Society* **160** (3) 596.
- Chen, L., Clerc, F. and Panangaden, P. (2019). Bisimulation for Feller-Dynkin processes. *Electronic Notes in Theoretical Computer Science* **347** 45–63. *Proceedings of the Thirty-Fifth Conference on the Mathematical Foundations of Programming Semantics*.
- Chen, L., Clerc, F. and Panangaden, P. (2020). Towards a classification of behavioural equivalences in continuous-time Markov processes. *Electronic Notes in Theoretical Computer Science. Proceedings of the Thirty-Sixth Conference on the Mathematical Foundations of Programming Semantics*.
- Clerc, F. (2021). *Bisimulation and Behavioural Equivalences for Continuous-Time Markov Processes*. PhD thesis, McGill University.
- Clerc, F., Fijalkow, N., Klin, B. and Panangaden, P. (2019). Expressiveness of probabilistic modal logics: A gradual approach. *Information and Computation* **267** 145–163.
- Danos, V., Desharnais, J., Laviolette, F. and Panangaden, P. (2006). Bisimulation and congruence for probabilistic systems. *Information and Computation* **204** (4) 503–523.
- Desharnais, J., Edalat, A. and Panangaden, P. (2002). Bisimulation for labeled Markov processes. *Information and Computation* **179** (2) 163–193.
- Desharnais, J. and Panangaden, P. (2003). Continuous stochastic logic characterizes bisimulation for continuous-time Markov processes. *Journal of Logic and Algebraic Programming* **56** 99–115. Special issue on Probabilistic Techniques for the Design and Analysis of Systems.
- Doberkat, E.-E. (2005). Semi-pullbacks for stochastic relations over analytic spaces. *Mathematical Structures in Computer Science* **15** (4), 647–670.
- Dudley, R. M. (1989). *Real Analysis and Probability*. Wadsworth and Brookes/Cole.
- Einstein, A. (1905). The theory of the brownian movement. *Annalen der Physik* **17** 549.
- Fijalkow, N., Klin, B. and Panangaden, P. (2017). The expressiveness of probabilistic modal logic revisited. In *Proceedings of the 44th International Colloquium on Automata Languages and Programming*.
- Karatzas, I. and Shreve, S. (2012). *Brownian motion and stochastic calculus*, vol. 113. Springer Science and Business Media.
- Larsen, K. G. and Skou, A. (1991). Bisimulation through probabilistic testing. *Information and Computation* **94** 1–28.
- Milner, R. (1980). *A Calculus for Communicating Systems*, vol. 92. Lecture Notes in Computer Science. Springer-Verlag.
- Milner, R. (1989). *Communication and Concurrency*. Prentice-Hall.
- Pachl, J. and Terraf, P. S. (2020). Semipullbacks of labelled Markov processes. *ArXiv e-prints*, 1706.02801. To be published in Logical Methods in Computer Science.
- Panangaden, P. (2009). *Labelled Markov Processes*. Imperial College Press.
- Park, D. (1981). Concurrency and automata on infinite sequences. In *Proceedings of the 5th GI Conference on Theoretical Computer Science*, vol. 104. Lecture Notes in Computer Science, pp. 167–183. Springer-Verlag.
- Chris, L., Rogers, G. and Williams, D. (2000). *Diffusions, Markov Processes and Martingales: Volume 1. Foundations*. 2nd edition, Cambridge University Press.
- Royden, H. L. and Fitzpatrick, P. M. (2010). *Real Analysis*, 4th edition. Printice-Hall Inc.
- Sangiorgi, D. (2009). On the origins of bisimulation and coinduction. *ACM Transactions on Programming Languages and Systems (TOPLAS)* **31** (4) 15.
- Terraf, P. S. (2011). Unprovability of the logical characterization of bisimulation. *Information and Computation* **209** (7) 1048–1056.
- Whitt, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and their Applications to Queues*. Springer Series in Operations Research. Springer-Verlag.

Appendix A. Proof of Proposition 11

This section is entirely based on Rogers and Williams (2000), and we will only cite the specific reference of the results that we use without citing the whole book throughout this section.

The Riesz representation theorem can be found as Theorem II.80.3 of Rogers and Williams (2000). The authors also offer the following useful extension (Theorem III.6.1):

Theorem 66. *A bounded linear functional ϕ on $C_0(E)$ may be written uniquely in the form*

$$\phi(f) = \int_E f(x) \mu(dx)$$

where μ is a signed measure on E of finite total variation.

As is stated in the Riesz representation theorem, that measure μ is inner regular. We have also found references of this theorem as Riesz-Markov-Kakutani representation theorem.

This theorem has the following corollary (Theorem III.6.2):

Corollary 67. *Suppose that $V : C_0(E) \rightarrow b\mathcal{E}$ is a (bounded) linear operator that is sub-Markov in the sense that $0 \leq f \leq 1$ implies $0 \leq Vf \leq 1$. Then, there exists a unique sub-Markov kernel (also denoted by) V on (E, \mathcal{E}) such that for all $f \in C_0(E)$ and $x \in E$*

$$Vf(x) = \int f(y) V(x, dy)$$

While the proof is left as an exercise in Rogers and Williams (2000), let us explicitly write it down:

Proof. For every x in E , write V_x for the functional $V_x(f) = Vf(x)$. This functional is bounded and linear which enables us to use Theorem 66: there exists a signed measure μ_x on E of finite total variation such that

$$V_x(f) = \int_E f(y) \mu_x(dy)$$

We claim that $V : (x, B) \mapsto \mu_x(B)$ is a sub-Markov kernel, that is

- For all x in E , $V(x, -)$ is a subprobability measure on (E, \mathcal{E})
- For all B in \mathcal{E} , $V(-, B)$ is \mathcal{E} -measurable.

The first condition directly follows from the definition of V : $V(x, -) = \mu_x$ which is a measure and furthermore $V(x, E) = \mu_x(E) = V_x(1)$ (where 1 is the constant function over the whole space E which value is 1). Using the hypothesis that V is Markov, we get that $V1 \leq 1$. We have to be more careful in order to prove that $V(x, B) \geq 0$ for every measurable set B : this is a consequence of the regularity of the measure μ_x (see Proposition 11 of Section 21.4 of Royden and Fitzpatrick (2010)). This shows that μ_x is a subprobability measure on (E, \mathcal{E}) .

Now, we have to prove that for every $B \in \mathcal{E}$, $V(-, B)$ is measurable. Recall that the set E is σ -compact: there exists countably many compact sets K_k such that $E = \bigcup_{k \in \mathbb{N}} K_k$. For $n \in \mathbb{N}$, define $E_n = \bigcup_{k=0}^n K_k$ and $B_n = B \cap E_n$.

For every $n \in \mathbb{N}$, there exists a sequence of functions $(f_j^n)_{j \in \mathbb{N}} \subset C_0(E)$ that converges pointwise to $\mathbb{1}_{B_n}$, that is for every $x \in E$, $\mu_x(B_n) = \lim_{j \rightarrow +\infty} Vf_j^n(x)$. Since the operator $V : C_0(E) \rightarrow b\mathcal{E}$, the maps Vf_j^n are measurable which means that $V(-, B_n) : x \mapsto \mu_x(B_n)$ is measurable.

Since for every $x \in E$, $\mu_x(B) = \lim_{n \rightarrow +\infty} \mu_x(B_n)$, this further means that $V(-, B) : x \mapsto \mu_x(B)$ is measurable. □

Appendix B. Details of the Proof of Theorem 58

Lemma 68. *The two quotient topological spaces $(E_1 \uplus E_3) / \sim$ and E_2 / \approx are bijective.*

Proof. First, let us construct a map $\psi : E_2 / \approx \rightarrow (E_1 \uplus E_3) / \sim$. We are going to use the universal property of the quotient. First, we define a map

$$\begin{aligned} k : E_2 &\rightarrow (E_1 \uplus E_3) / \sim \\ z &\mapsto \pi_{\sim}(i_1 \circ h(z)) \end{aligned}$$

Note that $\pi_{\sim}(i_1 \circ h(z)) = \pi_{\sim}(i_3 \circ g(z))$ by definition of the equivalence \sim .

Now, suppose that $z \approx z'$ (in E_2). We want to show that $k(z) = k(z')$. Since $z \approx z'$, there exists a sequence z_0, \dots, z_n such that $z_0 = z$, $z' = z_n$ and for every j , $g(z_{2j}) = g(z_{2j+1})$ and $h(z_{2j+1}) = h(z_{2j+2})$. This means that for every j , $i_3 \circ g(z_{2j}) = i_3 \circ g(z_{2j+1})$ and hence $k(z_{2j}) = k(z_{2j+1})$ and similarly $k(z_{2j+1}) = k(z_{2j+2})$. In particular, this shows that $k(z) = k(z_0) = k(z_n) = k(z')$.

Hence by the universal property of the quotient, there exists a unique continuous map ψ such that the following diagram commutes:

$$\begin{array}{ccc} E_2 & \xrightarrow{k} & (E_1 \uplus E_3) / \sim \\ \pi_{\approx} \downarrow & & \nearrow \psi \\ E_2 / \approx & & \end{array}$$

Second, let us construct a map $\psi' : (E_1 \uplus E_3) / \sim \rightarrow E_2 / \approx$. We are also going to use the universal property of the quotient. First, we define a map

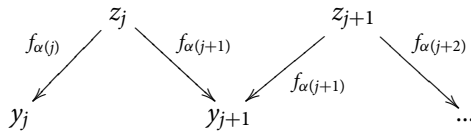
$$\begin{aligned} k' : E_1 \uplus E_3 &\rightarrow E_2 / \approx \\ i_1(x_1) &\mapsto \pi_{\approx}(z) \text{ where } z \in h^{-1}(i_1(x_1)) \\ i_3(x_3) &\mapsto \pi_{\approx}(z) \text{ where } z \in g^{-1}(i_3(x_3)) \end{aligned}$$

This is well-defined since:

- all the sets $h^{-1}(i_1(x_1))$ and $g^{-1}(i_3(x_3))$ are not empty (by surjectivity of g and h)
- if $z, z' \in h^{-1}(i_1(x_1))$, then $h(z) = h(z') = i_1(x_1)$ and hence $z \approx z'$. This means that $\pi_{\approx}(z) = \pi_{\approx}(z')$ (and similarly for the other case)

We will write $f_1 = h$ and $f_3 = g$

Now consider $x \sim x' \in E_1 \uplus E_3$. We want to show that $k'(x) = k'(x')$. The fact that $x \sim x'$ means that there exists y_0, \dots, y_n such that $y_j \in E_{\alpha(j)}$ (where $\alpha(j) = 1$ or 3) and $i_{\alpha(0)}(y_0) = x$ and $i_{\alpha(n)}(y_n) = x'$ and a sequence z_0, \dots, z_{n-1} in E_2 such that for every j , $f_{\alpha(j)}(z_j) = y_j$ and $y_{j+1} = f_{\alpha(j+1)}(z_{j+1})$.



Whether each downwards map is a g or an h depends only in which set E_1 or E_3 lies y_j . Note that these maps should really be \mapsto but the notations are already strenuous enough.

Now for such a j , we have that $y_{j+1} = f_{\alpha(j+1)}(z_j) = f_{\alpha(j+1)}(z_{j+1})$ and hence $k'(i_{\alpha(j)}(y_j)) = \pi_{\approx}(z_j) = \pi_{\approx}(z_{j+1})$ (note that for $j = 0$ and n , only one of those equalities make sense). This proves that $k'(x) = k'(i_{\alpha(0)}(y_0)) = k'(i_{\alpha(n)}(y_n)) = k'(x')$

Hence by the universal property of the quotient, there exists a unique continuous map ψ' such that the following diagram commutes:

$$\begin{array}{ccc} E_1 \uplus E_3 & \xrightarrow{\pi_{\approx}} & (E_1 \uplus E_3) / \sim \\ k' \downarrow & & \nearrow \psi' \\ E_2 / \approx & & \end{array}$$

Finally, using the uniqueness of those maps, we get that $\psi \circ \psi' = id$ and $\psi' \circ \psi = id$ which proves that the two quotient spaces are bijective. □

Lemma 69. *The space E_4 is locally compact and Hausdorff.*

Proof. For this proof, we will write π instead of π_{\approx} .

First note that the quotient map $\pi : E_2 \rightarrow E_2/\approx$ is open: indeed, for an open set U , $\pi^{-1}\pi(U) = \bigcup_{n \in \mathbb{N}} (h^{-1}hg^{-1}g(U))$. Since g is open, $g(U)$ is open and since g is continuous, $g^{-1}g(U)$ is open. Similarly, $h^{-1}h(g^{-1}g(U))$ is open.

Let us show that $\approx \subset E_2 \times E_2$ is closed. Consider a sequence (x_n, y_n) in \approx that converges to (x, y) . In particular, for every $n \in \mathbb{N}$, $\pi(x_n) = \pi(y_n)$. Furthermore, $\lim_n \pi(x_n) = \pi(x)$ and $\lim_n \pi(y_n) = \pi(y)$. By uniqueness of limit, $\pi(x) = \pi(y)$ and hence $(x, y) \in \approx$.

Since the quotient map is open and the equivalence \approx is closed, the space E_4 is Hausdorff.

Since the map π is open and the space E_2 is locally compact, the space E_4 is also locally compact. □

Lemma 70. *The maps ϕ_1 and ϕ_3 are surjective, continuous and open.*

Proof. We will show the results for $\phi_1 = i_1 \circ h : E_1 \rightarrow E_4$. First note that as the composition of two continuous maps, it is also a continuous map.

Second, consider $x_4 \in E_4$. Recall that π_{\approx} is defined as $E_2 \xrightarrow{h} E_1 \xrightarrow{i_1} E_1 \uplus E_3 \xrightarrow{h} E_4$. We know that it is a quotient map and thus surjective, that is there exists x_2 such that $\pi_{\approx}(x_2) = x_4$, and thus $\phi_1(h(x_2)) = x_4$. Hence ϕ_1 is surjective.

Finally, let us show that ϕ_1 is open. Consider U an open set in E_1 . We have that

$$\begin{aligned} \phi_1(U) &= \{y \mid \exists z \in U \ y = \phi_1(z)\} \\ &= \{y \mid \exists x \in h^{-1}(U) \ y = \phi_1(h(x))\} \quad \text{since } h \text{ is surjective} \\ &= \{y \mid \exists x \in h^{-1}(U) \ y = \pi_{\approx}(x)\} \quad \text{since } \pi_{\approx} = \phi_1 \circ h \\ &= \pi_{\approx}(h^{-1}(U)) \end{aligned}$$

Since the map h is continuous, the set $h^{-1}(U) \subset E_2$ is open. We have shown in the proof of Lemma 69 that the map π_{\approx} is open, which means that the set $\phi_1(U) \subset E_4$ is open. This concludes the proof. □

Lemma 71. *The map obs_4 is measurable.*

Proof. For this proof, we will write π instead of π_{\approx} .

It is enough to show that $obs_4^{-1}(A)$ is measurable for A a singleton in 2^{AP} . Now note that $obs_4^{-1}(A) = \pi(obs_2^{-1}(A))$. Furthermore, $obs_2^{-1}(A) = \bigcap_{i \in AP} B_i$ where B_i is either open or closed and corresponds to whether there is a 1 or a 0 at the i -th position in the singleton A . Note that since obs_2 is stable under \approx , we have that for an arbitrary set C , $\pi^{-1}\pi(obs_2^{-1}(C)) = obs_2^{-1}(C)$ and in particular $\pi^{-1}\pi(B_i) = B_i$.

- if B_i is open, then since π is open, $\pi(B_i)$ is open.
- if B_i is closed, write $U_i = B_i^c$ for its open complement.

$$\begin{aligned} [\pi(U_i)]^c &= [\pi(B_i^c)]^c \\ &= [\pi([\pi^{-1}\pi(B_i)]^c)]^c \quad \text{since } \pi^{-1}\pi(B_i) = B_i \\ &= [\pi\pi^{-1}([\pi(B_i)]^c)]^c \quad \text{since inverse images commute with complements} \\ &= [[\pi(B_i)]^c]^c \quad \text{since } \pi\pi^{-1} = id \\ &= \pi(B_i) \end{aligned}$$

Since $\pi(U_i)$ is open, we have that $\pi(B_i)$ is closed.

This shows that in both cases, $\pi(B_i)$ is measurable. Since

$$\begin{aligned} obs_4^{-1}(A) &= \pi(obs_2^{-1}(A)) \\ &= \pi\left(\bigcap_{i \in AP} B_i\right) \\ &= \pi\left(\bigcap_{i \in AP} \pi^{-1}\pi(B_i)\right) \quad \text{since } \pi^{-1}\pi(B_i) = B_i \\ &= \pi\pi^{-1}\left(\bigcap_{i \in AP} \pi(B_i)\right) \quad \text{since inverse images commute with countable intersections} \\ &= \bigcap_{i \in AP} \pi(B_i) \quad \text{since } \pi\pi^{-1} = id \end{aligned}$$

we know that $obs_4^{-1}(A)$ is measurable. Note that we have also shown that an atomic proposition partitions the space E_4 into one open set and one closed set. \square

Lemma 72. *The probability distribution \mathbb{P}_4 on Ω_4 is well-defined and*

$$\mathbb{P}_4^{x_4}(B_4) = \mathbb{P}_2^{x_2}(B_2) = \mathbb{P}_3^{x_3}(B_3) = \mathbb{P}_1^{x_1}(B_1)$$

Proof. Let us start with the measurability of the sets B_1, B_2 and B_3 . Those proofs are very similar so we will only do the case for $B_2 = \{\omega \in \Omega_2 \mid \pi_{\approx} \circ \omega \in B_4\}$. The proof is done by induction on the structure of B_4 .

- First, if $B_4 = (X_5^4)^{-1}(C) = \{\omega \in \Omega_4 \mid \omega(s) \in C\}$ for $C \in \mathcal{E}_4$, then $B_2 = (X_5^2)^{-1}(\pi_{\approx}^{-1}(C))$. And since π_{\approx} is continuous, we know that $\pi_{\approx}^{-1}(C) \in \mathcal{E}_2$ and hence $B_2 \in \mathcal{G}_2$.
- If $B_4 = A_4^C$ with $A_4 \in \mathcal{G}_4$ and $A_2 = \{\omega \in \Omega_2 \mid \pi_{\approx} \circ \omega \in A_4\} \in \mathcal{G}_2$, then $B_2 = \Omega_2 \setminus A_2$. And since $A_2 \in \mathcal{G}_2$, we get that $B_2 \in \mathcal{G}_2$.
- If $B_4 = \bigcup_{i \in \mathbb{N}} A_i$ where for every $i \in \mathbb{N}$, $A_i \in \mathcal{G}_4$ and $A'_i = \{\omega \in \Omega_2 \mid \pi_{\approx} \circ \omega \in A_i\} \in \mathcal{G}_2$, then $B_2 = \bigcup_{i \in \mathbb{N}} A'_i$. And since $A'_i \in \mathcal{G}_2$ for every $i \in \mathbb{N}$, we get that $B_2 \in \mathcal{G}_2$.

We can now show that if $\pi_{\approx}(x_2) = \pi_{\approx}(y_2)$, then $\mathbb{P}_2^{x_2}(B_2) = \mathbb{P}_2^{y_2}(B_2)$. If $x_2, y_2 \in E_2$ are such that $\pi_{\approx}(x_2) = \pi_{\approx}(y_2)$, then this means that there exists z^0, \dots, z^n such that $x_2 = z^0, y_2 = z^n$ and for every $i = 0$ to $n - 1$, either $g(z^i) = g(z^{i+1})$ or $h(z^i) = h(z^{i+1})$. Remember that h is an FD-homomorphism, which means that for every $j = 0$ to n ,

$$\mathbb{P}_1^{h(z^j)}(B_1) = \mathbb{P}_2^{z^j}(\hat{B})$$

where $\hat{B} = \{\omega \in \Omega_2 \mid h \circ \omega \in B_1\}$. Rewriting this, we get that $\hat{B} = \{\omega \in \Omega_2 \mid \phi_1 \circ h \circ \omega \in B_4\} = B_2$. In particular, we have that if $h(z^i) = h(z^{i+1})$, then $\mathbb{P}_2^{z^i}(B_2) = \mathbb{P}_1^{h(z^i)}(B_1) = \mathbb{P}_2^{z^{i+1}}(B_2)$. We have a similar result for g , and therefore,

$$\mathbb{P}_2^{x_2}(B_2) = \mathbb{P}_2^{z^0}(B_2) = \dots = \mathbb{P}_2^{z^n}(B_2) = \mathbb{P}_2^{y_2}(B_2)$$

These results show that \mathbb{P}_4 is well-defined.

Finally, let us show the last equality. We have $x_2 \in E_2$ and $x_1 \in E_1$ such that $\pi_{\approx}(x_2) = \phi_1(x_1)$. Since h is surjective, there exists $y_2 \in E_2$ such that $x_1 = h(y_2)$. This implies that $\pi_{\approx}(x_2) = \pi_{\approx}(y_2)$. Using previously shown results,

$$\mathbb{P}_2^{x_2}(B_2) = \mathbb{P}_2^{y_2}(B_2) \quad \text{and} \quad \mathbb{P}_1^{h(y_2)}(B_1) = \mathbb{P}_2^{y_2}(B_2)$$

that is $\mathbb{P}_1^{x_1}(B_1) = \mathbb{P}_2^{x_2}(B_2)$. This concludes this proof. \square