Upper-Expectation Bisimilarity and Łukasiewicz μ -Calculus

Matteo Mio

Computer Laboratory, University of Cambridge, UK

Abstract. Several notions of bisimulation relations for probabilistic nondeterministic transition systems have been considered in the literature. We study a novel testing-based behavioral equivalence called *upper expectation bisimilarity* and, using standard results from functional analysis, we develop its coalgebraic and algebraic theory and provide a logical characterization in terms of an expressive probabilistic modal μ -calculus.

1 Introduction

Directed-graph structures are sufficient for modeling nondeterministic programs and concurrent systems but cannot directly represent other important aspects of computation, such as *probabilistic behavior*, *timed transitions* and other *quantitative* information one might need to express. *Probabilistic nondeterministic transition systems* (PNTS), also known in the literature as (simple) Segala systems [2], concurrent Markov chains and probabilistic automata [29], have been identified in the last two decades as convenient mathematical structures, generalizing standard nondeterministic transition systems (NTS), to provide operational semantics to probabilistic nondeterministic languages (see, e.g., [2,30]).

A central concept in the theory of programming languages and concurrent systems is the notion of *behavioral equivalence*. An equivalence relation $\simeq \subseteq S \times S$ between states of a system is, informally speaking, a behavioral equivalence if $s \simeq t$ implies that s and t satisfy the same class of properties of interest. Of course different classes of properties induce different notions of equivalence. The paradigmatic example of behavioral equivalence for ordinary NTS's is Milner and Park's *bisimilarity* [21]. Among its good properties, bisimilarity enjoys the following: **B1**) two states are bisimilar if and only if they satisfy the same properties expressed, e.g., in Kozen's modal μ -calculus [17] or in other (weaker but useful in practice) branching-time logics such as CTL, CTL* [1] and the basic modal logic **K** or its labeled version, the Hennessy–Milner modal logic. Furthermore, **B2**) bisimilarity is a congruence for a wide family of process algebras, all specified following one of the many rule formats (e.g., GSOS, tyft/tyxt, *etc*) and **B3**) enjoys a rich mathematical theory based on coinduction, playing a role of paramount importance in *coalgebra* [28]. Lastly, but importantly, **B4**) bisimilarity can also be explained in terms of Milner's standard metaphor of *push-button experiments* on systems [21]. Such experiments provide an abstract, yet intuitive, testing semantics for bisimilarity.

In the context of PNTS's several notions of behavioral equivalence, based on the technical machinery of coinduction, have been considered in the literature [29,20,31,3,9,10]. The one which has attracted most attention so far was introduced by Segala in [29] and is referred to in this paper as *standard bisimilarity*. This is a mathematically natural notion (cf. B3) and, indeed, it has been rediscovered using the methods of coalgebra [30]. In [29] another behavioral equivalence (strictly coarser than standard bisimilarity), which we refer to as *convex bisimilarity*, is introduced by Segala as a preferable notion because, as we shall discuss in Section 2, it is strongly motivated by the concept of *probabilistic scheduler*. Importantly, both equivalences are congruence relations (cf. B2) for the wide class of PGSOS process algebras, which includes virtually all Milner's CCS-style (probabilistic) process operators of practical interest [2].

Following a traditional approach based on interpreting formulas as *sets* of states (or probability distributions over states [10]), several Hennessy-Milner-style modal logics, capable of characterizing either standard or convex bisimilarity have been investigated (see, e.g., [29,20,10,11]). However all these logics (even when enriched with fixed-point operators as in [10]) lack the expressive power required to formulate many natural and practically useful properties of PNTS's such as, e.g., those expressible in the popular temporal logics PCTL and PCTL* of [4]. Unlike their non-probabilistic counterparts, CTL and CTL* [1], these two logics induce non-standard and different behavioral equivalences [31]:

$$\text{standard bisimilarity} \subsetneq \text{convex bisimilarity} \subsetneq \text{PCTL}^* \subsetneq \text{PTCL}$$

Despite significant efforts, the field still lacks a satisfactory logical framework (cf. B1) for expressing practically useful properties of systems and, at the same time, reasoning about behavior up-to satisfactory notions of behavioral equivalence. The problem is further complicated by the lack of a (widely accepted) testing semantics (cf. B4) for either standard bisimilarity, convex bisimilarity or any of the other proposed equivalences studied in the literature (e.g., [11,6,3]).

Starting from the late 90's, following an alternative *quantitative* approach based on interpreting formulas as maps $\llbracket \phi \rrbracket : S \rightarrow \mathbb{R}$ from states to real values, fixed-point modal logics (quantitative μ -calculi) for PNTS's have been studied (see, e.g., [15,9,26]). The present paper contributes to a programme of ongoing research, one of whose overall aims is to investigate the extent to which quantitative μ -calculi play as fundamental a rôle in the probabilistic-nondeterministic setting as that of Kozen's μ -calculus in the nondeterministic setting.

Following the quantitative approach, the author has recently introduced in [22,24] the first fixed-point based logic, called Łukasiewicz μ -calculus ($\mathbb{L}\mu$), capable of encoding PCTL. It has been shown in [22,23] how this (and similar) logic can be given a *game semantics* generalizing the well-known parity-game semantics of Kozen's μ -calculus. Compositional verification techniques have been developed in [25], exploiting the elegant fixed-point semantics. Most recently, model checking algorithms have been investigated in [24].

Contributions. This paper adds to the mounting evidence that the $\mathbb{R}$ -valued approach to temporal logics for PNTS's is a convenient framework by showing that

$L\mu$ captures what is arguably the preferred notion of bisimilarity for PNTS's: convex bisimilarity. A strong novelty of our approach is that to obtain this result one needs to take a fresh perspective on convex bisimilarity. To this end, a nonstandard behavioral equivalence called *upper expectation* (UE) *bisimilarity*, which naturally arises from a very abstract testing scenario based on $\mathbb{R}$ -valued experiments (cf. property **B4**), is introduced. Unlike other related works, our experiments are not given by, e.g., the formulas of a given logic (such as the μ -calculus of [9]) nor by terms of some process algebra (see, e.g., [11]) but, instead, are modeled abstractly by functions $f: S \rightarrow \mathbb{R}$ from program states to real values.

We prove that UE-bisimilarity coincides with the behavioral equivalence induced by the expressive logic $L\mu$ (cf. property **B1**). Furthermore, we show that this logic characterizes the so-called *behavioral* (*Hausdorff*) *metric*, a concept of fundamental importance in the theory of approximation of probabilistic systems pioneered by Panangaden [27]. Thus the $\mathbb{R}$ -valued approach to logics for PNTS's may be considered equally suitable as a mechanism for characterizing process equivalence as other non-quantitative logics advocated for this specific purpose.

As a main result we prove that, while UE-bisimilarity is generally coarser than convex bisimilarity, the two notions coincide on a wide class of systems:

$$\text{convex bisimilarity} = \text{UE-bisimilarity} = \text{\text{Łukasiewicz } \mu\text{-Calculus}}$$

As a matter of fact, we argue that UE-bisimilarity constitutes a natural relaxation of convex bisimilarity based on elementary topological considerations. Indeed, a very interesting feature of our work is that results from linear algebra and functional analysis are crucial to link UE-bisimilarity with the topological notions of *closedness*, *compactness* and *convexity*. This, in turn, allows the identification of UE-bisimilarity with the *coalgebraic* notion of cocongruence of the appropriate functor type (cf. property **B3**) and to relate it with Segala's convex bisimilarity.

In Section 5 we present two algebraic results about UE-bisimilarity: a sound and complete equational characterization of the behavior of PNTS's, obtained once again by applying a known representation theorem from functional analysis, and a congruence result with respect to the CCS-style process-algebra operators definable in the PGSOS format of [2] (cf. property **B2**).

Collectively our results provide strong mathematical foundations for UE-bisimilarity comparable with those of Milner and Park's bisimilarity and shed light on the mathematical nature of PNTS's and on the quantitative approach to temporal logics for PNTS's.

2 Coalgebra and PNTS's

We employ the basic language of coalgebra in the description of systems and behavioral equivalences. We refer to [28] for a gentle introduction to the subject.

Definition 1. *Let $\mathbf{Set}$ be the category of sets and functions between them. The endofunctor $\mathcal{P}$ (powerset) on $\mathbf{Set}$ is defined as: $\mathcal{P}(X) = \{A \mid A \subseteq X\}$ and $(\mathcal{P}(f))(A) = f[A] = \{f(x) \mid x \in A\}$, for all sets X, Y and functions $f: X \rightarrow Y$.*

Definition 2. A nondeterministic transition system (NTS) is a F -coalgebra $(X, \alpha : X \rightarrow F(X))$ of the functor $F = \mathcal{P}$.

In applications one most often encounters *labeled* NTS's, or *Kripke structures*, i.e., NTS's endowed with a set P of propositional letters $p \in P$ interpreted as predicates. For the sake of simplicity we just consider plain NTS's and their probabilistic generalizations. The results we develop extend straightforwardly to the labeled and propositional extensions.

Coalgebra provides an abstract notion of behavioral equivalence of F -coalgebras called *cocongruence* which is based on the idea of lifting relations on the set X of states to relations on $F(X)$ [18]. The notion of cocongruence for the functor $\mathcal{P}$ coincides with ordinary Milner and Park's bisimulation for NTS's [18].

Definition 3 ([30]). The endofunctor $\mathcal{D}$ (discrete probability distributions) on **Set** is defined as: $\mathcal{D}(X) = \{\mu : X \rightarrow [0, 1] \mid \sum_x \mu(x) = 1\}$, and $(\mathcal{D}(f))(\mu) = f[\mu] = y \mapsto \mu(f^{-1}(y))$ where, for all sets $S \subseteq X$, we define $\mu(S) = \sum_{x \in S} \mu(x)$.

Note that the composite functor $\mathcal{PD}$ maps a set X to the collection of all sets of discrete probability distributions on X .

Definition 4. A probabilistic nondeterministic transition system (PNTS) is a $\mathcal{PD}$ -coalgebra $(X, \alpha : X \rightarrow \mathcal{PD}(X))$. We write $x \rightarrow \mu$ to specify that $\mu \in \alpha(x)$.

The intended interpretation is that the system, at some state $x \in X$, can evolve by nondeterministically choosing one of the *accessible* distributions $\mu \in \alpha(x)$ and then continuing its execution from the state $y \in X$ with probability $\mu(y)$. PNTS's can be visualized, using graphs labeled with probabilities. For example the PNTS (X, α) having set of states $X = \{x, x_1, x_2\}$ and transition map $\alpha(x) = \{\mu_1, \mu_2\}$ and $\alpha(x_1) = \alpha(x_2) = \emptyset$, with $\mu_1(x_1) = \mu_2(x_2) = 0.2$ and $\mu_1(x_2) = \mu_2(x_1) = 0.8$, can be depicted as in Figure 1. The combination of nondeterministic choices immediately followed by probabilistic ones, allows the modeling of concurrent probabilistic programming languages in a natural way [2].

PNTS's with a definition equivalent to the coalgebraic one given above were introduced by Segala [29] who also defined two notions of bisimilarity for PNTS's. We refer to *standard bisimilarity* as the stronger (i.e., finer) of the two:

Definition 5 (Standard Bisimulation). Given a PNTS (X, α) , a standard bisimulation is an equivalence relation $E \subseteq X \times X$ such that if $(x, y) \in E$ then

- if $x \rightarrow \mu$ then there exists ν such that $y \rightarrow \nu$ and $(\mu, \nu) \in \hat{E}$, and
- if $y \rightarrow \nu$ then there exists μ such that $x \rightarrow \mu$ and $(\mu, \nu) \in \hat{E}$,

where $(\mu, \nu) \in \hat{E}$ holds if $\mu(A) = \nu(A)$ (see Definition 3) for all sets $A \subseteq X$ which are unions of E -equivalence classes.

The definition looks technical but has a simple interpretation. Two states x, y of a PNTS (X, α) are standard-bisimilar if the two sets of reachable distributions $\alpha(x)$ and $\alpha(y)$ are equal modulo E , i.e., if $\{[\mu]_{\hat{E}} \mid \mu \in \alpha(x)\} = \{[\mu]_{\hat{E}} \mid \mu \in \alpha(y)\}$, where $(\mu, \nu) \in \hat{E}$ means that if one identifies E -related states, then μ and ν

become *equal*, i.e., they assign the same probabilities to all events $A \subseteq X/E$ (i.e., unions of equivalence classes). Using this argument one can show that the notion of standard bisimilarity coincides with that of cocongruence of $\mathcal{PD}$ -coalgebras [30]. As an example, consider the two PNTS's rooted at x and y respectively, depicted in Figure 1, and assume the processes x_1 and x_2 to be observationally different¹. The processes x and y are not standard bisimilar. This is because y can lead to a probability distribution μ_3 which cannot be matched by x .

Fig. 1. Example of states (x, y) not standard bisimilar

It has been argued by Segala (see, e.g., [29]) that standard bisimilarity is too strict a behavioral equivalence when PNTS's are used to model nondeterministic probabilistic programs/systems. In this setting, the nondeterminism in the system is supposed to model all the possible choices which can be made by, e.g., an external *scheduler*. It is natural, however, to assume that schedulers can themselves use probabilistic methods to perform their choices. Thus, a *probabilistic scheduler* could choose to pick, from the state x , the successor distributions μ_1 and μ_2 with equal probability $\frac{1}{2}$, and consequently reach states x_1 and x_2 with equal probabilities $\frac{1}{2}(0.2 + 0.8) = 0.5$. Thus a scheduler, by choosing probabilistically between μ_1 and μ_2 , can mimic the choice of $\mu_3 = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2$.

Definition 6. Let X be a set. A convex combination of elements of $\mathcal{D}(X)$ is a distribution ν of the form $\nu(x) = \sum_{i=1}^n \lambda_i \cdot \mu_i(x)$, for $\mu_i \in \mathcal{D}(X)$, $\lambda_i \in [0, 1]$ and $\sum_i \lambda_i = 1$. The convex hull of $A \in \mathcal{PD}(X)$, denoted by $H(A)$, is the collection of all convex combinations of elements in A . The set A is convex if $A = H(A)$.

We are now ready to introduce the second notion of bisimilarity introduced by Segala [29,20], which we refer to² as *convex bisimilarity*.

Definition 7 (Convex Bisimulation). Given a PNTS (X, α) , a convex bisimulation is an equivalence relation $E \subseteq X \times X$ such that if $(x, y) \in E$ then

- if $x \rightarrow_C \mu$ then there exists ν such that $y \rightarrow_C \nu$ and $(\mu, \nu) \in \hat{E}$, and
- if $y \rightarrow_C \nu$ then there exists μ such that $x \rightarrow_C \mu$ and $(\mu, \nu) \in \hat{E}$,

where $x \rightarrow_C \mu$ holds if and only if $\mu \in H(\alpha(x))$.

¹ By this we mean that $(x_1, x_2) \notin E$ for all bisimulations E . Of course this can be implemented by adding structure to the system (e.g., a single edge from x_1 to μ_1). Our assumption thus simply abstracts away from such additional details.

² The adjective *probabilistic* is often adopted in the literature [29,31]. We prefer the adjective *convex* which transparently reflects the mathematics behind the notion.

Thus convex bisimilarity is obtained by replacing $\rightarrow$ with $\rightarrow_C$ in Definition 5. Probabilistic schedulers take the place of ordinary schedulers and can simulate any convex combination of the reachable probability distributions.

Remark 1. Note that Definition 7 remains unaltered if $\rightarrow_C$ is replaced by $\rightarrow$ (as originally proposed by Segala in [29]) in the left side of the two conditions.

If probabilistic schedulers are assumed, Example 1 shows how, as observed by Segala, standard-bisimilarity is too strict because it distinguishes between states that, under the intended interpretation, ought to be identified. This fact does not mean that cocongruence (i.e., standard bisimilarity) is not the “right” notion of behavioral equivalence for $\mathcal{PD}$ -coalgebras. Rather, it suggests that $\mathcal{PD}$ -coalgebras do not precisely model the class of systems we have in mind.

Definition 8. *The endofunctor $\mathcal{P}_c\mathcal{D}$ (convex sets of probability distributions) on **Set** is defined as follows: $\mathcal{P}_c\mathcal{D}(X) = \{A \mid A \in \mathcal{PD}(X) \text{ and } A = H(A)\}$ and $(\mathcal{D}(f))(A) = f[A] = \{f[\mu] \mid \mu \in A\}$, where $f[\mu]$ is defined as in Definition 3.*

Remark 2. The functoriality of $\mathcal{P}_c\mathcal{D}$ is well known. In fact more is true, and $\mathcal{P}_c\mathcal{D}$ carries a monad structure [32]. This property has been extensively studied, especially in the field of domain theory (see, e.g., [32,13]), and recognized as important in the setting of probabilistic-nondeterminism.

It is clear that a $\mathcal{P}_c\mathcal{D}$ -coalgebra (X, α) is just a particular kind of $\mathcal{PD}$ -coalgebra such that, for all $x \in X$, the set $\alpha(x)$ is convex. Formally (see, e.g., [30] for an introduction to this concept) $\text{id} : \mathcal{PD} \rightarrow \mathcal{PD}$ is an injective natural transformation. Furthermore the convex hull operation gives us a natural way to convert a $\mathcal{PD}$ -coalgebra into a $\mathcal{P}_c\mathcal{D}$ -coalgebra (formally, $H : \mathcal{PD} \rightarrow \mathcal{P}_c\mathcal{D}$ is a surjective natural transformation). Transforming a $\mathcal{PD}$ -coalgebra (i.e., a PNTS) (X, α) into the $\mathcal{P}_c\mathcal{D}$ -coalgebra $(X, H_X \circ \alpha)$ precisely corresponds to the substitution of the arrow relation $\rightarrow$ in Definition 5 with the relation $\rightarrow_C$ in Definition 7. It is now straightforward to verify that an equivalence relation $E \subseteq X \times X$ is a convex bisimulation in the PNTS (X, α) if and only if E is a cocongruence (i.e., standard bisimilarity) in the $\mathcal{P}_c\mathcal{D}$ -coalgebra $(X, H_X \circ \alpha)$. Thus convex bisimilarity can be seen as standard bisimilarity *modulo* (H_X is surjective) the *behavioral equation* $A \approx B$ whenever $H(A) = H(B)$, for all $A, B \in \mathcal{PD}(X)$, capturing the the behavior of probabilistic schedulers.

Remark 3. Although $\mathcal{P}_c\mathcal{D}$ -coalgebras are the models naturally corresponding to convex bisimilarity, we can always work, concretely, with ordinary PNTS’s (i.e., $\mathcal{PD}$ -coalgebras) (X, α) , perhaps represented as finite graphs, and tacitly replace α with $H_X \circ \alpha$ (i.e., $\rightarrow$ with $\rightarrow_C$). This is convenient since, generally, the convex hull of a finite set is uncountable.

The discussion carried out in this section serves to clarify that no *a priori* categorical or coalgebraic argument exists supporting a notion of behavioral equivalence in favor of another, when modeling computing systems. Coalgebra provides, e.g., the mathematically deep notion of cocongruence for a functor, but the choice of an appropriate functor is part of the modeling process.

3 Upper Expectation Bisimilarity

We saw how convex bisimilarity naturally arises from the observation that schedulers may make probabilistic choices. We also discussed how it can be understood coalgebraically in terms of cocongruences of $\mathcal{P}_c\mathcal{D}$ -coalgebras. However it is not possible to claim, on the sole basis of these facts, that convex bisimilarity is a convenient notion of behavioral equivalence for PNTS's. Probabilistic schedulers constitute a good reason to consider two convex bisimilar states as behaviorally equivalent. But it is not clear why one might want to distinguish between two states that are not convex bisimilar. We illustrate the problem by means of the simple example of Figure 2. As usual we assume the three states x_1 , x_2 and x_3 to be observationally distinct (cf. Footnote 1). The two states x and y are not convex bisimilar because μ_3 is not a convex combination of μ_1 and μ_2 . It is not

Fig. 2. Example of states (x, y) not convex bisimilar

simple, however, to find a *concrete*³ reason to distinguish between the two states. As a matter of fact, one can prove (see, e.g., [31]) that the states x and y satisfy the same properties formulated in the expressive logics PCTL and PCTL* of [4].

Remark 4. While modal logics, carefully crafted to capture convex bisimilarity (and even standard bisimilarity) can be defined [20,29,10,11], it is certainly interesting to look at the distinguishing power of popular temporal logics for PNTS's, (of which PCTL and PCTL* are main examples) capable of expressing branching properties of probabilistic concurrent systems useful in practice.

We now introduce a simple, yet realistic, experimental scenario which allows one to distinguish the two states x and y of Figure 2. Suppose we are allowed to make repeated experiments (in the sense of Milner's push-button metaphor [21]) on the PNTS's of Figure 1. After a sufficiently large number ($n \rightarrow \infty$) of experiments at, e.g., the state x , we observe that an event S (e.g., $S = \{x_1\}$, representing the occurrence of terminal state x_1) happened with *upper probability*⁴ $\frac{m}{n}$. We can then make the following reasonable assessment: the PNTS

³ This deliberately vague adjective could be seconded by, e.g., *experimental*, *testing-based*, *etc.*

⁴ Our scenario is of course based on the common frequentist interpretation of probabilities as limits of relative frequencies in a large number of trials. Technically, due to the nondeterminism involved, a (unique) limit may not exist. Thus, what is actually experimentally observed is that the sequence of relative frequencies eventually settles below an upper value (supremum limit) λ .

at state x can exhibit behavior S (i.e., end up in x_1) with *at least* probability $\frac{m}{n}$. It seems then natural to stipulate that two states x and y are equivalent if, for every event S , the state x can exhibit behavior (event) S with maximum probability λ if and only if y can. It is simple to verify (the relevant events S to be considered are the subsets of $\{x_1, x_2, x_3\}$) that the two states x and y of Figure 2 are equivalent in this sense.

However it is also realistic to assume that in the experiments carried on PNTS's, one is not just allowed to observe the occurrences of events and thus, after a sufficient number of experiments estimate their upper probabilities by means of relative frequencies. Rather, one is allowed to associate a real valued information $r_i \in \mathbb{R}$ to the outcome of each experiment i and, after n experiments, observe the average value $\frac{1}{n} \sum_{i=0}^n r_i$. This enhanced scenario is better explained by a simple example. Consider again the states x and y of Figure 2 and consider the function $g: \{x_1, x_2, x_3\} \rightarrow \mathbb{R}$ defined as $g(x_1)=0.6$, $g(x_2)=0$ and $g(x_3)=0.5$. The function g represents the experiment in the sense that if, after letting the scheduler choose a transition (i.e., pushing the button, in Milner's metaphor) the state x_i is reached, then the real number $g(x_i)$ is registered as result. Thus, for instance, if $\{x_1, x_2, x_3, x_2, x_1\}$ was the outcome of five experiments, the numerical sequence $\{0.6, 0.5, 0, 0, 0.6\}$ and its average $0.34 = \frac{1.70}{5}$ would be our observation. Note that the simpler observation of an event S can be modeled by the experiment $\chi_S: X \rightarrow \{0, 1\}$, i.e., by the characteristic function of S .

Definition 9. Let X be a set, $\mu \in \mathcal{D}(X)$ and $f: X \rightarrow \mathbb{R}$. The expected value of f under μ , written $E_\mu(f)$, is defined as $E_\mu(f) = \sum_x \mu(x)f(x)$.

The expected values of g under μ_1 , μ_2 and μ_3 are 0.38, 0.35 and 0.39, respectively, and this readily means that, for a sufficiently large number of trials, the average resulting from experiments g on state x is *necessarily* smaller or equal than 0.38, while in state y it *can be* strictly greater than 0.38 (and at most 0.39). Thus, it is *possible* that an agent, by means of a sufficiently large number of repeated experiments g , *may* be able to distinguish between the two states x and y . This discussion leads to the following definitions.

Definition 10. Let X be a set and $A \in \mathcal{PD}(X)$ a set of probability distributions on X . The upper expectation functional $ue_A: (X \rightarrow \mathbb{R}) \rightarrow \mathbb{R}$, mapping functions $X \rightarrow \mathbb{R}$ to real numbers, is defined as $ue_A(f) = \bigsqcup_{\mu \in A} E_\mu(f)$.

Thus, $ue_A(f)$ represents the maximum (supremum limit) expected value of f which *may* be achieved when choosing probability distributions in A .

Definition 11. Given a PNTS (X, α) , an upper expectation (UE) bisimulation is an equivalence relation $E \subseteq X \times X$ such that if $(x, y) \in E$ then the equality $ue_{\alpha(x)}(f) = ue_{\alpha(y)}(f)$ holds for all E -invariant functions $f: X \rightarrow \mathbb{R}$, i.e., such that if $(z, w) \in E$ then $f(z) = f(w)$.

We restrict to E -invariant experiments f following the idea that E -related (i.e., UE-bisimilar) states ought to be identified. The choice of considering upper expectations seems one-sided, but one could equally well choose lower expectations (le_A) observing that $le_A(f) = 1 - ue_A(1 - f)$. This is an instance of the

may/must duality well known in classical bisimulation theory, modal logic and concurrency theory [21]. It follows from our discussion that restricting Definition 11 to $\{0, 1\}$ -valued functions $g: X \rightarrow \{0, 1\}$ would weaken the definition to that of an equivalence relation not capable of, e.g., distinguishing the states x and y of Figure 2. This remark reveals that $\mathbb{R}$ -valued experiments ($f: X \rightarrow \mathbb{R}$) generally provide more information than just observation of events ($g: X \rightarrow \{0, 1\}$).

The definition of UE-bisimulation gives reasons for distinguishing states based on the existence of a witnessing experiment f which *may* have an expected outcome in one state which *cannot* be matched by the other state.

3.1 Relation between Convex and UE-Bisimilarity

UE-bisimilarity arises naturally from the simple testing scenario discussed above and, as we shall discuss in Section 4, enjoys a remarkable natural connection with real-valued modal logics. It is thus worth to develop its theory and compare it with that of convex bisimilarity. In this section we show that the two notions coincide on a wide class of systems by means of an alternative characterization of UE-bisimilarity. Recall that $\mathcal{P}_c\mathcal{D}$ -coalgebras (cf. Remark 3) can be thought of as $\mathcal{P}\mathcal{D}$ -coalgebras modulo the behavioral equation $A \approx B$ if $H(A) = H(B)$, for $A, B \in \mathcal{P}\mathcal{D}(X)$. Following the same idea, to understand UE-bisimilarity coalgebraically one needs to consider the behavioral equation $A \approx B$ if $ue_A = ue_B$ (pointwise equality). As it turns out, the two equations coincide under very mild conditions. In rest of the paper we restrict attention to a fairly simple (yet important) class of PNTS's, as this greatly simplifies our discussion. In Remark 5 we briefly outline how our results can be generalized.

Convention 1. *We restrict attention to PNTS's having a finite state space, i.e., $\mathcal{P}\mathcal{D}$ -coalgebras $(X, \alpha: X \rightarrow \mathcal{P}\mathcal{D}(X))$ having a finite carrier set $X = \{x_1, \dots, x_n\}$ (endowed with the discrete topology). This allows one to view the space of functions $X \rightarrow \mathbb{R}$ as the Euclidean space $\mathbb{R}^n$ and each $\mu \in \mathcal{D}(X)$ as the n -dimensional vector $\mu = [\mu(x_1), \dots, \mu(x_n)]$. Note that $\mathcal{D}(X)$ is a closed and bounded (hence compact [16]) subset of $\mathbb{R}^n$ and that $\alpha(x)$, for $x \in X$, can be infinite.*

The following is a known result from functional analysis [14, §10.2] and optimization theory which generalizes to the setting described in Remark 5.

Theorem 1. *Let X be a finite set and $A, B \in \mathcal{P}\mathcal{D}(X)$. Then $ue_A = ue_B$ iff $\text{cl}(H(A)) = \text{cl}(H(B))$, where $\text{cl}(C)$ denotes the topological closure of the set C .*

It is a standard result in linear algebra that the closure of a convex set is itself convex (see, e.g., [19, §8.4]). For this reason the set $\text{cl}(H(A))$ is called the *closed convex hull* of A . In what follows we denote with $\bar{H}$ the operation $A \mapsto \text{cl}(H(A))$.

The result of Theorem 1 can be used to prove the following alternative characterization of UE-bisimulation (cf. Definition 7).

Theorem 2. *Given a PNTS (X, α) , an equivalence relation $E \subseteq X \times X$ is a UE-bisimulation iff for all $(x, y) \in E$, it holds that:*

- if $x \rightarrow_{CC} \mu$ then there exists ν such that $y \rightarrow_{CC} \nu$ and $(\mu, \nu) \in \hat{E}$, and
- if $y \rightarrow_{CC} \nu$ then there exists μ such that $x \rightarrow_{CC} \mu$ and $(\mu, \nu) \in \hat{E}$,

where $x \rightarrow_{CC} \mu$ if and only if $\mu \in \overline{H}(\alpha(x))$.

Thus UE-bisimilarity can be obtained by replacing $\rightarrow_C$ with $\rightarrow_{CC}$ in Definition 7 of convex bisimilarity. As a consequence UE-bisimilarity can be strictly coarser than convex bisimilarity. The following proposition, however, reveals that the two notions coincide for a wide class of PNTS's.

Proposition 1. *Let (X, α) be a PNTS such that $\alpha(x)$ is closed for all $x \in X$. Then $E \subseteq X \times X$ is a convex bisimulation if and only if it is a UE-bisimulation.*

Restricting to PNTS's of this kind can hardly be seen as a limitation in concrete applications. Every finite set is closed, thus every PNTS representable as a finite graph satisfies Convention 1 and the closedness condition of Proposition 1. Furthermore, the restriction to closed sets seems natural as agrees with the well established *motto* “observable properties are open sets” suggesting that only those sets A and B which can be separated by open sets should be distinguished.

Remark 5. Among the possible ways of relaxing the constraint of Convention 1, a very general (and mathematically convenient) approach consists in modeling PNTS's as $\mathcal{PD}$ -coalgebras in the category of compact Hausdorff topological spaces. Here $\mathcal{D}$ maps the space X to the space of probability measures on X , endowed with the weak*-topology, and $\mathcal{P}$ maps X to the space of its closed (hence compact⁵) subsets, endowed with the Vietoris topology [16]. Experiments on X are now modeled by the Banach space $C(X)$ of continuous functions $X \rightarrow \mathbb{R}$ [19]. Due to the lack of space, we just mention in this remark that it possible to define the functor $\mathcal{P}_{cc}\mathcal{D}$ mapping X to its set of convex closed subsets and show that UE-bisimilarity coincides with the notion of cocongruence of $\mathcal{P}_{cc}\mathcal{D}$ -coalgebras. All the results presented in this paper generalize to this setting at the cost of mathematical complications needed to deal with infinite dimensional spaces.

As discussed in Section 2, the notion of probabilistic scheduler gives good reasons for considering two convex bisimilar states as behaviorally equivalent. On the other hand UE-bisimilarity provides witnesses (experiments) $f : X \rightarrow \mathbb{R}$ which can be used to distinguish states that are not UE-bisimilar (cf. example of Figure 2). Thus the two *a posteriori* equivalent (under the mild closedness assumption) viewpoints complement each other in a nice way.

Remark 6. In the significantly different setting of two-player stochastic concurrent games, a behavioral equivalence called *game bisimilarity* is discovered in [9] as the kernel of a behavioral metric d induced by a quantitative $[0, 1]$ -valued fixed-point logic $\text{qL}\mu$. Up to the necessary modifications, game bisimilarity can be shown to coincide with UE-bisimilarity. The authors of [9] argue in favor of game bisimilarity on the basis of the naturalness of the logic $\text{qL}\mu$ (which is a

⁵ A closed subset of compact set is itself compact [16]. From the modeling point of view, this provides a natural topological generalization of finite-branchingness.

weak logic not capable of encoding PCTL). Our explanation in terms of the simple metaphor of $\mathbb{R}$ -valued experiments is perhaps useful in clarifying and further motivating this notion using logic-free arguments. No connection between game bisimilarity and convexity is discussed in [9]. See also Section 4.1 below.

4 Real Valued Modal Logics and Łukasiewicz μ -Calculus

The result of Theorem 1 states that the closed convex hull operation $A \mapsto \overline{H}(A)$ and the upper expectation $A \mapsto ue_A$ are essentially the same operation.

Proposition 2 ([14]). *Let X be a finite set and $A \in \mathcal{PD}(X)$. Then $\overline{H}(A) = \{\mu \in \mathcal{D}(X) \mid \forall f: X \rightarrow \mathbb{R}. (E_\mu(f) \leq ue_A(f))\}$.*

Thus from the functional ue_A it is possible to construct $\overline{H}(A)$ and, by Theorem 1, from $\overline{H}(A)$ one obtains $ue_A = ue_{\overline{H}(A)}$. Hence, the transition map α of a PNTS (X, α) , with $\alpha(x)$ convex closed for all $x \in X$ (i.e., a $\mathcal{P}_{cc}\mathcal{D}$ -coalgebra, cf. remarks 3 and 5 and Theorem 2), can be seen both as a function $(x \mapsto \alpha(x))$ of type $X \rightarrow \mathcal{P}_{cc}\mathcal{D}(X)$ and as a function $(x \mapsto ue_{\alpha(x)})$ of type $X \rightarrow ((X \rightarrow \mathbb{R}) \rightarrow \mathbb{R})$. Equivalently (by currying) the transition map α can alternatively be seen as the function transformer $\diamond_\alpha : (X \rightarrow \mathbb{R}) \rightarrow (X \rightarrow \mathbb{R})$ defined as:

$$(\diamond_\alpha f)(x) = ue_{\alpha(x)}(f) \stackrel{\text{Def 10}}{=} \bigsqcup_{x \rightarrow \mu} E_\mu(f) \quad (1)$$

It is remarkable here that the function transformer $\diamond_\alpha$ happens to coincide with the interpretation of the *diamond modality* in all $\mathbb{R}$ -valued modal logics for PNTS's in the literature [15,26,8,23,12]. The semantics of a formula ϕ of these logics, interpreted on a PNTS (X, α) , is a function $\llbracket \phi \rrbracket : X \rightarrow \mathbb{R}$ and, in particular, $\llbracket \diamond \phi \rrbracket = \diamond_\alpha(\llbracket \phi \rrbracket)$. While it is obvious that the PNTS (X, α) induces $\diamond_\alpha$ (just as in the definition), it is far from clear that from $\diamond_\alpha$ one can reconstruct the $\mathcal{P}_{cc}\mathcal{D}$ -coalgebra (X, α) . The fact that the core of a $\mathbb{R}$ -valued logic (i.e., the interpretation of the basic modality $\diamond$) automatically arises from the very elementary observation (motivating UE-bisimilarity) that $\mathbb{R}$ -valued experiments on PNTS's are useful and, due to their greater observational power (cf. example of Figure 2) should replace ordinary Boolean predicates, sheds light on the mathematical nature of the quantitative approach to logics for PNTS's.

Following our discussions, *formulas* can then be thought of as *experiments* and the value $\llbracket \diamond \phi \rrbracket(x)$ as the maximal expected value of experiment ϕ performed after “pushing the button” (in Milner’s terminology) at state x . E.g., the experiment distinguishing the states x and y of Figure 2 corresponds to the formula $\diamond \phi$, for some ϕ crafted in such a way that $\llbracket \phi \rrbracket(x_1) = 0.6$, $\llbracket \phi \rrbracket(x_2) = 0$ and $\llbracket \phi \rrbracket(x_3) = 0.5$.

The many concrete $\mathbb{R}$ -valued modal logics in the literature are obtained by considering other connectives (which can then be thought of as ways of compositionally constructing complex experiments from simpler ones) to be used in combination with $\diamond$. For example the constant $\underline{1}$ ($\llbracket \underline{1} \rrbracket(x) = 1$) and the connective $\sqcup$ ($\llbracket \phi \sqcup \psi \rrbracket(x) = \max\{\llbracket \phi \rrbracket(x), \llbracket \psi \rrbracket(x)\}$) are considered in all the logics we

are aware of. Different choices of connectives have distinct advantages over each other in terms of, e.g., model checking complexity, expressivity, game semantics, compositional reasoning methods, *etc.* In this paper we focus on the recently introduced quantitative fixed-point logic of [24,22], called Łukasiewicz μ -calculus ($L\mu$), because (unlike the other quantitative logics cited above) it is sufficiently expressive to encode other popular temporal logics for PNTS, such as PCTL. Here we limit ourselves to the basic definition of $L\mu$ and we refer to [24,25,22] for motivational discussions and theoretical results about $L\mu$ including: model checking algorithms, game semantics and proof systems for verification.

Definition 12 (Syntax). *Formulas are generated by the following grammar: $\phi ::= \Diamond\phi \mid v \mid \perp \mid \lambda\phi \mid \neg\phi \mid \phi \sqcup \phi \mid \phi \oplus \phi \mid \mu v.\phi$, where $\lambda \in [0, 1] \cap \mathbb{Q}$, v ranges over a countable set $\mathbf{Var}$ of variables and (as usual in fixed point logics) bound variables must occur under the scope of an even number of negations.*

Definition 13 (Semantics). *Given a PNTS (X, α) and an interpretation $\rho : \mathbf{Var} \rightarrow (X \rightarrow [0, 1])$ of the variables, the semantics of ϕ is defined as the map $\llbracket \phi \rrbracket_\rho : X \rightarrow [0, 1]$ defined by structural induction on formulas as:*

$$\begin{aligned} &\bullet \llbracket \Diamond\phi \rrbracket_\rho = \Diamond_\alpha(\llbracket \phi \rrbracket_\rho) & \bullet \llbracket v \rrbracket_\rho = \rho(v) & \bullet \llbracket \phi \sqcup \psi \rrbracket_\rho(x) = \max\{\llbracket \phi \rrbracket_\rho(x), \llbracket \psi \rrbracket_\rho(x)\} \\ &\bullet \llbracket \lambda\phi \rrbracket_\rho(x) = \lambda \cdot \llbracket \phi \rrbracket_\rho(x) & \bullet \llbracket \perp \rrbracket_\rho(x) = 1 & \bullet \llbracket \phi \oplus \psi \rrbracket_\rho(x) = \min\{1, \llbracket \phi \rrbracket_\rho(x) + \llbracket \psi \rrbracket_\rho(x)\} \\ &\bullet \llbracket \neg\phi \rrbracket_\rho(x) = 1 - \llbracket \phi \rrbracket_\rho(x) & \bullet \llbracket \mu v.\phi \rrbracket_\rho = \text{lfp}(f \mapsto \llbracket \phi \rrbracket_{\rho[f/v]}) \end{aligned}$$

where lfp denotes the least fixed point operator and $\rho[f/v]$ denotes the update of the interpretation ρ at variable v defined as expected [24].

The operations of $L\mu$, except the modality $\Diamond(-)$ and the fixed-point operator $\mu v.(-)$, are the operations of Łukasiewicz (fuzzy) logic, the logic of MV-algebras.

The logic $L\mu$, as all other $\mathbb{R}$ -valued logics for PNTS's we are aware of, is *sound* (or *adequate*) with respect to UE-bisimilarity.

Theorem 3 (Soundness). *Let (X, α) be a PNTS and $E \subseteq X \times X$ a UE-bisimulation. For all $L\mu$ formulas ϕ and $(x, y) \in E$ it holds that $\llbracket \phi \rrbracket(x) = \llbracket \phi \rrbracket(y)$.*

Unlike most other logics, however, $L\mu$ enjoys the following strong property which generalizes to the setting described in Remark 5.

Theorem 4 (Denseness of $L\mu$). *Let (X, α) be a PNTS and $E \subseteq X \times X$ a UE-bisimulation. The set of functions $\{\llbracket \phi \rrbracket \mid \phi \text{ is a closed } L\mu \text{ formula}\}$ is dense (w.r.t the sup-norm on $X \rightarrow \mathbb{R}$) in the set of functions $X \rightarrow [0, 1]$ with are E -invariant (cf. Definition 11). The same result holds even if the fixed-point free fragment of $L\mu$ is considered.*

This means that for every $\epsilon > 0$ and every experiment $f : X \rightarrow [0, 1]$ which cannot distinguish between UE-bisimilar states, there exists a closed (fixed-point free) $L\mu$ formula ϕ such that for all $x \in X$, $|f(x) - \llbracket \phi \rrbracket(x)| < \epsilon$. Thus, up-to approximation to an arbitrary degree of precision, the formulas of $L\mu$ syntactically capture the set of experiments on PNTS's that respect UE-bisimilarity. The following completeness (or *expressivity*) result is a simple consequence of Theorem 4.

Proposition 3 (Completeness). *Let (X, α) be a PNTS and $x, y \in X$ be two states which are not UE-bisimilar. Then there exists some (fixed-point free) $L\mu$ formula ϕ such that $\llbracket \phi \rrbracket(x) \neq \llbracket \phi \rrbracket(y)$.*

4.1 Behavioral Metrics

Behavioral metrics for probabilistic systems, first investigated by Panangaden in the context of Markov processes [27], are based on the intuition that small changes in the probabilities of the system corresponds, roughly speaking, to small changes in behavior. Thus one might wish to relate behavior between states not in terms of equivalence relations but rather in terms of metrics⁶ $d: X \times X \rightarrow [0, 1]$ capturing *how much* two states behave differently. This approach generalizes that based on ordinary equivalence relations, as one can always consider the kernel $\ker(d) = \{(x, y) \mid d(x, y) = 0\}$ of the metric, and proved to be useful in developing a general theory of *approximation* for probabilistic systems [27]. In the context of PNTS's a robust notion of behavioral metric is given by the so-called Hausdorff metric which is based on the following idea. For a given PNTS (X, α) , and for any⁷ metric m on $\mathcal{D}(X)$, one can consider the Hausdorff metric d_H^m on the space of (compact) closed subsets of $\mathcal{D}(X)$. Then it is natural to define a metric on states as $d(x, y) = d_H^m(\alpha(x), \alpha(y))$.

An elegant characterization of the Kantorovich-Hausdorff metric d_H^K (where K is the Kantorovich metric on $\mathcal{D}(X)$ [16,27]) using the logic $\mathbf{qL}\mu$ of [9] (cf., Remark 6) can be obtained as $d_H^K = \sup \{ |\llbracket \phi \rrbracket(x) - \llbracket \phi \rrbracket(y)|, \phi \in \mathbf{qL}\mu \}$. This relies on the connectives of $\mathbf{qL}\mu$ carefully chosen so that (following earlier ideas of Panangaden [27]) denotations of formulas are Lipschitz (not expansive) functions. This, however, implies that $\mathbf{qL}\mu$ does not satisfy the denseness property of $\mathbf{L}\mu$ and, as a consequence, cannot express PCTL properties. Since the logic $\mathbf{qL}\mu$ is readily a fragment of $\mathbf{L}\mu$, the Łukasiewicz μ -calculus can be used to characterize the metric d_H^K . However it is immediate to observe that, because of the denseness and completeness properties of $\mathbf{L}\mu$, the distance logically defined as above with ϕ ranging over all $\mathbf{L}\mu$ formulas is necessarily trivial (i.e., $d(x, y) = 0$ if x and y are UE-bisimilar and 1 otherwise). The following result shows how the TV-Hausdorff metric d_H^{TV} (where TV is the Total Variation metric on $\mathcal{D}(X)$ [16]) admits an elegant characterization in terms of Diamond-guarded $\mathbf{L}\mu$ formulas.

Theorem 5. *Let (X, α) be a PNTS quotiented by UE-bisimilarity (cf. Remark 6) and assume without loss of generality (cf. Theorem 1) that $\alpha(x)$ is convex closed for every $x \in X$. Define $d_L(x, y) = \sup \{ |\llbracket \Diamond \phi \rrbracket(x) - \llbracket \Diamond \phi \rrbracket(y)|, \phi \in \mathbf{L}\mu \}$. Then $d_L(x, y) = d_H^{TV}(\alpha(x), \alpha(y))$.*

Proof. The proof exploits a result about convex closed sets from functional analysis [19, §8.4] and the denseness property of $\mathbf{L}\mu$ (Theorem 4) in the sup-norm.

The definition of a behavioral metric for PNTS's based on the TV-metric on $\mathcal{D}(X)$ appears to be novel and seems worth of further investigation. Indeed the

⁶ Technically, d is a pseudo-metric as states $x \neq y$ with $d(x, y) = 0$ are admitted. The function d becomes an authentic metric on the state space quotiented by $\ker(d)$.

⁷ Many natural metrics on $\mathcal{D}(X)$ exist. E.g., any norm on $\mathbb{R}^n$ (cf. Convention 1) induces a metric on $\mathcal{D}(X)$. Since $\mathbb{R}^n$ is finite dimensional, they all induce the same topology. The sup-norm on $\mathbb{R}^n$ (cf. Theorem 4) induces the *total variation* metric $T(\mu, \nu) = \max_x \{ |\mu(x) - \nu(x)| \}$ [16,19].

TV-metric is well known in probability theory and has important applications in statistics. Furthermore, it is natural to interpret $d_L(x, y)$ as a value related to the probability of distinguishing between x and y in a *one-shot* experiment ϕ , modeled by the formula $\Diamond\phi$ (“push the button” once and perform experiment ϕ). We plan to investigate this viewpoint, and the possible relations with the (information-theoretic) one-shot attack models of [5], in future research.

5 Algebraic Aspects of UE-Bisimilarity

As discussed in Section 4, the transition map $\alpha : X \rightarrow \mathcal{P}_{cc}\mathcal{D}(X)$ of a $\mathcal{P}_{cc}\mathcal{D}$ -coalgebra (X, α) can be equivalently presented as the function transformer $\Diamond_\alpha : (X \rightarrow \mathbb{R}) \rightarrow (X \rightarrow \mathbb{R})$ defined as in Equation 1. Clearly, not all transformers F of this type can be representations of $\mathcal{P}_{cc}\mathcal{D}$ -coalgebras (e.g., $F(x \mapsto 0) = x \mapsto 1$). The following theorem, based on a type of Riesz representation theorem for convex closed sets from functional analysis [14], provides a precise characterization of those function transformers F that arise from $\mathcal{P}_{cc}\mathcal{D}$ -coalgebra.

Theorem 6. *Let (X, α) be a $\mathcal{P}_{cc}\mathcal{D}$ -coalgebra. Denote with $\sqsubseteq$ and $+$ the point-wise order and sum of the vector space $X \rightarrow \mathbb{R}$, respectively, and with $\lambda(_)$ the operation of multiplication by scalars. The function transformer $\Diamond_\alpha$ satisfies the following properties: • (monotone) if $f \sqsubseteq g$ then $\Diamond_\alpha f \sqsubseteq \Diamond_\alpha g$, • ($\Diamond_\alpha(\underline{1})$ is Boolean) $\Diamond_\alpha(x \mapsto 1) \in X \rightarrow \{0, 1\}$, • (sublinear) $\Diamond_\alpha(f + g) \sqsubseteq \Diamond_\alpha(f) + \Diamond_\alpha(g)$, • (positive affine homogenous) $\Diamond_\alpha(\lambda_1 f + \lambda_2 \underline{1}) = \lambda_1 \Diamond_\alpha(f) + \lambda_2 \Diamond_\alpha(\underline{1})$, for all $\lambda_1 \geq 0$ and $\lambda_2 \in \mathbb{R}$. Furthermore, every function transformer $F : (X \rightarrow \mathbb{R}) \rightarrow (X \rightarrow \mathbb{R})$ with these properties arise as $F = \Diamond_\alpha$ from a unique $\mathcal{P}_{cc}\mathcal{D}$ -coalgebra (X, α) .*

The theorem can be generalized to the setting described in Remark 5. In a follow-up of this paper it will be shown how, on the basis of this result, it is possible to develop an algebraic account of $\mathcal{P}_{cc}\mathcal{D}$ -coalgebras in the form of a correspondence between $\mathcal{P}_{cc}\mathcal{D}$ -coalgebras and certain types of algebras (with a rich vector space structure resembling that of the function space $X \rightarrow \mathbb{R}$). We expect this result will be of help in designing *compositional* verification methods for PNTS’s based on *equational reasoning* and *axiomatizations* of $\mathbb{R}$ -valued logics for PNTS’s.

We conclude this section by stating a result of central importance for the practical applicability of UE-bisimilarity in programming languages which can be proved by means of standard process-algebra techniques (see, e.g., [7, §3]).

Theorem 7. *UE-bisimilarity is a congruence relation for all process algebras specified by the probabilistic-nondeterministic PGSOS format of [2].*

Acknowledgements. The author would like to thank Alex Simpson, Alexandra Silva, Jan Rutten, Marcello Bonsangue, Helle Hvid Hansen, Henning Basold and the anonymous reviewers for their helpful comments and suggestions. The support of Advanced Grant ECSYM of the ERC is acknowledged with gratitude.

References

1. Baier, C., Katoen, J.P.: Principles of Model Checking. The MIT Press (2008)
2. Bartels, F.: On Generalised Coinduction and Probabilistic Specification Formats: Distributive Laws in Coalgebraic Modelling. PhD thesis, CWI (2004)
3. Bernardo, M., De Nicola, R., Loret, M.: Revisiting trace and testing equivalences for nondeterministic and probabilistic processes. In: Birkedal, L. (ed.) FOSSACS 2012. LNCS, vol. 7213, pp. 195–209. Springer, Heidelberg (2012)
4. Bianco, A., de Alfaro, L.: Model checking of probabilistic and nondeterministic systems. In: Thiagarajan, P.S. (ed.) FSTTCS 1995. LNCS, vol. 1026, pp. 499–513. Springer, Heidelberg (1995)
5. Chatzikokolakis, K., Palamidessi, C., Panangaden, P.: On the Bayes risk in information-hiding protocols. *Journal of Computer Security* 16(5) (2008)
6. Crafa, S., Ranzato, F.: A spectrum of behavioral relations over LTSs on probability distributions. In: Katoen, J.-P., König, B. (eds.) CONCUR 2011. LNCS, vol. 6901, pp. 124–139. Springer, Heidelberg (2011)
7. D’Argenio, P.R., Gebler, D., Lee, M.D.: Axiomatizing bisimulation equivalences and metrics from probabilistic SOS rules. In: Muscholl, A. (ed.) FOSSACS 2014. LNCS, vol. 8412, pp. 284–298. Springer, Heidelberg (2014)
8. de Alfaro, L., Majumdar, R.: Quantitative solution of omega-regular games. *Journal of Computer and System Sciences* 68, 374–397 (2004)
9. de Alfaro, L., Majumdar, R., Raman, V., Stoelinga, M.: Game refinement relations and metrics. *Logical Methods in Computer Science* 4 (2008)
10. Deng, Y., van Glabbeek, R.: Characterising probabilistic processes logically. In: Fermüller, C.G., Voronkov, A. (eds.) LPAR-17. LNCS, vol. 6397, pp. 278–293. Springer, Heidelberg (2010)
11. Deng, Y., van Glabbeek, R., Hennessy, M., Morgan, C.: Testing finitary probabilistic processes. In: Bravetti, M., Zavattaro, G. (eds.) CONCUR 2009. LNCS, vol. 5710, pp. 274–288. Springer, Heidelberg (2009)
12. Desharnais, J., Gupta, V., Jagadeesan, R., Panangaden, P.: Weak bisimulation is sound and complete for PCTL. In: Brim, L., Jančar, P., Křetínský, M., Kučera, A. (eds.) CONCUR 2002. LNCS, vol. 2421, pp. 355–370. Springer, Heidelberg (2002)
13. Goubault-Larrecq, J.: Prevision domains and convex powercones. In: Amadio, R.M. (ed.) FOSSACS 2008. LNCS, vol. 4962, pp. 318–333. Springer, Heidelberg (2008)
14. Huber, P.: Robust Statistics. Wiley (1991)
15. Huth, M., Kwiatkowska, M.: Quantitative analysis and model checking. In: Proc. of LICS (1997)
16. Kechris, A.S.: Classical Descriptive Set Theory. Springer (1994)
17. Kozen, D.: Results on the propositional mu-calculus. *Theoretical Computer Science*, 333–354 (1983)
18. Kurz, A.: Logics for Coalgebras and Applications to Computer Science. PhD thesis, Ludwig Maximilian University of Munich (2000)
19. Lax, P.: Functional Analysis. Wiley Interscience (2002)
20. Lynch, N., Segala, R.: Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing* 2, 250–273 (1995)
21. Milner, R.: A Calculus of Communication Systems. LNCS, vol. 92. Springer, Heidelberg (1980)
22. Mio, M.: Game Semantics for Probabilistic μ -Calculi. PhD thesis, School of Informatics, University of Edinburgh (2012)

23. Mio, M.: Probabilistic Modal μ -Calculus with Independent product. *Logical Methods in Computer Science* 8(4) (2012)
24. Mio, M., Simpson, A.: Łukasiewicz mu-calculus. In: *Proc. of Workshop on Fixed Points in Computer Science. EPTCS*, vol. 126 (2013)
25. Mio, M., Simpson, A.: A proof system for compositional verification of probabilistic concurrent processes. In: Pfenning, F. (ed.) *FOSSACS 2013. LNCS*, vol. 7794, pp. 161–176. Springer, Heidelberg (2013)
26. Morgan, C., McIver, A.: A probabilistic temporal calculus based on expectations. In: *Proc. of Formal Methods Pacific* (1997)
27. Panangaden, P.: *Labelled Markov processes*. Imperial College Press (2009)
28. Sangiorgi, D., Rutten, J.: *Advanced topics in bisimulation and coinduction*. Cambridge University Press (2011)
29. Segala, R.: *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis. MIT (1995)
30. Sokolova, A.: Probabilistic Systems Coalgebraically: A survey. *Theoretical Computer Science* 412(38) (2011)
31. Song, L., Zhang, L., Godskesen, J.C.: Bisimulations Meet PCTL Equivalences for Probabilistic Automata. In: Katoen, J.-P., König, B. (eds.) *CONCUR 2011. LNCS*, vol. 6901, pp. 108–123. Springer, Heidelberg (2011)
32. Tix, R., Keimel, K., Plotkin, G.D.: Semantic domains for combining probability and non-determinism. *Electronic Notes in Theoretical Computer Science* (2005)