Upper-Expectation Bisimilarity
and Real-valued Modal Logics

Matteo Mio * **

CWI-Amsterdam,
`miomatteo@gmail.com`

Abstract. Several notions of bisimulation relations for probabilistic non-deterministic transition systems have been considered in the literature. We consider a novel testing-based behavioral equivalence called *upper-expectation bisimilarity* and develop its theory using standard results from linear algebra and functional analysis. We show that, for a wide class of systems, our new notion coincides with Segala's *convex bisimilarity*. We develop logical characterizations in terms of expressive probabilistic modal μ-calculi and a novel real-valued modal logic. We prove that upper-expectation bisimilarity is a congruence for the wide family of process algebras specified following the probabilistic GSOS rule format.

Keywords: Quantitative Modal Logics, Bisimulation, Convexity.

1 Introduction

Directed-graph structures are sufficient for modeling nondeterministic programs and concurrent systems but can not be used to represent other important aspects of computation, such as *probabilistic behavior*, *timed transitions* and other *quantitative* information one might need to express. To address this limitation, since the late 80's, a lot of research has focused on the identification of appropriate structures for expressing these quantitative aspects (see, e.g., [36,24,4]), and in particular for modeling probabilistic behavior. One of the most successful such models is today known under several names: (simple) Segala systems [3], concurrent Markov chains [24], probabilistic automata [51] or just *probabilistic nondeterministic transition systems* (PNTS). Today PNTS's are the mathematical structures, generalizing standard nondeterministic transition systems (NTS), most often used to provide operational semantics to probabilistic and nondeterministic languages [26,35,3].

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* [42]. 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 the modal μ-calculus [31] or in other (weaker but useful in practice) branching-time logics such as CTL, CTL* [55], and the basic modal logic **K** [5] or its labeled version, the Hennessy–Milner modal logic [25]. Furthermore, **B2)** bisimilarity is a congruence for a wide family of process algebras, all specified following one of the many rule formats: GSOS [6], tyft/tyxt [22], *etc.* This means that if programs p and q are bisimilar then, for any program context $C[x]$, the two composite programs $C[p]$ and $C[q]$ are also bisimilar. This is the crucial property that authorizes the substitution of equals for equals in programs, a process of fundamental importance in, e.g., optimization and compositional program development and verification. **B3)** Bisimilarity also enjoys a rich mathematical theory based on coinduction and plays a role of paramount importance in *coalgebra* [50]. Lastly, but importantly, **B4)** bisimilarity can also be explained in terms of Milner's standard metaphor of *push-buttons experiments* on systems [42]. 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 [51,39,54,40,15,16]. The one which has attracted most attention so far was introduced by Segala in [51] 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 theory (see., e.g., [53]). Importantly, standard bisimilarity is a congruence relation (cf. B2) for the wide class of PGSOS process algebras, which virtually includes all Milner's CCS-style (probabilistic) process operators of practical interest [3]. However standard bisimilarity happens to be strictly finer than the equivalence induced by important temporal logics for expressing useful properties of PNTS, such as, PCTL, PCTL* [4,2] and the fixed-point modal logics (μ-calculi) of [28,15,41,43]. This is a rather unsatisfactory fact (cf. B1) because one does not want to distinguish between programs that satisfy the same properties of interest. Furthermore no (widely accepted) testing semantics (cf. B4) for standard bisimilarity exists.

In [51] another behavioral equivalence for PNTS's, which we refer to as *convex bisimilarity*, is introduced by Segala. This notion, as we shall discuss in Section 3.1, is strongly motivated by the concept of *probabilistic scheduler*. Convex bisimilarity is coarser than standard bisimilarity. However it is known to be strictly finer than the equivalence induced by the logic PCTL* which in turn admits a coinductive characterization [54]. Other notions of bisimulation for PNTS's have been recently introduced in [17], [11] and [40].

Contributions. In this paper we analyze yet another notion of behavioral equivalence which we refer to as *upper expectation (*UE*) bisimilarity*. By application

of known results in linear algebra and functional analysis, we show that UE-bisimilarity and convex bisimilarity coincide under very mild assumptions. As it is often the case, having two complementary descriptions of the same concept turns out to be useful. As we discuss in Section 4, UE-bisimilarity arises naturally from a very abstract testing scenario based on $\mathbb{R}$-valued experiments (cf. property **B4** above). Unlike other similar works, our experiments are not given by, e.g., the formulas of a given logic (such as the μ-calculus of [15]) nor by terms of some process algebra (see, e.g., [17]) but, instead, are modeled by functions $f : S \to \mathbb{R}$ from program states to real values.

We argue in Section 5 that this abstraction, besides being mathematically convenient, sheds light on the mathematical foundations of real-valued logics (including the μ-calculi mentioned above) which have been subject of increasing interest in the last decade [41,49,19,14,44]. In Section 5.2 we develop the theory of a novel real-valued modal logic, which we call **R** (in honor of Frigyes Riesz), inspired by the powerful results from functional analysis presented in Section 5.1. We define a model-theoretic semantics for **R** based on PNTS's and an algebraic semantics. We prove that the two coincide. To the best of our knowledge, no algebraic (i.e., axiomatic) semantics of other $\mathbb{R}$-valued modal logics for PNTS's have appeared in the literature. In standard modal logic, particular classes of models (e.g., reflexive NTS's) can be obtained by adding axioms (e.g., $\Box \phi \to \phi$) to the base modal logic **K** [5]. We show that Markov processes and ordinary NTS's, which can be seen as particular classes of PNTS's, can be captured by adding appropriate axioms to the basic logic **R**. This confirms the naturalness of **R** which we see as one of the main contributions of this work.

The logic **R** is valuable for its mathematical simplicity but cannot express useful properties of systems such as: termination goals, liveness constraints, *etc.* The $\mathbb{R}$-valued modal μ-calculi of [43] are better suited to this task, as they (strictly) subsume the probabilistic logic PCTL of [4] which is the specification language adopted in, e.g., the PRISM verification framework [34]. In Section 5.3 we prove that UE-bisimilarity coincides with the equivalence relation induced by the μ-calculi of [43]. Thus the quantitative approach to probabilistic μ-calculi may be considered equally suitable as a mechanism for characterising process equivalence as other non-quantitative μ-calculi advocated for this purpose (see, e.g., [16]). We also provide in Section 5.4 a logical characterization of the Hausdorff behavioral metric for PNTS's, a concept of fundamental importance in the theory of approximation of probabilistic systems pioneered by Panangaden [49]. Collectively, the results of Section 5 provide strong logical foundations for UE-bisimilarity (cf. property **B1**) comparable with those of Milner and Park's bisimilarity.

We prove in Section 6 that UE-bisimilarity is a congruence relation with respect to the CCS-style (communicating) parallel composition operator. This result is known from [39] where the property is proved for (the equivalent under mild conditions) convex bisimilarity and parallel composition is defined in terms of an equivalent automata-theoretic synchronized product operation. Our proof, however, is interesting because it uses process-algebra methods and read-

ily extends to every process-operator definable in the PGSOS format of [3] which virtually includes all CCS-style process operators of practical interest (cf. property **B2**).

In Sections 2 and 3, summarizing standard and well known results, we discuss how UE-bisimilarity can be understood coalgebraically in terms of cocongruence for an appropriate functor. This confirms the naturalness and mathematical robustness (cf. property **B3**) of this notion of behavioral equivalence.

It seems likely that our results will be of help towards the difficult task of designing and verifying probabilistic concurrent programs and protocols. The importance of bisimulation in software development (e.g., refinement techniques) and verification (e.g., system minimization) is well known. The logical characterizations of bisimulation are of key importance (see, e.g., [52,12]) in the development of *compositional* verification methods, which are crucial in the analysis of industrial-size software. Congruence results, such as the one we obtained for PGSOS process algebras (which can be thought of as abstract prototypes of concurrent programming languages) constitute the basis for the development of design and verification techniques based on equational reasoning. We also suggest that our natural testing semantics for UE-bisimilarity might be related to information-theoretic notions of attacks (cf. Section 5.4) which have received much attention in the recent literature [10].

2   Elementary Notions of Coalgebra

We start by recalling the definition of Milner and Park's bisimulation in the context of ordinary nondeterministic transition systems.

Definition 1 ([42]). *A* nondeterministic transition system *(NTS) is a pair* $(X, \rightarrow)$ *where* X *is a set of states and* $\rightarrow \subseteq X \times X$ *is a transition relation. An equivalence relation* $E \subseteq X \times X$ *is a* bisimulation *if* $(x, y) \in E$ *implies that:*

  – *if* $x \rightarrow x'$ *then there exists* $y \rightarrow y'$ *such that* $(x', y') \in E$*, and*
  – *if* $y \rightarrow y'$ *then there exists* $x \rightarrow x'$ *such that* $(x', y') \in E$*.*

States x *and* y *are* bisimilar *if* $(x, y) \in E$ *for some bisimulation* E*.*

In applications one most often encounters *labeled* NTS's, where an L-indexed set $\{\xrightarrow{a}\}_{a \in L}$ of relations is considered for some set of labels L, or *Kripke structures*, where the NTS is endowed with a set of propositional letters 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 (see, e.g., Section 6) and propositional extensions.

It is going to be convenient to employ the basic language of coalgebra [50] in the description of systems and behavioral equivalences. We refer to [50] for a gentle introduction to the subject.

Definition 2. *Let* **Set** *be the category of sets and functions between them. The endofuntor* $\mathcal{P}$ *(powerset) on* **Set** *is defined as:*

- $\mathcal{P}(X) = \{A \mid A \subseteq X\}$,
- $\big(\mathcal{P}(f)\big)(A) = f[A] = \{f(x) \mid x \in A\}$

for all sets X, Y and functions $f : X \to Y$.

We can now restate the definition of NTS's using coalgebraic terminology. The equivalence between the two definitions is well known (see, e.g., [50]) and straightforward to verify.

Definition 3. *A* nondeterministic transition system *(NTS) is a $\mathcal{P}$-coalgebra $(X, \alpha : X \to \mathcal{P}(X))$.*

Definition 4. *Let F, G be two endofunctors on* **Set***. A* natural transformation *from F to G, written $\eta : F \to G$, is a collection of functions $\eta_X : F(X) \to G(X)$, indexed by sets $X \in$ **Set***, such that for all functions $f : Y \to Z$ it holds that $\eta_Z \circ G(f) = F(f) \circ \eta_Y$.*

Thus, a natural transformation is a method η for transforming the action of a functor F into that of the functor G in a uniform way. Following standard ideas (see, e.g., [53]), we shall use this concept (cf. Propositions 1 and 3) to formalize the idea of transformation of systems of some type into systems of another type. Coalgebra provides an abstract notion of behavioral equivalence based on the categorical concept of cocongruence [33,13].

Definition 5. *Let $X \in$ **Set** and $E \subseteq X \times X$ be an equivalence relation. The map $q_E : X \to X/E$, mapping elements $x \in X$ to their E-equivalence class $[x]_E$, is called the* quotient map *of E.*

Definition 6. *Given a* **Set***-endofunctor F and a F-coalgebra (X, α), an equivalence relation $E \subseteq X \times X$ is a called a* cocongruence *if for all $(x, y) \in E$ it holds that $\big(\alpha(x), \alpha(y)\big) \in \hat{E}$, where the* lifted relation *$\hat{E} \subseteq F(X) \times F(X)$ is defined as $\ker(F(q_E)) = \{(A, B) \mid F(q_E)(A) = F(q_E)(B)\}$.*

When a functor F is fixed, the definition can be made explicit by expanding the definition of $\hat{E}$. Consider, e.g., the functor $\mathcal{P}$, an equivalence relation $E \subseteq X \times X$ and two sets $A, B \in \mathcal{P}(X)$ with $A = \{x_i\}_{i \in I}$ and $B = \{y_j\}_{j \in J}$. Then $(A, B) \in \hat{E}$ holds iff the two sets $A/E = \{[x_i]_E\}_{i \in I}$ and $B/E = \{[y_j]\}_{j \in J}$ are *equal* as sets. Thus the relation E is a cocongruence if, for every $(x, y) \in E$, the two sets $\alpha(x)$ and $\alpha(y)$ of reachable states are *equal modulo E*. Hence, cocogruences are those equivalence relations that preserve the transition (i.e., coalgebraic) structure of the system. Note that $(A, B) \in \hat{E}$ holds iff for all $x \in A$ there exists $y \in B$ such that $(x, y) \in E$ and *viceversa*. Therefore the abstract notion of cocongruence for the functor $\mathcal{P}$ coincides with ordinary Milner and Park's bisimulation for NTS's.

3   PNTS's and Bisimulations

Definition 7 ([53]). *The endofuntor $\mathcal{D}$ (discrete probability distributions) on* **Set** *is defined as:*

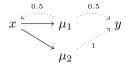- $\mathcal{D}(X) = \{\mu \colon X \to [0,1] \mid \sum_x \mu(x) = 1\}$
- $\big(\mathcal{D}(f)\big)(\mu) = f[\mu], \quad y \xmapsto{f[\mu]} \sum\{\mu(x) \mid x \in f^{-1}(y)\}$

for all sets X, Y and functions $f \colon X \to Y$. For $x \in X$ we denote with $\delta(x)$ (or δ_x) the probability Dirac distribution specified by $\delta_x(y) = 1$ if $y = x$ and 0 otherwise. For a set $A \subseteq X$, we write $\mu(A)$ for $\sum_{x \in A} \mu(x)$.

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

Definition 8. *A* probabilistic nondeterministic transition system *(PNTS) is a $\mathcal{P}\mathcal{D}$-coalgebra $(X, \alpha \colon X \to \mathcal{P}\mathcal{D}(X))$. We write $x \to \mu$ 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* probability distributions μ, i.e., such that $\mu \in \alpha(x)$, 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 $X = \{x, y\}$ and transition map $\alpha(x) = \{\mu_1, \mu_2\}$ and $\alpha(y) = \emptyset$, with $\mu_1(x) = \mu_1(y) = \frac{1}{2}$ and $\mu_2(y) = 1$, can be depicted as:

The combination of nondeterministic choices immediately followed by probabilistic ones, allows the modeling of concurrent probabilistic programming languages in a natural way [3,26]. As remarked earlier, one often considers labeled or propositional generalizations of PNTS's. All our results extend to these (and other similar) generalizations in a straightforward way (cf. Section 6).

PNTS's with a definition equivalent to the coalgebraic one given above were introduced by Segala [51] who also defined two notions of bisimilarity for PNTS's. One, the stronger (i.e., finer) of the two, which we refer to as *standard bisimilarity* here, is a natural extension of Milner and Park's bisimilarity for NTS's based on the insights of Larsen and Skou [36].

Definition 9 (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 $x \to \mu$ then there exists ν such that $y \to \nu$ and $\mu =_E \nu$, and*
- *if $y \to \nu$ then there exists μ such that $x \to \mu$ and $\mu =_E \nu$,*

where $\mu =_E \nu$ holds if $\mu(A) = \nu(A)$ (see Definition 7) for all sets $A \subseteq X$ which are unions of E-equivalence classes.

Remark 1. The notion of standard bisimilarity coincides with that of cocongruence for the functor $\mathcal{P}\mathcal{D}$. This is simple to verify by expanding the definition of $\hat{E}$ as in Definition 6. For two sets $A = \{\mu_i\}_{i \in I}$ and $B = \{\nu_j\}_{j \in J}$ we have

$(A, B) \in \hat{E}$ if and only if $A/E = \{[\mu_i]_E\}_{i \in I}$ and $B/E = \{[\nu_j]_E\}_{j \in J}$ are equal as sets where, for each $\mu \in \mathcal{D}(X)$, we denoted with $[\mu]_E$ the probability distribution $q_E[\mu]$ (see definitions 5 and 7).

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 bisimulation, i.e., are equal once their elements $\mu, \nu \in \mathcal{D}(X)$ are considered modulo E. At the level of probability distributions, $\mu =_E \nu$ means that if one glues together (i.e., identifies) E-related states (i.e., applies the quotient map $x \mapsto [x]_E$), then μ and ν become *equal*, i.e., they assign the same probabilities to all events $A \subseteq X/E$. 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 different[1]. The processes x and y are not standard bisimilar. This is because y can lead to a probability distribution μ_3 which can not be matched by x.
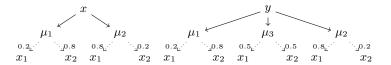

Fig. 1: Example of states (x, y) not standard bisimilar.

3.1 Convex Bisimilarity

It has been argued by Segala (see, e.g., [51]) 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. For instance the scheduling algorithm of an operating system might use probabilistic protocols to perform fair choices in an efficient way. 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 10. *Let X be a set. A* convex combination *of elements $\mu_i \in \mathcal{D}(X)$, for $i \in \{0, \dots, n\}$, is a probability distribution ν of the form $\nu(x) = \sum_i \lambda_i \mu_i(x)$,*

---

[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.

for $\lambda_i \in [0,1]$ and $\sum_i \lambda_i = 1$. The convex hull *of $A \in \mathcal{PD}(X)$, denoted by $\mathrm{H}(A)$, is the collection of all possible convex combinations of elements in A. The set A is* convex *if $A = \mathrm{H}(A)$. The entire set $\mathcal{D}(X)$, the emtpyset and the singletons $\{\mu\}$, for $\mu \in \mathcal{D}(X)$, are examples of convex sets.*

We are now ready to introduce the second notion of bisimilarity introduced by Segala [51,39], which we refer to[2] as *convex bisimilarity*, based on the concept of probabilistic schedulers.

Definition 11 (Convex Bisimulation). *Given a PNTS (X, α), a* convex bisim- ulation *is an equivalence relation $E \subseteq X \times X$ such that if $(x, y) \in E$ then*

- *if $x \to_C \mu$ then there exists ν such that $y \to_C \nu$ and $\mu =_E \nu$,*
- *if $y \to_C \nu$ then there exists μ such that $x \to_C \mu$ and $\mu =_E \nu$,*

where $x \to_C \mu$ if and only if $\mu \in \mathrm{H}(\alpha(x))$.

Thus convex bisimilarity is obtained by replacing $\to$ with $\to_C$ in Definition 9. Probabilistic schedulers take the place of ordinary schedulers and, by means of probabilistic choices, can simulate any convex combination of the reachable probability distributions.

Remark 2. Polynomial time algorithms for computing convex bisimilarity, based on linear programming, are studied in [9].

If probabilistic schedulers are assumed, Example 1 shows how, as observed by Segala in the first place, standard-bisimilarity is too strict because it distin- guishes between states that, under the intended interpretation, ought to be identified. This fact does not mean that cocogruence (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. We now briefly discuss how convex bisimilarity can be understood coalgebraically.

Definition 12. *The endofuntor $\mathcal{P}_c\mathcal{D}$ on* **Set** *is defined as:*

- $\mathcal{P}_c\mathcal{D}(X) \overset{\mathrm{def}}{=} \{A \mid A \in \mathcal{PD}(X) \text{ and } A = \mathrm{H}(A)\}$
- $(\mathcal{D}(f))(A) \overset{\mathrm{def}}{=} f[A] = \{f[\mu] \mid \mu \in A\}$

where $f[\mu]$ is defined as in Definition 7.

Hence $\mathcal{P}_c\mathcal{D}(X)$ is the collection of all convex subsets of $\mathcal{D}(X)$. This means that $\mathcal{P}_c\mathcal{D}(X) \subseteq \mathcal{PD}(X)$. Furthermore the action of $\mathcal{P}_c\mathcal{D}$ on morphisms f is the same as that of $\mathcal{PD}$ (i.e., $\mathcal{P}_c\mathcal{D}(f)$ is $\mathcal{PD}(f)$ restricted to convex sets) and indeed $f[A]$ is a convex set whenever A is convex, as it is simple to verify.

---

[2] The adjective *probabilistic* is often adopted in the literature [51,54]. We prefer the adjective *convex* which transparently reflects the mathematics behind the notion.

Remark 3. The fact that $\mathcal{P}_c\mathcal{D}$ is a functor is well known. In fact more is true, and $\mathcal{P}_c\mathcal{D}$ carries a monad structure [56]. This property has been extensively studied, especially in the field of domain theory [56,21,46], 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 $\mathcal{P}\mathcal{D}$-coalgebra such that, for all $x \in X$, the set $\alpha(x)$ is convex. Furthermore the convex hull operation gives us a natural way to convert a $\mathcal{P}\mathcal{D}$-coalgebra into a $\mathcal{P}_c\mathcal{D}$-coalgebra.

Proposition 1. *The identity map* $\mathrm{id}_X : \mathcal{P}_c\mathcal{D}(X) \to \mathcal{P}\mathcal{D}$, *defined as* $\mathrm{id}_X(A) = A$, *is an injective natural transformation from* $\mathcal{P}_c\mathcal{D}$ *to* $\mathcal{P}\mathcal{D}$. *The map* $\mathrm{H}_X : \mathcal{P}\mathcal{D}(X) \to \mathcal{P}_c\mathcal{D}$, *defined as* $\mathrm{H}_X(A) = \mathrm{H}(A)$, *is a surjective natural transformation from* $\mathcal{P}\mathcal{D}$ *to* $\mathcal{P}_c\mathcal{D}$.

Transforming a $\mathcal{P}\mathcal{D}$-coalgebra (i.e., a PNTS) (X, α) into the $\mathcal{P}_c\mathcal{D}$-coalgebra $(X, \mathrm{H}_X \circ \alpha)$ precisely corresponds to the substitution of the arrow relation $\to$ in Definition 9 with the relation $\to_C$ in Definition 11. 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 $(X, \mathrm{H}_X \circ \alpha)$. The process of transforming a $\mathcal{P}\mathcal{D}$-coalgebra (X, α) into a $\mathcal{P}_c\mathcal{D}$-coalgebra generally causes a loss of information. For instance, the two systems of Figure 1 are both mapped to the same $\mathcal{P}_c\mathcal{D}$-coalgebra. This transformation process is really a quotient operation (H_X is surjective) induced by the *behavioral equation* $A \approx B$ whenever $\mathrm{H}(\mathrm{A}) = \mathrm{H}(\mathrm{B})$, for all $A, B \in \mathcal{P}\mathcal{D}(X)$, capturing the the behavior of probabilistic schedulers.

Remark 4. 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., $\mathcal{P}\mathcal{D}$-coalgebras) (X, α), perhaps represented as finite graphs, and tacitly replace α with $\mathrm{H}_X \circ \alpha$ (i.e., $\to$ with $\to_C$). This is convenient since, generally, the convex hull of a finite set is uncountable.

The discussio carried out in this subsection 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.

4 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 on $\mathcal{P}_c\mathcal{D}$-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
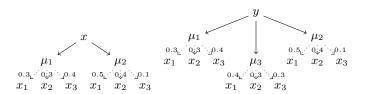

Fig. 2: Example of states (x, y) not convex bisimilar.

not simple, however, to find a *concrete*[3] reason to distinguish between the two states. As a matter of fact, it has been proved in [54] that the states x and y satisfy the same properties formulated in the expressive logic PCTL* of [4].

Remark 5. While modal logics, carefully crafted to capture convex bisimilarity (and even standard bisimilarity) can be defined [39,51,16,17], it is certainly interesting to look at the distinguishing power of popular temporal logics for PNTS's, (of which PCTL* is a main example) capable of expressing branching properties of probabilistic concurrent systems useful in practice (see, e.g., [34]).

To explain why the two states x and y have indeed similar behavior, we now introduce a simple experimental scenario. Suppose we are allowed to make repeated experiments (in the sense of Milner's push-buttons metaphor [42]) on the PNTS's of Figure 1. After $n \to \infty$ experiments (that is, for n as big as desired) we observe that an event A (e.g., $A = \{x_1\}$, representing the occurrence of terminal state x_1) happened m times. We can then make the following reasonable[4] assessment: the PNTS *may* exhibit behavior A (i.e., end up in x_1) with probability $\frac{m}{n}$. It seems then natural to stipulate that two states x and y are equivalent if, for each event A, the state x can exhibit behavior (event) A with probability (frequency) λ if and only if y can. This leads to the following definition.

Definition 13. *Let X be a set and $S \in \mathcal{PD}(X)$ a set of probability distributions on X. The* upper probability functional $up_S : \mathcal{P}(X) \to [0, 1]$, *mapping subsets $A \subseteq X$ to real numbers in $[0, 1]$, is defined as the supremum value $up_S(A) = \bigsqcup_{\mu \in S} \mu(A)$.*

---

[3] This is deliberately a vague adjective which could be seconded by, *experimental, operational, testing-based,* ...

[4] This is of course based on the common frequentist interpretation of probabilities as limits of relative frequencies in a large number of trials.

The adjective *functional* is motivated by the fact that $\mathcal{P}(X)$ can be seen as the space of characteristic functions $X \to \{0,1\}$. Thus up_S maps functions (predicates) to real numbers.

Definition 14. *Given a PNTS (X, α), an* upper probability (UP) *bisimulation is an equivalence relation $E \subseteq X \times X$ such that if $(x, y) \in E$ then the equality $up_{\alpha(x)}(A) = up_{\alpha(y)}(A)$ holds for all sets $A \subseteq X$ which are unions of E-equivalence classes.*

Remark 6. The choice of considering upper probabilities seems one-sided, but one could equally well choose lower probabilities (lp_S) observing that $lp_S(A) = 1 - up_S(X \setminus A)$.

The idea behind the definition is simple. If x and y are UP-bisimilar, than if x can exhibit event A with probability λ then also y can, and *viceversa*. At the same time, if x and y are not UP-bisimilar, there exists an event A such that an agent, by performing on x and y a sufficiently large number of repeated experiments looking after the occurrences of event A, *may* observe a probability of A on x that *can not* be achieved in y. The restriction on A being union of E-equivalence classes is natural, since we wish to identify (i.e., apply $x \mapsto [x]_E$) E-related (i.e., UP-bisimilar) states.

We introduced UP-bisimilarity to understand in what sense the behavior of the states x and y of Figure 2 are similar. It is now simple to verify that x and y are UP-bisimilar. Thus x and y can be not distinguished by means of experiments. As it turns out, UP-bisimilarity has been recently derived (with different names) in [40] and [54], albeit following a different conceptual path. It is shown in [54] that UP-bisimilarity is strictly coarser than the equivalences induced by the logic PCTL* and its weaker fragment PCTL. This is a significant drawback of UP-bisimilarity as one definitely wants to distinguish between states that can be separated by useful temporal logics such as PCTL and PCTL*. However we derived UP-bisimilarity following a straightforward testing metaphor, which we now refine to get a better behaved alternative.

Suppose indeed that in the experiments carried on PNTS's, one is not just allowed to observe the occurrences of events and thus, after enough ($n \to \infty$) experiments, estimate their probabilities by means of relative frequencies. Rather, one can 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 new scenario is better explained by a simple example. Consider again the states x and y of Figure 2 and consider the function $g : \{x_1, x_2, x_3\} \to \mathbb{R}$ defined as $g(x_1) = 60$, $g(x_2) = 0$ and $g(x_3) = 50$. 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 $\{60, 0, 50, 0, 60\}$ and its average $34 = \frac{170}{5}$ would be our observation.

Definition 15. *Let X be a set, $\mu \in \mathcal{D}(X)$ and $f : X \to \mathbb{R}$. The* expected value *of f under μ, written $E_\mu(f)$, is defined as $E_\mu(f) = \sum_x \mu(x) f(x)$.*

Note that the expected values of g under μ_1, μ_2 and μ_3 are 38, 35 and 39, respectively. This readily means that, for $n \to \infty$, the average resulting from experiments g on state x is *necessarily* smaller or equal than 38, while in state y it *can be* strictly greater than 38 (and at most 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 16. *Let X be a set and $S \in \mathcal{PD}(X)$ a set of probability distributions on X. The* upper expectation functional $ue_S : (X \to \mathbb{R}) \to \mathbb{R}$, *mapping functions $X \to \mathbb{R}$ to real numbers, is defined as $ue_S(f) = \bigsqcup_{\mu \in S} E_\mu(f)$.*

Definition 17. *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 $f : X \to \mathbb{R}$, i.e., such that if $(z, w) \in E$ then $f(z) = f(w)$.*

Remark 7. Up to minor modifications, the notion of UE bisimilarity already appeared in the literature [15] under the name of *game bisimilarity*, in the (significantly) different setting of two-player stochastic (concurrent) games. See Remark 9 below for further discussion.

We restrict to E-invariant experiments f following the usual idea that E-related (i.e., UE-bisimilar) states ought to be identified. Note how UP-bisimilarity is obtained by restricting the set of $\mathbb{R}$-valued functions f to $\{0, 1\}$-valued functions $g : X \to \{0, 1\}$, in the definition of UE-bisimulation. Indeed note that g is E-invariant if and only if g is the characteristic function of some set $Y \subseteq X$ which is union of E-equivalence classes. This observation, together with the fact that the two states x and y of Figure 2 are UP-bisimilar but not UE-bisimilar, reveals that $\mathbb{R}$-valued experiments (modeled by $f : X \to \mathbb{R}$) generally provide more information than just observations of events (modeled by $g : X \to \{0, 1\}$).

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. The dual *modalities* may/must (can/cannot) considered in our testing scenario are, of course, well known key concepts in classical bisimulation theory, modal logics and concurrency theory [42,5].

4.1 Relation between Convex and UE-bisimilarity

UE-bisimilarity arises naturally from the simple testing scenario discussed above and, as we shall discuss in Section 5, 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 4) can be thought as $\mathcal{PD}$-coalgebras modulo the behavioral equation $A \approx B$ if $H(A) = H(B)$, for $A, B \in$

$\mathcal{PD}(X)$. 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 the technical development.

Convention 41 *We restrict attention to PNTS's having a finite state space, i.e., $\mathcal{PD}$-coalgebras $(X, \alpha \colon X \to \mathcal{PD}(X))$ having a finite carrier set $X = \{x_1, \dots, x_n\}$. This allows one to view the space of functions $X \to \mathbb{R}$ 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*) subset of $\mathbb{R}^n$ and that $\alpha(x)$, for $x \in X$, can be infinite.*

In Section 7 we briefly discuss how our results can be generalized to infinite systems at the cost of mathematical complications since infinite dimensional topological spaces need to be considered.

   The following is a known result in optimization theory and applied statistics (see, e.g., [27, §10.2], [57] and [23]).

Theorem 1 ([27]). *Let X be a finite set and $A, B \in \mathcal{PD}(X)$. Then $ue_A = ue_B$ if and only if $\mathrm{cl}(\mathrm{H}(A)) = \mathrm{cl}(\mathrm{H}(B))$, where $\mathrm{cl}(C)$ 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., [37, §8.4]). For this reason the set $\mathrm{cl}(\mathrm{H}(A))$ is called the *closed convex hull* of A. In what follows we denote with $\overline{\mathrm{H}}$ the operation $A \mapsto \mathrm{cl}(\mathrm{H}(A))$.

   We can use the result of Theorem 1 to prove the following alternative characterization of UE-bisimulation (cf. Definition 11).

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

   – *if $x \to_{CC} \mu$ then there exists ν such that $y \to_{CC} \nu$ and $\mu =_E \nu$,*
   – *if $y \to_{CC} \nu$ then there exists μ such that $x \to_{CC} \mu$ and $\mu =_E \nu$,*

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

Proof. For $A \in \mathcal{PD}(X)$, denote with $A/E \in \mathcal{PD}(X/E)$ the set $A/E = \{q[\mu] \mid \mu \in A\}$ (cf. discussion after Definition 9), where q is the quotient map of E. We can now rephrase the assertion of the theorem as follows. The relation E is a UE-bisimulation iff for every $(x, y) \in E$ it holds that $\overline{\mathrm{H}}(\alpha(x))/E = \overline{\mathrm{H}}(\alpha(y))/E$.

   Observe that the map $\mu \mapsto q[\mu]$ is linear, hence continuous since $\mathbb{R}^n$ is finite dimensional. This implies that, for every convex closed set $A \in \mathcal{PD}(X)$, the set $\overline{\mathrm{H}}(A)/E$ is a convex closed set in $\mathcal{PD}(X/E)$. Note that there exists a one-to-one correspondence between E-invariant functions $f \colon X \to \mathbb{R}$ and arbitrary functions $g \colon X/E \to \mathbb{R}$. The definition of UE-bisimilarity as in Definition 17 can then be rephrased as: $(x, y) \in E$ implies $ue_{\alpha(x)/E} = ue_{\alpha(y)/E}$. The result then follows by Theorem 1.

Thus UE-bisimilarity can be obtained by replacing $\to_C$ with $\to_{CC}$ in Definition 11 of convex bisimilarity. As a consequence UE-bisimilarity can be strictly coarser than convex bisimilarity. The following lemma, however, reveals that the two notions coincide for a wide class of PNTS's.

Proposition 2. *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.*

Proof. If $\alpha(x) \subseteq \mathcal{D}(X)$ is closed then it is compact since $\mathcal{D}(X)$ is bounded. The convex hull of a compact set in $\mathbb{R}^n$ is itself closed and compact. Thus $\mathrm{H}(\alpha(x)) = \overline{\mathrm{H}}(\alpha(x)))$ and the result follows.

Restricting to PNTS's of this kind can hardly be seen as a limitation in concrete applications. First, every finite set is closed, thus every PNTS representable as a finite graph satisfies Convention 41 and the closedness condition of Proposition 2. Furthermore every NTS and every Markov process (see [53] and definition below) can be seen as PNTS's satisfying the closedness condition.

Definition 18. *A (discrete)* Markov process *(or Markov* chain*) is a* MP*-coalgebra of the functor* $\mathrm{MP}(X) = \{*\} + \mathcal{D}(X)$.

Proposition 3. *The set-indexed collection of maps $\eta_X : \mathrm{MP}(X) \to \mathcal{PD}(X)$ and $\gamma_X : \mathcal{P}(X) \to \mathcal{PD}(X)$ defined as:*

- $\eta_X(*) = \emptyset$ *and* $\eta_X(\mu) = \{\mu\}$
- $\gamma_X(\{x_i\}_{i \in I}) = \overline{\mathrm{H}}(\{\delta(x_i)\}_{i \in I})$, *with $\delta(x_i)$ as in Definition 7*

are injective natural transformations.

Thus we shall consider every NTS and Markov Process as the (unique) corresponding PNTS with the closedness property.

Remark 8. 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 the space of convex compact closed subsets of $\mathcal{D}(X)$ and consider the expected natural transformations id : $\mathcal{P}_{cc}\mathcal{D} \to \mathcal{P}_c\mathcal{D}$ and cl : $\mathcal{P}_c\mathcal{D} \to \mathcal{P}_{cc}\mathcal{D}$. It is then simple to verify that the notion of UE-bisimilarity coincides with that of cocongruence on $\mathcal{P}_{cc}\mathcal{D}$-coalgebras (cf. Remark 4 and associated discussions). Note that the natural transformations of Proposition 3 restrict to type $\mathrm{MP} \to \mathcal{P}_{cc}\mathcal{D}$ and $\mathcal{P} \to \mathcal{P}_{cc}\mathcal{D}$, respectively.

As discussed in Section 11, 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) $f : X \to \mathbb{R}$ 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 9. As already mentioned in Remark 7, in the important work of [15], game bisimilarity is discovered as the kernel of a behavioral (pseudo)metric d

induced by the quantitative $[0,1]$-valued logic $qL\mu^\ominus$ (which, as we will discuss in Section 5.3, is a weak logic not capable of encoding, e.g., the logic PCTL of [4]). The authors of [15] argue in favor of game bisimilarity on the basis of the naturalness of the logic $qL\mu^\ominus$. Our explanation in terms of real-valued experiments is perhaps useful in clarifying and further motivating this notion using logic-free arguments based on the simple (and, as we will discuss in Section 5.3, more general) metaphor of $\mathbb{R}$-valued experiments. No connection between UE-bisimilarity and convex bisimilarity is discussed in [15].

5 Real Valued Modal Logics

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 4 ([27]). *Let X be a finite set and $A \in \mathcal{PD}(X)$. Then $\{\mu \in \mathcal{D}(X) \mid \forall f : X \to \mathbb{R}.\big(E_\mu(f) \leq ue_A(f)\big)\} = \overline{H}(A)$.*

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, we can look at 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. Remark 4), both as a function $(x \mapsto \alpha(x))$ of type $X \to \mathcal{P}_{cc}\mathcal{D}(X)$ and as a function $(x \mapsto ue_{\alpha(x)})$ of type $X \to \big((X \to \mathbb{R}) \to \mathbb{R}\big)$. Equivalently (by currying) the transition map α can seen as the function transformer $\Diamond_\alpha : (X \to \mathbb{R}) \to (X \to \mathbb{R})$ defined as:

$$\big(\Diamond_\alpha f\big)(x) = ue_{\alpha(x)}(f) \overset{\text{Def 16}}{=} \bigsqcup_{x \to \mu} E_\mu(f) \tag{1}$$

It is remarkable here that the function transformer $\Diamond_\alpha$ happens to coincide with the interpretation (i.e., semantics) of the *diamond modality* in all quantitative logics for PNTS's in the literature [28,41,14,44,18]. While it is obvious that the PNTS (X, α) induces $\Diamond_\alpha$ (just as in the definition), the fact that from $\Diamond_\alpha$ one can reconstruct the $\mathcal{P}_{cc}\mathcal{D}$-coalgebra (X, α) is far from clear. Some of the consequences of this observation are explored in Section 5.2.

$\mathbb{R}$-valued modal logics are based on the fundamental idea of replacing ordinary Boolean predicates (i.e., $\{0,1\}$-valued functions) with quantitative ones. Thus, the semantics of a formula ϕ, interpreted on a PNTS (X, α), is a function $[\![\phi]\!] : X \to \mathbb{R}$ and, in particular, $[\![\Diamond\phi]\!] = \Diamond_\alpha([\![\phi]\!])$. Following our discussions, *formulas* can then be thought of as *experiments* and the value $[\![\Diamond\phi]\!](x)$ as the maximal (limit) expected value of experiment ϕ performed after "pushing the button" (in Milner's terminology) at x.

Example 1. 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 $[\![\phi]\!](x_1) = 60$, $[\![\phi]\!](x_2) = 0$ and $[\![\phi]\!](x_3) = 50$.

The many concrete $\mathbb{R}$-valued modal logics appearing in the literature are obtained by considering other connectives to be used in combination with $\lozenge$. For example the constant $\underline{1}$ ($[\![\underline{1}]\!](x) = 1$) and the connective $\sqcup$ ($[\![\phi \sqcup \psi]\!](x) = \max\{[\![\phi]\!](x), [\![\psi]\!](x)\}$) 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.* All $\mathbb{R}$-valued modal logics we are aware of are *sound* with respect to UE-bisimilarity:

Definition 19 (Soundness). *An $\mathbb{R}$-valued modal logic is* sound *if, for every PNTS (X, α), UE-bisimulation E and formula ϕ it holds that if $(x, y) \in E$ then $[\![\phi]\!](x) = [\![\phi]\!](y)$, i.e., $[\![\phi]\!]$ is E-invariant.*

The following is a natural *desideratum* for a modal logic:

Definition 20 (Strong Completeness). *A $\mathbb{R}$-valued modal logic is* strongly complete *if, for every (X, α), the set of functions $[\![\phi]\!] : X \to \mathbb{R}$, with ϕ ranging over formulas, is* dense[5] *in the set $X \to \mathbb{R}$ of functions that respect UE-bisimilarity, i.e., that are E-invariant for every UE-bisimulation E.*

Thus a logic is strongly complete if every experiment $X \to \mathbb{R}$ (which can not separate UE-bisimilar states) can be approximated, with an arbitrary level of precision, by (the denotation of) some formula ϕ. Note that we consider approximations since a logic typically consists of a countable set of formulas, while $X \to \mathbb{R}$ is uncountable. Strong completeness is a natural notion and, as we will discuss in the next sections, a mathematically well behaved one. Note that, another weaker form of completeness could be considered, simply requiring that, whenever $[\![\phi]\!](x) = [\![\phi]\!](y)$ for all ϕ, then x and y are UE-bisimilar. As mentioned earlier (cf. Remark 9) the logic qL$\mu^\ominus$ of [15] is weakly complete.

We will show in Section 5.3 that the probabilistic μ-calculi of [43] are sound and strongly complete. Before proving this, however, we develop the theory of a novel modal logic in Section 5.2 designed on the basis of first principles coming from deep results in linear algebra and functional analysis discussed in Section 5.1.

Remark 10. It is important to appreciate how we derived the core of a $\mathbb{R}$-valued logic for PNTS's (i.e., the interpretation of the basic modality $\lozenge$) from the very elementary and concrete observation (motivating UE-bisimilarity) that $\mathbb{R}$-valued experiments on PNTS's are useful and, due to their greater observational power, should replace ordinary $\{0, 1\}$-observations (cf. example of Figure 2).

5.1 Representation Theorems

The function space $X \to \mathbb{R}$ is an object of supreme importance in linear algebra and functional analysis [37]. It is the real vector space (with sums and scalar

---

[5] We consider the uniform metric on $X \to \mathbb{R}$, i.e., the one induced by the sup norm. Thus $\mathcal{F}$ is dense in $X \to \mathbb{R}$ if for every $f : X \to \mathbb{R}$ and $\epsilon > 0$, there exists $g \in \mathcal{F}$ such that $\max_{x \in X} |f(x) - g(x)| < \epsilon$.

multiplication defined pointwise), often denoted by $C(X)$, of all $\mathbb{R}$-valued con-tinuous[6] functions on X normed[7] by the *sup norm* (also known as L^∞ norm [37]), defined as $\|f\|_\infty = \max_x |f(x)|$. The subspace of $C(X) \to \mathbb{R}$ consisting of *linear functionals* (i.e., such that $F(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 F(f_1) + \lambda_2 F(f_2)$) is known as the *continuous linear dual of $C(X)$*, and denoted by $C(X)^*$. The following is the finite-dimentional variant of the Riesz Representation Theorem, a result of fundamental importance in analysis and measure theory (see, e.g., [37, §A.1]).

Theorem 3 (Riesz Representation Theorem). *Let X be a finite set and $\mu \in \mathcal{D}(X)$. Then E_μ is a monotone (i.e., if $f \sqsubseteq g$ then $E_\mu(f) \leq E_\mu(g)$) linear functional. Dually, every monotone linear functional $F \in C(X)^*$ such that $F(\underline{1}) = 1$ (where $\underline{1}(x) = 1$) is of the form E_μ for some $\mu \in \mathcal{D}(X)$*

Since we are interested in convex closed sets of probability distributions, rather then probabity distributions, we now consider the following variant of Riesz's theorem.

Theorem 4 ([27]). *Let X be a finite set and $A \in \mathcal{PD}(X)$ be a convex closed and nonempty set. Then ue_A is a monotone functional $F : C(X) \to \mathbb{R}$ which is subadditive (i.e., $F(f + g) \leq F(f) + F(g)$), positively affinely homogeneous (i.e., for all $\lambda_1 \geq 0$ and $\lambda_2 \in \mathbb{R}$, $F(\lambda_1 f + \lambda_2 \underline{1}) = \lambda_1 F(f) + \lambda_2 F(\underline{1})$) and $F(\underline{1}) = 1$. Dually, every functional $F : C(X) \to \mathbb{R}$ with these properties is of the form ue_A for some convex closed nonempty set $A \in \mathcal{PD}(X)$.*

Note how Theorem 3 can be seen as specializing Theorem 4 for singletons $A = \{\mu\}$. The following is another useful specialization of Theorem 4.

Theorem 5. *Let X be a finite set, $A \in \mathcal{P}(X)$ a nonempty subset of X and $B = \overline{\mathrm{H}}(\{\delta(x) \mid x \in A\})$ the convex hull of Dirac's over A, which is closed since A is finite. The functional ue_B preserve sups, i.e., $ue_B(f \sqcup g) = ue_B(f) \sqcup ue_B(g)$ where $f \sqcup g$ denotes the pointwise l.u.b. defined as $(f \sqcup g)(x) = f(x) \sqcup g(x)$. Dually, every F as in Theorem 4 which preserves sups is of the form ue_B.*

Proof. The first part of the Theorem is simple to verify. For the other direction, let $C \subseteq \mathcal{D}(X)$ be the convex closed set, given by Theorem 4, such that $F = ue_C$. Suppose towards a contradiction that C is not of the form B. Let $Y = \{x \mid \exists \mu \in C.\mu(x) > 0\} \subseteq X$ so that every probability distribution in C is a convex combination of Dirac's δ_x for $x \in Y$. For each $x \in Y$ denote with $1_x : X \to \{0, 1\}$ the indicator function of x, so that $E_\mu(1_x) = \mu(x)$. Since C is not of the form B, there exists some $x \in Y$ such that $ue_C(1_x) < 1$. Let $\mu \in C$ maximizing 1_x in C, i.e., such that $ue_C(1_x) = E_\mu(1_x)$. The existence of μ follows from the compactness of C. Let $\lambda = \mu(x)$. Pick some other $y \in Y$ such that $\mu(y) > 0$ and define $g = \lambda 1_y$. It follows that $E_\mu(1_x \sqcup g) > \lambda$ and this implies $ue_C(1_x \sqcup g) > \lambda$. But, by construction, $ue_C(1_x) = \lambda$ and $ue_C(g) \leq \lambda$. A contradiction with the assumption that $F = ue_C$ preserves sups.

---

[6] Every function $X \to \mathbb{R}$ is continuous since X is assumed to be finite and endowed with the discrete topology.

[7] Recall that every norm of a finite dimensional space induces the same topology. In other words, the notion of closedness is norm-invariant in $\mathbb{R}^n$.

Note that these representation results deal with nonempty sets. For our purposes it is useful to generalize them to include also the representation of $ue_\emptyset$ ($f \mapsto 0$). Note that all functionals F considered so far are monotone and positively affinely homogeneous (p.a.h.).

Lemma 1. *Let X be a finite set and $F : C(X) \to \mathbb{R}$ be p.a.h. and monotone. If $F(\underline{1}) = 0$ then $F(f) = 0$, for all $f \in C(X)$.*

Proof. Find $\lambda \geq 0$ such that $-\lambda\underline{1} \sqsubseteq f \sqsubseteq \lambda\underline{1}$. By positive homogeneity, $F(-\lambda\underline{1}) = F(\lambda\underline{1}) = 0$. Apply monotonicity.

Thus, by relaxing the requirement $F(\underline{1}) = 1$ in the three representation theorems to $F(\underline{1}) \in \{0, 1\}$, we obtain functionals which can also represent the empty set.

5.2 The Riesz Modal Logic

In this section we introduce a novel $\mathbb{R}$-valued modal logic which we name **R** in honor of Frigyes Riesz who introduced what today are known as *lattice ordered vector spaces* or *Riesz spaces* [20]. We refer to [29,38] as standard references to the theory of Riesz spaces. Here we limit ourselves to the basic definitions and results.

Definition 21. *A* Riesz space *is a real (or rational) vector space R endowed with an order relation $\sqsubseteq$ which is a lattice and compatible with the vector space structure, i.e., if $f, g \in R$ then $f \sqsubseteq g$ implies $f + h \sqsubseteq g + h$ for all $h \in R$ and $f \sqsupseteq \underline{0}$ implies $\lambda f \sqsupseteq \underline{0}$, for all $\lambda \geq 0$, where $\underline{0}$ denotes the zero element of the vector space R. As usual $f \sqcup g$ denotes the least upper bound (lub) of f and g and similarly for $f \sqcap g$ (glb). It is always the case that $f \sqcap g = -(-f \sqcup -g)$. Note that $f \sqsubseteq g$ if and only if $f \sqcup g = g$.*

The vector space $X \to \mathbb{R}$, with functions ordered pointwise, is an example of Riesz space. The language of Riesz spaces, combining linear algebra and order theory, is sufficiently rich to express the representation theorems of Section 5.1.

Definition 22. *Let R be a Riesz space. The* positive part *of $f \in R$ is defined as $f^+ = f \sqcup \underline{0}$. We denote with R^+ the set $\{f^+ \mid f \in R\}$. We say R is* Archimedean *if, for all $f \in R^+$, the greatest lower bound of $\{(\frac{1}{n+1})f \mid n \in \mathbb{N}\}$ exists and equals $\underline{0}$. An element $f \in R^+$ is a* strong unit *if for all $g \in R^+$ there exists some $n \in \mathbb{N}$ such that $nf \sqsupseteq g$. We say R is* unitary *if it contains a strong unit.*

Note that $X \to \mathbb{R}$ is Archimedean and unitary with the constant function $\underline{1}$ ($x \mapsto 1$) as strong unit.

Definition 23 (Boolean Elements). *Let R be a Riesz space with strong unit u. Define the binary operation $\oplus$ of* truncated sum *as $f \oplus g = (f + g) \sqcap u$. An element $f \in R$ is* Boolean *if $f \oplus f = f$. Denote with B the set of Boolean elements in R. It is known (see, e.g., [47]) that $(B, \sqcup, \neg, \underline{u})$, with $\neg b = u - b$ is a Boolean algebra. If R is of the form $X \to \mathbb{R}$, with the pointwise order and strong unit $\underline{1}$, then B is the Boolean algebra of functions $X \to \{0, 1\}$.*

The following result, known as (finite dimensional) Yosida representation theorem, is fundamental in the theory of Riesz spaces.

Theorem 6 (p. 152, [38]). *Every Riesz space $(R, \sqsubseteq)$ which is n-dimensional as a vector space, Archimedean and unitary is of the form $X \to \mathbb{R}$ with $|X| = n$, pointwise ordering and strong unit $\underline{1}$.*

We now use this result, together with the representation theorems of Section 5.1, to develop a uniform algebraic account of PNTS's, Markov processes and NTS's (cf. Proposition 3). The following definition is similar to that of Boolean algebras with operators, the algebraic counterpart of standard modal logic [5].

Definition 24. *A* modal Riesz space *is a triple $(R, \sqsubseteq, \Diamond)$ where $(R, \sqsubseteq)$ is a Riesz space with strong unit (denoted by $\underline{1}$) and $\Diamond$ is an unary operation on R. We say that $(R, \sqsubseteq, \Diamond)$ is of* type-PNTS *if:*

(Monotone) if $f \sqsubseteq g$ then $\Diamond f \sqsubseteq \Diamond g$
(Sublinear) $\Diamond(f + g) \sqsubseteq \Diamond(f) + \Diamond(g)$
(Pos. Affine Homogenous) $\Diamond(\lambda_1 f + \lambda_2 \underline{1}) = \lambda_1 \Diamond(f) + \lambda_2 \Diamond(\underline{1})$
   for all $\lambda_1 \geq 0$ and $\lambda_2 \in \mathbb{R}$
($\Diamond \underline{1}$ is Boolean) $\Diamond \underline{1} \oplus \Diamond \underline{1} = \Diamond \underline{1}$

It is of type-MP *if it is of type-PNTS and further satisfies*

(Linear) $\Diamond(\lambda_1 f + \lambda_2 g) = \lambda_1 \Diamond(f) + \lambda_2 \Diamond(g)$, for all $\lambda_1, \lambda_2 \in \mathbb{R}$

and of type-NTS if it is of type-PNTS *and further satisfies*

(Sup-preserving) $\Diamond(f \sqcup g) = \Diamond(f) \sqcup \Diamond(g)$.

The following results should now be expected.

Proposition 5. *Let X be a finite set and (X, α) a PNTS such that $\alpha(x)$ is convex closed, for all $x \in X$. Let $R = (C(X), \sqsubseteq, \Diamond_\alpha)$, with $\sqsubseteq$ the pointwise order and $\Diamond_\alpha$ as in Equation 1. Then R is a finitely dimensional Archimedean modal Riesz space of type-PNTS. If (X, α) is a Markov process then R is of type-MP. If (X, α) is a NTS then R is of type-NTS.*

Theorem 7. *Let $(R, \sqsubseteq, \Diamond)$ be a finite dimensional Archimedean modal Riesz space of type-PNTS. Then there exists a unique (up to isomorphism) PNTS (X, α), with $\alpha(x)$ convex closed for all $x \in X$, such that $(R, \sqsubseteq, \Diamond)$ is of the form $(C(X), \sqsubseteq, \Diamond_\alpha)$. If R is of type-MP, then (X, α) is a Markov process. If R is of type-NTS, then (X, α) is a NTS.*

Proof. By Theorem 6 R is of the form $(X \to \mathbb{R})$ for some finite X. Of course the set X is unique up to isomorphism. Thus the $\Diamond$ operation has type $(X \to \mathbb{R}) \to (X \to \mathbb{R})$. By currying, consider the equivalent map $\diamond : X \to (X \to \mathbb{R}) \to \mathbb{R}$. For every $x \in X$, the map $\diamond(x)$ is monotone, subadditive, p.a.h. and $\diamond(\underline{1})(x) \in \{1, 0\}$. By Theorem 4 there exists a unique convex closed set C_x represented by $\diamond(x)$. The desired PNTS is then (X, α) with $\alpha(x) = C_x$. The cases for MP and NTS's are dealt in the same way by using Theorem 4 and Theorem 5.

The Riesz modal logic **R** which we now introduce is a syntactic counterpart of modal Riesz spaces.

Definition 25 (Logic R). *The formulas of the logic* **R** *are generated by the grammar:*

$$\phi ::= \underline{1} \mid \phi + \phi \mid q\phi \mid \phi \sqcup \phi \mid \Diamond\phi$$

where $q \in \mathbb{Q}$. We use the expected derived formulas $-\phi = (-1)\phi$, $\underline{0} = \underline{1} - \underline{1}$, $\phi \sqcap \psi = -(-\phi \sqcup -\psi)$, $\phi \oplus \psi = (\phi + \psi) \sqcap \underline{1}$, $\phi^{+} = \phi \sqcup \underline{0}$. We write $\phi \in \mathbf{R}$ if ϕ is a formula of the logic **R**.

Definition 26 (Model $\mathbb{R}$-valued Semantics). *Given a PNTS (X, α), we define the semantics $[\![\phi]\!]_{\alpha} : X \to \mathbb{R}$ of $\phi \in \mathbf{R}$ as:*

- $[\![\underline{1}]\!]_{\alpha}(x) = 1$
- $[\![\phi + \psi]\!]_{\alpha}(x) = [\![\phi]\!]_{\alpha}(x) + [\![\psi]\!]_{\alpha}(x)$
- $[\![\phi \sqcup \psi]\!]_{\alpha}(x) = \max\{[\![\phi]\!]_{\alpha}(x), [\![\psi]\!]_{\alpha}(x)\}$
- $[\![q\phi]\!]_{\alpha}(x) = q[\![\phi]\!]_{\alpha}(x)$
- $[\![\Diamond\phi]\!]_{\alpha} = \Diamond_{\alpha}([\![\phi]\!]_{\alpha})$, *with $\Diamond_{\alpha}$ as in Equation 1.*

Proposition 6 (Soundness). *Let (X, α) be a PNTS. For every formula ϕ and UE-bisimulation E the map $[\![\phi]\!]_{\alpha}$ is E-invariant.*

Proof. The proof goes by induction on the structure of ϕ. The interesting case is $[\![\Diamond\psi]\!]_{\alpha}$. By definition of UE-bisimilarity, $(x, y) \in E$ holds if for all E-invariant $f : X \to \mathbb{R}$, the equality $\bigsqcup_{x \to \mu} E_{\mu}(f) = \bigsqcup_{y \to \nu} E_{\nu}(f)$ holds. By induction hypothesis $[\![\psi]\!]_{\alpha}$ is E-invariant. Apply Definition 26.

Remark 11. Note that Proposition 6 can be rephrased by saying that, given a PNTS (X, α), every UE-bisimulation E is contained in the equivalence relation $\ker(\mathbf{R})$ defined as follows: $\ker(\mathbf{R}) = \{(x, y) \mid \forall\phi.([\![\phi]\!]_{\alpha}(x) = [\![\phi]\!]_{\alpha}(y))\}$.

We now prove that $\ker(\mathbf{R})$ is itself a UE-bisimulation and thus, by Proposition 6, the greatest one. The proof will also show that **R** is strongly complete in the sense of Definition 20. In the proof we make use of the following important result which, for the sake of simplicity, we specialize to finite sets (see, e.g., [29, §13]).

Proposition 7 (Stone-Weierstrass Theorem for Riesz spaces). *Let X be a finite set and R a Riesz subspace of $X \to \mathbb{R}$ that separates points, i.e., such that for every $x \neq y \in X$ there exists $f \in R$ such that $f(x) \neq f(y)$. Then R is dense in $X \to \mathbb{R}$.*

Theorem 8. *Let (X, α) be a PNTS. Then $\ker(\mathbf{R})$ is a UE-bisimulation.*

Proof. Denote with E the equivalence relation $\ker(\mathbf{R})$. Note that the map $[\![\phi]\!]_{\alpha}$, for every $\phi \in \mathbf{R}$, can be seen as a map of type $C(X/E) = X/E \to \mathbb{R}$ since, by definition, $[\![\phi]\!]_{\alpha}$ is E-invariant. For every $x, y \in X$ such that $(x, y) \notin E$ there exists some $\phi \in \mathbf{R}$ such that $[\![\phi]\!]_{\alpha}(x) \neq [\![\phi]\!]_{\alpha}(y)$. Thus **R** separates points in X/E. It is clear that **R** carries a Riesz space structure and that $\{[\![\phi]\!]_{\alpha} \mid \phi \in \mathbf{R}\}$ is a

Riesz subspace of $C(X/E)$. Thus, by Theorem 7, the (denotations of) formulas of $\mathbf{R}$ are dense $C(X/E)$.

We need to prove that if $(x, y) \in E$ then, for every E-invariant $f : X \to \mathbb{R}$, the equality $\bigsqcup_{x \to \mu} E_\mu(f) = \bigsqcup_{y \to \nu} E_\nu(f)$ holds. This equality can equivalently be expressed as $\big(\Diamond_\alpha(f)\big)(x) = \big(\Diamond_\alpha(f)\big)(y)$. Assume towards a contradiction that $\big(\Diamond_\alpha(f)\big)(x) > \big(\Diamond_\alpha(f)\big)(y)$. Since f is E-invariant, it can be seen as a map of type $X/E \to \mathbb{R}$. Since $\mathbf{R}$ is dense in this space, there exists $\phi \in \mathbf{R}$ (approximating f) such that $\big(\Diamond_\alpha(\phi)\big)(x) > \big(\Diamond_\alpha(\phi)\big)(y)$. By definition, this means that $[\![\Diamond\phi]\!]_\alpha(x) > [\![\Diamond\phi]\!]_\alpha(y)$. A contradiction with the hypothesis that $(x, y) \in \ker(\mathbf{R})$.

Corollary 1. *The logic $\mathbf{R}$ is strongly complete.*

We considered an $\mathbb{R}$-valued semantics for $\mathbf{R}$ but, in light of Proposition 5 and Theorem 7, an equivalent algebraic semantics can be defined.

Definition 27. *Let $(R, \sqsubseteq, \Diamond)$ be a modal Riesz space with strong unit denoted by u. Denote by $(\!|\_|\!)_R : \mathbf{R} \to R$ the expected homomorphism, from formulas to R, specified by $(\!|\underline{1}|\!) = u$. The element $(\!|\phi|\!)_R$ is called the* algebraic semantics *of ϕ interpreted in R.*

Definition 28 (Equational Theory). *For a PNTS (X, α), we write $(X, \alpha) \models \phi = \psi$ if $[\![\phi]\!]_\alpha = [\![\psi]\!]_\alpha$. For a collection $\mathcal{M}$ of PNTS's we write $\mathcal{M} \models \phi = \psi$ if $(X, \alpha) \models \phi = \psi$ for all $(X, \alpha) \in \mathcal{M}$. Similarly, for a modal Riesz space R, we write $R \Vdash \phi = \psi$ if $(\!|\phi|\!)_R = (\!|\psi|\!)_R$ and, for a collection $\mathcal{R}$ of modal Riesz space, we write $\mathcal{M} \models \phi = \psi$ if $R \Vdash \phi = \psi$ for all $R \in \mathcal{R}$.*

The following is a direct consequence of Proposition 5 and Theorem 7.

Corollary 2. *Let $\mathcal{M}$ be the collection of all PNTS (resp. Markov processes, NTS's) (X, α) with X finite and $\alpha(x)$ closed. Let $\mathcal{R}$ be the collection of finite dimensional Archimedean modal Riesz space of type-PNTS (resp. type-MP, type-NTS). Then for all $\phi, \psi \in \mathbf{R}$ it holds that $\mathcal{M} \models \phi = \psi$ if and only if $\mathcal{R} \Vdash \phi = \psi$.*

As we anticipated in the introduction, we are not aware of any other $\mathbb{R}$-valued or Boolean (i.e., $\{0, 1\}$-valued) logic for PNTS's equipped with an algebraic (i.e., axiomatic) semantics. In the context of Markov processes, on the other hand, the Boolean logic of Panangaden *et. al.* [32] is completely axiomatized and a duality between (generally infinite) models and algebras is proved. It is going to be interesting to study the relation between their logic and $\mathbf{R}$. With respect to other logics for Markov processes, we mention that the $\mathbb{R}$-valued (in fact $[0, 1]$-valued) logic of [49] can be easily seen as a fragment of $\mathbf{R}$.

Note that $\mathbb{R}$-valued logics are interesting also in the context of NTS's, despite the fact that ordinary Boolean ones (such as the modal logic $\mathbf{K}$) suffice to characterize Milner and Park's bisimilarity. See, e.g., [19], for a (fixed-point) $\mathbb{R}$-valued modal logic, for expressing quantitative properties of ordinary NTS's. Thus, in light of its algebraic presentation and completeness result, the logic $\mathbf{R}$ is also useful as a foundation for quantitative logics for NTS's. Note that the standard modal logic $\mathbf{K}$ can be seen as a Boolean fragment of $\mathbf{R}$ (cf. Definition

23). The logic $\mathbf{K}$ can be axiomatized by the axioms of Boolean algebra plus $\phi \sqsupseteq \psi \Rightarrow \Diamond\phi \sqsupseteq \Diamond\psi$ (monotonicity), $\Diamond\underline{0} = \underline{0}$ (which follows easily since $\Diamond$ is a p.a.h. functional) and $\Diamond(\phi \sqcup \psi) = \Diamond(\phi) \sqcup \Diamond(\psi)$ (cf. Axiom of sup-preservation) [5]. We plan to study the relations between $\mathbf{K}$ and $\mathbf{R}$ in future works.

The correspondence between model and algebraic semantics, defined on three important classes of systems, is likely to be a valuable tool for future investigations. As an example, we suggest that the algebraic axiomatization could be of help towards design of *structural proof systems* (e.g., *sequent calculi*) for the logic $\mathbf{R}$. Such systems proved to be useful in the development of techniques for *compositional verification* of nondeterministic and probabilisitic systems. See, e.g., [52] for a paradigmatic (but restricted to NTS's) example on this line of research based on the standard modal logic $\mathbf{K}$ and, e.g., [45,8], for recent developments.

Remark 12. We freely adopted the term *logic* for the real valued formalism of $\mathbf{R}$. This is well established terminology in computer-science literature (see, e.g., [41,28,19]) as well as in mathematical logic. For example, Łukasiewicz logic [47] and, more generally, *fuzzy logics*, are well known examples of real-valued formalisms. Furthermore, if a logic is understood as a language for expressing properties of models, then $\mathbf{R}$ is rightfully a logic in that its formulas denote *experiments* to be performed on models, and these can be regarded as *quantitative predicates*.

5.3 Probabilistic Modal μ-Calculi

The choice of primitives of the logic $\mathbf{R}$ has been guided by the Representation theorems of Section 5.1. The language of Riesz spaces is a simple, yet mathematically rich [29,38,20], formalism capable of formulating these powerful results from functional analysis. We see the value of the logic $\mathbf{R}$ mainly as a tool for theoretical investigations. However, as many other basic modal logics for NTS's including the modal logic $\mathbf{K}$, the logic $\mathbf{R}$ cannot express many interesting properties of systems such as: termination goals, safety and liveness constraints, *etc.* Richer logics, capable of expressing these properties, can be obtained by enriching the basic modal systems with additional operators. The modal μ-calculus of Dexter Kozen [31] is a very expressive logic for NTS's obtained by extending $\mathbf{K}$ with least (μ) and greatest (ν) fixed point operators [55,7]. The modal μ-calculus enjoys a remarkably rich theory [1].

In the last decade several fixed point modal logics (henceforth, *probabilistic μ-calculi*) for expressing properties of PNTS's have been considered. We refer to [43] as a reference to this area of research. Among the different approaches that have been followed to developing analogues of the modal μ-calculus, the most relevant is based on enriching some basic quantitative $\mathbb{R}$-valued modal logic with least and greatest fixed point operators.

Definition 29. *The formulas of the logics* qL, $\mathrm{qL}^{\odot}$, $\mathrm{qL}^{\ominus}$ *and* L *are defined as:*

$$
\begin{array}{ll}
\text{(qĿ)} & \phi ::= \underline{1} \mid \phi \sqcup \phi \mid q\phi \mid \neg\phi \mid \Diamond\phi \\
\text{(qĿ}^{\odot}\text{)} & \phi ::= \underline{1} \mid \phi \sqcup \phi \mid q\phi \mid \neg\phi \mid \Diamond\phi \mid \phi \cdot \phi \\
\text{(qĿ}^{\ominus}\text{)} & \phi ::= \underline{1} \mid \phi \sqcup \phi \mid q\phi \mid \neg\phi \mid \Diamond\phi \mid \phi \ominus q \\
\text{(Ł)} & \phi ::= \underline{1} \mid \phi \sqcup \phi \mid q\phi \mid \neg\phi \mid \Diamond\phi \mid \phi \oplus \phi
\end{array}
$$

where $q \in [0,1] \cap \mathbb{Q}$ is a rational number. The semantics of a formula ϕ interpreted in a PNTS (X, α) is defined as a map $[\![\phi]\!]_{\alpha} : X \to \mathbb{R}$ specified as (subscripts α are omitted): $[\![\underline{1}]\!]$, $[\![\phi \sqcup \psi]\!]$, $[\![q\phi]\!]$ and $[\![\Diamond\phi]\!]$ are defined as in $\boldsymbol{R}$; $[\![\neg\phi]\!](x) = 1 - [\![\phi]\!](x)$, $[\![\phi \cdot \psi]\!](x) = [\![\phi]\!](x)[\![\psi]\!](x)$ (product on reals), $[\![\phi \ominus q]\!](x) = \max\{0, [\![\phi]\!](x) - q\}$, and $[\![\phi \oplus \psi]\!](x) = \min\{1, [\![\phi]\!](x) + [\![\psi]\!](x)\}$.

Note that qĿ$^{\ominus}$ is a fragment of Ł ($\phi \ominus q$ can be encoded as $\neg(\neg\phi \oplus q)$) and that Ł is a fragment of $\boldsymbol{R}$ ($\phi \oplus \psi$ can be encoded as $(\phi + \psi) \sqcap \underline{1}$). The operations of the logic Ł coincide with those available[8] in Łukasiewicz logic [47] and this justifies the choice of notation. The logic $\boldsymbol{R}$ and qĿ$^{\odot}$, on the other hand, appear to be incomparable. Furthermore, observe that $[\![\phi]\!]_{\alpha}$ ranges over $[0,1]$, i.e., has always type $[\![\phi]\!]_{\alpha} : X \to [0,1]$, and that the interpretation of every connective (except $\neg$) is pointwise monotone. The latter two observations allow one to develop a theory of fixed points on top of these logics based on the order-theoretic Knaster-Tarski fixed-point theorem. The syntax of the logics is extended to include *variables* Var, ranged over by the letter v, and least $(\mu v.\phi)$ and greatest $(\nu v.\phi)$ fixed point operators which, as expected, bind the variable v in ϕ. As customary in fixed-point logics [55], one restricts attention to *positive formulas*, where every occurrence of variable v occurs in the scope of an even number of negations. The semantics of formulas is defined as follows.

Definition 30. *For a PNTS (X, α), an interpretation of the variables is an assignment $\rho : \mathtt{Var} \to (X \to [0,1])$. For every $f : X \to [0,1]$ we write $\rho[f/v]$ for the interpretation specified as $\rho[f/v](w) = f$ if $w = v$ and $\rho[f/v](w) = \rho(w)$ otherwise. The semantics of a fixed point formula ϕ interpreted in (X, α) with ρ is the map specified extending Definition 29 with: $[\![v]\!]_{\rho} = \rho(v)$, $[\![\mu v.\phi]\!]_{\rho} = \mathrm{lfp}(f \mapsto [\![\phi]\!]_{\rho[f/v]})$ and, similarly, $[\![\nu v.\phi]\!]_{\rho} = \mathrm{gfp}(f \mapsto [\![\phi]\!]_{\rho[f/v]})$, where f ranges over $X \to [0,1]$.*

Since $X \to [0,1]$ (ordered pointwise) is a complete lattice, least and greatest points of arbitrary monotone operators exists. We denote with qĿμ, qĿ$\mu^{\odot}$, qĿ$\mu^{\ominus}$ and Łμ the μ-calculi obtained in this way. Historically, qĿμ has been the first logic of this family to be studied [28,41]. The logic qĿ$\mu^{\ominus}$ has been considered in, e.g., [14,15] and is based on the quantitative modal logic (for Markov processes) of Panangaden [49]. It is possible to show (see, e.g., [44] that these logics are not sufficiently expressive to encode other important temporal logics for PNTS's such as the *probabilistic CTL* (PCTL) of [4]. To address this limitation, the two logics qĿ$\mu^{\odot}$ and Łμ are introduced[9] in [43]. The author shows how qĿ$\mu^{\odot}$ and Łμ can encode the *qualitative fragment* of PCTL and *full* PCTL respectively. Our first result of this section is the soundness, with respect to UE-bisimilarity, of all μ-calculi discussed above.

---

[8] Actually, the operation of scalar multiplication by $q \in [0,1]$, gives Ł a so-called *Riesz MV-algebra* structure. See, e.g., [48].

[9] The logic Łμ is called pĿ$\mu_{\oplus}$ in [43].

Proposition 8 (Soundness). *Let (X, α) be a PNTS, E an UE-bisimulation and ρ an interpretation such that $\rho(v)$ is E-invariant for all $v \in \mathtt{Var}$. If $(x, y) \in E$ then, for every closed formula ϕ of $qL\mu$ it holds that $\llbracket \phi \rrbracket_\rho(x) = \llbracket \phi \rrbracket_\rho(y)$. Similarly for $qL\mu^\ominus$, $qL\mu^\odot$ and $L\mu$ formulas.*

Proof. The proof is by induction of the complexity of ϕ as in Proposition 6. For the case $\phi = \mu v.\psi$, by Knaster-Tarski theorem, we have $\llbracket \phi \rrbracket_\rho = \llbracket \phi \rrbracket_\rho^\alpha$, where α ranges over the ordinals and $\llbracket \psi \rrbracket_\rho^{\alpha+1} = \llbracket \psi \rrbracket_{\rho[\llbracket \psi \rrbracket_\rho^\alpha / v]}$ and $\llbracket \psi \rrbracket_\rho^\beta = \bigsqcup_{\alpha < \beta} \llbracket \psi \rrbracket_\rho^\beta$, where β is a limit ordinal. It is clear that the pointwise supremum of a familiy E-invariant functions is E-invariant. Furthermore, by induction hypothesis on ψ and α, also $\llbracket \psi \rrbracket_\rho^{\alpha+1}$ is E-invariant. Thus $\llbracket \phi \rrbracket$ is E-invariant as desired. The case for $\phi = \nu v.\psi$ is similar.

It is shown in [15] that $qL\mu^\ominus$ (or even its fixed-point free fragment $qL^\ominus$) is weakly complete with respect to UE-bisimilarity: if $\llbracket \phi \rrbracket(x) = \llbracket \phi \rrbracket(y)$ for all $qL\mu^\ominus$ formulas ϕ, then x and y are UE-bisimilar. However the logic $qL\mu^\ominus$ is not strongly complete (cf. Section 5.4 below) in the sense[10] of Definition 20. We now prove that, instead, the logics $qL\mu^\odot$ and L, and thus also their fixed-point extensions, are strongly complete.

Theorem 9. *The logics L and $qL^\odot$ are strongly complete.*

Proof. The proof for L is based on a Stone-Weierstrass Theorem for RMV-algebras (see, e.g., [48]), stating that every RMV-subalgebra (cf. Footnote 8) of $X \to [0, 1]$ that separates points is dense in $X \to [0, 1]$. The proof is then identical to that of Theorem 8. For what concerns $qL^\odot$ it is shown in [43] (Lemma 3.3.16) that for for every $f, g \in X \to [0, 1]$, the function $f \oplus g$ can be approximated by a sequence of functions expressible in the language of $qL^\odot$.

Thus the expressive logics $qL\mu^\odot$ and $L\mu$ can be used to denote (approximations of) all possible $[0, 1]$-valued experiments (cf. Example 1). This, together with the fact that $L\mu$ can encode the logic PCTL, provides a strong characterization of UE-bisimilarity in terms of expressive, and potentially useful in practice, temporal logics for verification.

5.4 Logical Characterization of the Hausdorff Metric

As mentioned above, the logic $qL\mu^\ominus$ of [15] (and the similar $[0, 1]$-valued logic of Panangaden [49]) is not strongly complete but just weakly complete. The connectives of $qL\mu^\ominus$ are carefully chosen so that the denotation of formulas are Lipschitz (i.e., not expansive) functions (see [49]). This property is crucial

---

[10] Actually, Definition 20 requires (the denotation of) formulas to be dense in $X \to \mathbb{R}$. In the context of $[0, 1]$-value logics, we instead require density in $X \to [0, 1]$. Note how this adapted notion of strong completeness still implies weak completeness.

in proving that the logically defined *behavioral metric*[11] on states of a PNTS (X, α)

$$d_L(x, y) = \bigsqcup_{\phi \in \mathrm{qL}\mu^{\ominus}} |[\![\phi]\!]_\alpha(x) - [\![\phi]\!]_\alpha(y)| \tag{2}$$

coincides with the so-called *Hausdorff* behavioral metric which is based on the following idea. For a given PNTS (X, α), and for any[12] metric m on $\mathcal{D}(X)$, one can consider the Hausdorff metric d_H^m on the space of closed subsets of $\mathcal{D}(X)$ (see, e.g., [30]). Then it is natural to define a metric on states as $d(x, y) = d_H^m(\alpha(x), \alpha(y))$. When m is the Kantorovich metric (see, e.g., [30]) on $\mathcal{D}(X)$, then d_H^m and d_L, as above, coincide. This, as claimed earlier, implies that $\mathrm{qL}\mu^{\ominus}$ is not strongly complete: take x, y such that $0 < d(x, y) < 1$ so that there is no $\phi \in \mathrm{qL}\mu^{\ominus}$ such that $[\![\phi]\!](x) = 1$ and $[\![\phi]\!](y) = 0$.

The introduction of behavioral metrics for probabilistic systems is strongly motivated by the need of *approximation methods* (based on the idea that small changes of the probabilities in a PNTS corresponds to small changes in behavior), a line of research pioneered by Panangaden [49]. The fact that $\mathrm{qL}\mu^{\ominus}$ characterizes the Hausdorff behavioral metric is thus a remarkable property. Given the strong completeness of L (or **R**, $\mathrm{qL}^{\odot}$), no interesting (i.e., not discrete) metric can be obtained as in (2). We now show, however, that also the richer logic L (and **R**, $\mathrm{qL}^{\odot}$) can be used to logically characterize the *Hausdorff* behavioral metric in an interesting fashion. This is once again proven by applying results from linear algebra.

Theorem 10. *Let (X, α) be a PNTS. Define d_L as:*

$$d_L(x, y) = \bigsqcup_{\phi \in \mathrm{L}} |[\![\Diamond\phi]\!]_\alpha(x) - [\![\Diamond\phi]\!]_\alpha(y)|$$

Then d_L is (equivalent to) $d(x, y) = d_H^m(\alpha(x), \alpha(y))$, where m is the total variation metric.

Proof. Assume without loss of generality (by Theorem 1) that $\alpha(x)$ is convex closed for all $x \in X$. It is known (see, e.g., [37, §8.4]) that for convex closed sets the equality $d_H^m(A, B) = \bigsqcup_{\|f\|_\infty \leq 1} |ue_A(f) - ue_B(f)|$ holds, where $\|f\|_\infty \leq 1$ means that f has sup-norm ≤ 1, i.e., $f : X \to [-1, 1]$. If $A = \emptyset$ then $d_L(x, y) = d(x, y)$, as it is simple to verify. Thus assume $A, B \neq \emptyset$. Since ue_A and ue_B are positive affinely homogeneous, for every $\|f\| \leq 1$ we have, for $g = \frac{1}{2}f + \frac{1}{2}$, that $ue_A(g) = \frac{1}{2}ue_A(f) + \frac{1}{2}$ and similarly for $ue_B(g)$. Note that $g : X \to [0, 1]$. Thus $|ue_A(g) - ue_B(g)| =$

---

[11] Technically, $d(x, y)$ is a pseudo-metric because if x, y are UE-bisimilar then $d(x, y) = 0$, even if $x \neq y$. The function $d(x, y)$ is an authentic metric on the state space quotiented by UE-bisimilarity (cf. Remark 2). To ignore this pedantic distinction, in this section we just simple assume that the largest UE-bisimulation in (X, α) is the identity relation.

[12] Many natural metrics on $\mathcal{D}(X)$ exist. E.g., any norm on $\mathbb{R}^n \cong X \to \mathbb{R}$ (cf. Convention 41) induces a metric on $\mathcal{D}(X)$. The sup-norm on $\mathbb{R}^n$ induces the so-called *total variation* metric on $\mathcal{D}(X)$. See also Remark 7.

$\frac{1}{2}|ue_A(f) - ue_B(f)|$. Hence the metric $d'(A, B) = \bigsqcup_{g:X\to[0,1]} |ue_A(g) - ue_B(g)|$ is equivalent to $d_H^m(A, B)$, i.e., the two metrics induce the same topology. The desired result follows by density of L in $X \to [0,1]$ and the fact that $ue_{\alpha(x)}(\llbracket\phi\rrbracket) = \llbracket\Diamond\phi\rrbracket_\alpha(x)$.

Remark 13. Due to the lack of space, we just remark here that, given a finite PNTS (X, α), it is simple to calculate the value $d_L(x, y)$ as in Theorem 10 by means of Linear Programming.

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 [10], in future research.

6   Congruence for PGSOS

A *congruence* on an algebra $\mathcal{A}$ is an equivalence relation that respects the constructors of the algebraic structure: if $a_i \equiv b_i$, for $i \in \{1, \dots, n\}$, then $f(\boldsymbol{a}) \equiv f(\boldsymbol{b})$ for any operation f of arity n in the signature of $\mathcal{A}$. When $\mathcal{A}$ is a collection of programs and each f is a program constructor, a congruence relation captures a notion of behavioral equivalence and validates the principle that substituting behaviorally equal sub-programs a_i of a compound program $f(\boldsymbol{a})$ with behaviorally equivalent sub-programs b_i, results in a program $f(\boldsymbol{b})$ that is equivalent to $f(\boldsymbol{a})$. We show in this section that UE-bisimilarity is a congruence with respect to the important program constructor of (communicating) parallel composition. The proof is valid for arbitrary systems (i.e., potentially infinite PNTS's) and can readily be adapted to show that UE-bisimilarity is a congruence with respect to the wide family of program constructors specified following the PGSOS rule format of [3] which virtually includes all CCS-style operators of practical interest. To make our discussion interesting, we now consider *labeled* PNTS's, i.e., structures $(X, \{\alpha_a\}_{a\in L})$ such that (X, α_a) is a PNTS, for every label $a \in L$. We define UE-bisimulations on labeled PNTS's as expected.

Definition 31. *Given a labeled PNTS $(X, \{\alpha_a\}_{a\in L})$, an equivalence relation $E \subseteq X \times X$ is a UE-bisimulation if it is a UE-bisimulation for the PNTS (X, α_a), for all $a \in L$.*

The operational semantics of the parallel operator operator ($||$) is specified as expected where, as customary, we consider a set of labels L containing a distinct label τ and pairs a complementary labels $a, \overline{a}$ (with $\overline{\overline{a}} = a$).

$$\frac{x \xrightarrow{a} \mu}{x||y \xrightarrow{a} \mu||y} \ (||\ \text{L}) \qquad \frac{y \xrightarrow{a} \mu}{x||y \xrightarrow{a} x||\mu} \ (||\ \text{R}) \qquad \frac{y \xrightarrow{a} \mu \qquad y \xrightarrow{\overline{a}} \nu}{x||y \xrightarrow{\tau} \mu||\nu} \ (||\ \text{Comm})$$

Following standard approaches (see, e.g., [52]) this leads to the corresponding notion of process algebra for $||$.

Definition 32. *A* process algebra *for the parallel operator $\|$ is a labeled PNTS $(X, \{\alpha_a\}_{a \in L})$ together with an interpretation ι of $\|$, i.e., a function $\iota : X \times X \to X$, such that $\iota(x, y) \xrightarrow{a} \mu$ holds iff:*

1. *($\| L$): $x \xrightarrow{a} \mu_1$ and $\mu = i(\mu_1 \times \delta_y)$, or*
2. *($\| R$): $y \xrightarrow{a} \mu_1$ and $\mu = i(\delta_x \times \mu_1)$, or*
3. *($\| Comm$): $a = \tau$, $x \xrightarrow{b} \mu_1$, $y \xrightarrow{\bar{b}} \mu_2$ and $\mu = i(\mu_1 \times \mu_2)$*

hold, where $x \xrightarrow{a} \mu$ iff $\mu \in \alpha_a(x)$, $\mu \times \nu$ denotes the product probability distribution $((x, y) \mapsto \mu(x)\nu(y))$ and $i(\mu \times \nu)$ is defined as $i(\mu \times \nu)(x) = \sum_{z \in \iota^{-1}(x)} (\mu \times \nu)(z)$, with $z \in X \times X$.

Theorem 11. *For a process algebra $\langle (X, \{\alpha_a\}_{a \in L}), \iota \rangle$, the UE-bisimilarity relation is a congruence for $\|$.*

Proof. Let E be the relation of UE-bisimilarity, i.e., the greatest UE-bisimulation on X (cf. Theorem 8). In what follows we just write $x \| y$ in place of $i(x, y)$. Define the relation $R = \{ (x \| y, x' \| y), (y \| x, y \| x') \mid (x, x') \in E, \ y \in X \}$. Since E is an equivalence relations so is R. We prove the theorem by showing that R is a UE-bisimulation and, as such, contained in E. We need to prove that for every $(z, z') \in R$ and every R-invariant function $f : X \to \mathbb{R}$ the following two implications hold:

i if $z \xrightarrow{a} \mu$ then for all $\epsilon > 0$, there is $z' \xrightarrow{a} \nu$ such that $E_\mu(f) \leq E_\nu(f) + \epsilon$,

ii if $z' \xrightarrow{a} \nu$ then for all $\epsilon > 0$ there is $z \xrightarrow{a} \mu$ such that $E_\nu(f) \leq E_\mu(f) + \epsilon$,

for all $a \in L$. Note how we reformulated $ue_{\alpha_a(z)}(f) = ue_{\alpha_a(z')}(f)$ as the equivalent conjunction of (i) and (ii).

We just consider the case $z = x \| y$ and $z' = x' \| y$, with $(x, x') \in E$, and show (i) above. The other cases are handled in a similar way. If $a \neq \tau$, we need to distinguish two cases:

1. ($\| R$): $y \xrightarrow{a} \mu_1$ and $\mu = x \| \mu_1$, and
2. ($\| L$): $x \xrightarrow{a} \mu_1$ and $\mu = \mu_1 \| y$

where we simply denoted with $x \| \mu_1$ and $\mu_1 \| y$ the probability measures $\iota(\delta_x \times \mu_1)$ and $\iota(\mu_1 \times \delta_y)$, respectively.

In the case ($\| R$), choose ν as $\nu = x' \| \mu_1$. It is then the case that $\mu(x \| y') = \nu(x' \| y')$, for all $y' \in X$. Since $(x, x') \in E$ and f is E-invariant, the equality $f(x \| y) = f(x' \| y)$ holds, for all $y \in X$. It then follows that $E_\mu(f) = E_\nu(f)$.

Consider now the case ($\| L$). Define $g_y : X \to [0, 1]$ as $g_y(z) = f(z \| y)$. Note that g_y is E-invariant because f is R-invariant. Since $(x, x') \in E$, there exists some $x' \xrightarrow{a} \nu_1$ such that $E_{\mu_1}(g_y) \leq E_{\nu_1}(g_y) + \epsilon$. Take $\nu = \nu_1 \times \delta_y$ so that $x' \| y \to \nu$ holds. It is straightforward to verify that $E_{\mu_1}(g_y) = E_\mu(f)$ and, similarly, $E_{\nu_1}(g_y) = E_\nu(f)$. Thus the desired result follows.

If $a = \tau$, we need to consider, in addition to the cases ($\| R$) and ($\| L$), the third case triggered by the rule ($\| Comm$). Thus assume

3. ($\| Comm$): $x \xrightarrow{a} \mu_1$, $y \xrightarrow{\bar{a}} \mu_2$ and $\mu = \mu_1 \| \mu_2$.

Define $g : X \to [0, 1]$ as $g(x) = \sum_y \mu_2(y) f(x \| y)$. Note, as above, that g is E-invariant because f is R-invariant. Since $(x, x') \in E$, there exists some $x' \xrightarrow{a} \nu_1$ such that $E_{\mu_1}(g) \leq E_{\nu_1}(g) + \epsilon$. As ν take the probability measure $\nu = \nu_1 \| \mu_2$. From definitions we have $E_{\mu_1}(g) = E_\mu(f)$ and $E_{\nu_1}(g) = E_\nu(f)$, as desired.

7 Generalization to Infinite Systems

The important assumption of Convention 41 allowed a significant simplification of the results presented in this work. The theory can be generalized to cover infinite systems at the cost of topological complications. The generalized setting, however, allow one to fully appreciate the powerful mathematical set-up, based on linear algebra and functional analysis, on which the results of this work are developed. Here, we just sketch the main ideas.

One can define PNTS's on the category of compact Hausdorff topological spaces by replacing $\mathcal{P}$ with the functor $\mathcal{V}$ (mapping a space X to the space $K(X)$ of its compact closed subsets, endowed with the Vietoris topology [30]) and $\mathcal{D}$ by $\mathcal{M}_{=1}$ (mapping a space X to the space of probability measures on X, endowed with the weak*-topology [30]). Experiments on X are now modeled by the space $C(X)$ of continuous functions $X \to \mathbb{R}$. This is a Riesz (Banach) space when endowed with the sup-norm [37] and ordered pointwise. Both Theorem 1 and the representation theorems of Section 5.1 can be generalized to the new setting (see, e.g., [57]). Similarly, counterparts of Theorem 6 and Theorem 10 hold for general compact Hausdorff spaces (see, e.g., [38,29] and [37]). The results of Section 5 can then be proved *mutatis mutandis*.

References

1. A. Arnold and D. Niwinski. *Rudiments of μ-calculus*. Studies in Logic. North-Holland, 2001.
2. C. Baier and J. P. Katoen. *Principles of Model Checking*. The MIT Press, 2008.
3. F. Bartels. *On Generalised Coinduction and Probabilistic Specification Formats : Distributive Laws in Coalgebraic Modelling*. PhD thesis, Centrum voor Wiskunde en Informatica, Amsterdam, 2004.
4. A. Bianco and L. de Alfaro. Model checking of probabilistic and nondeterministic systems. In *Foundations of Software Technology and Theoretical Computer Science*, number 1026 in Lecture Notes in Computer Science, pages 499–513. Springer-Verlag, 1995.
5. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge Tracts in Theoretical Computer Science, 2002.
6. B. Bloom, S. Istrail, and A. R. Meyer. Bisimulation can't be traced. *Journal of the ACM (JACM)*, 42:232–268, 1995.
7. J. C. Bradfield and C. Stirling. Modal logics and mu-calculi: an introduction. In *Handbook of Process Algebra*. Elsevier, 2001.
8. L. Cardelli, K. Larsen, and M. Radu. Modular Markovian logic. In *Proceedings of the 38th ICALP*, pages 380–391, 2011.

9. S. Cattani and R. Segala. Decision algorithm for probabilistic bisimulation. In *CONCUR*, 2012.

10. K. Chatzikokolakis, C. Palamidessi, and P. Panangaden. On the Bayes risk in information-hiding protocols. *Journal of Computer Security*, 16(5), 2008.

11. S. Crafa and F. Ranzato. A spectrum of behavioral relations over LTS's on probability distributions. In *CONCUR*, 2001.

12. M. Dam and D. Gurov. Compositional verification of ccs processes. In *PSI 1999*, pages 247–256, London, UK, 2000. Springer-Verlag.

13. V. Danos, J. Desharnais, F. Laviolette, and P. Panangaden. Bisimulation and cocongruence for probabilistic systems. *Information and Computation*, 204(5):503–523, 2006.

14. L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. *Journal of Computer and System Sciences, Volume 68, Issue 2*, pages 374 – 397, 2004.

15. L. de Alfaro, R. Majumdar, V. Ramanc, and M. Stoelinga. Game refinement relations and metrics. *Logical Methods in Computer Science*, 4, 2008.

16. Y. Deng and R. van Glabbeek. Characterising probabilistic processes logically. In *Logic for programming, artificial intelligence and reasoning*, volume 6397 of *Lecture Notes in Computer Science*, 2010.

17. Y. Deng, R. van Glabbeek, M. Hennessy, and C. Morgan. Testing finitary probabilistic processes. In *CONCUR*, 2009.

18. J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Weak bisimulation is sound and complete for PCTL. In *Proceedings of the 13nd International Conference on Concurrency Theory*, volume 2421 of *Lecture Notes in Computer Science*, pages 355–370, 2002.

19. D. Fischer, E. Grädel, and Ł. Kaiser. Model checking games for the quantitative μ-calculus. *Theory of Computing Systems*, 47(3), 2010.

20. R. Frigyes. Sur la décomposition des opérations fonctionelles linéaires. *Atti congresso internazionale matematici (Bologna 1928)*, Zanichelli:143–148, 1930.

21. J. Goubault-Larrecq. Prevision domains and convex powercones. In *Foundations of Software Science and Computation Structures*, 2008.

22. J. F. Groote and F. Vaandrager. Structured operational semantics and bisimulation as a congruence. In *Automata, Languages and Programming*, Lecture Notes in Computer Science, 1989.

23. J. Y. Halpern and R. Pucella. Characterizing and reasoning about probabilistic and non-probabilistic expectation. *Journal of the ACM*, 54(3), 2007.

24. H. Hansson. *Time and Probability in Formal Design of Distributed Systems*. Elsevier, 1994.

25. M. Hennessy and R. Milner. Algebraic laws for nondeterminism and concurrency. *Journal of the ACM (JACM)*, 32:137–161, 1985.

26. O. M. Herescu and C. Palamidessi. Probabilistic asynchronous π-calculus. In *Proceedings of FOSSACS 2000*, pages 146–160, 2000.

27. P. Huber. *Robust Statistics*. Wiley, 1991.

28. M. Huth and M. Kwiatkowska. Quantitative analysis and model checking. In *Proc. of 12th LICS*, 1997.

29. E. D. Jonge and A. V. Rooij. *Introduction to Riesz Spaces*, volume 78. Mathematical Centre Tracts, Amsterdam, 1977.

30. A. S. Kechris. *Classical Descriptive Set Theory*. Graduate Texts in Mathematics. Springer Verlag, 1994.

31. D. Kozen. Results on the propositional mu-calculus. In *Theoretical Computer Science*, pages 333–354, 1983.

32. D. Kozen, K. G. Larsen, R. Mardare, and P. Panangaden. Stone duality for Markov processes. In *Proceeding of LICS*, 2013.
33. A. Kurz. *Logics for Coalgebras and Applications to Computer Science*. PhD thesis, Ludwig Maximilian University of Munich, 2000.
34. M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: Verification of probabilistic real-time systems. In *CAV*, LNCS, 2011.
35. M. Kwiatkowska, G. Norman, D. Parker, and M. Vigotti. Probabilistic mobile ambients. *Theoretical Computer Science*, pages 12–13, 2009.
36. K. G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94:1–28, 1991.
37. P. Lax. *Functional Analysis*. Wiley Interscience, 2002.
38. W. A. J. Luxemburgh and A. C. Zaanen. *Riesz Spaces*, volume 1. North-Holland Mathematical Library, 1971.
39. N. Lynch and R. Segala. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2:250–273, 1995.
40. M. L. Marco Bernardo, Rocco De Nicola. Revisiting trace and testing equivalences for nondeterministic and probabilistic processes. In *Proceeding of FoSSaCS*, 2012.
41. A. McIver and C. Morgan. Results on the quantitative μ-calculus qMμ. *ACM Transactions on Computational Logic*, 8(1), 2007.
42. R. Milner. *A Calculus of Communicating Systems*. Springer-Verlag New York, Inc., 1982.
43. M. Mio. *Game Semantics for Probabilistic μ-Calculi*. PhD thesis, School of Informatics, University of Edinburgh, 2012.
44. M. Mio. Probabilistic Modal μ-Calculus with Independent product. *Logical Methods in Computer Science*, 8(4), 2012.
45. M. Mio and A. Simpson. A proof system for reasoning about probabilistic concurrent processes. In *Proceedings of 16th FoSSaCS*, 2013.
46. M. Mislove. Nondeterminism and probabilistic choice: Obeying the law. In *CONCUR*, 2000.
47. D. Mundici. *Advanced Lukasiewicz Calculus and MV-Algebras*. Trends in Logic. Springer-Verlag, 2011.
48. A. D. Nola and I. Leustean. Riesz MV-algebras and their logic. In *Proceeding of EUSFLAT*, 2011.
49. P. Panangaden. *Labelled Markov processes*. Imperial College Press, 2009.
50. D. Sangiorgi and J. Rutten. *Advanced topics in bisimulation and coinduction*, volume 52 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2011.
51. R. Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, M.I.T., 1995.
52. A. Simpson. Sequent calculi for process verification: Hennessy-Milner logic for an arbitrary GSOS. *Journal of Logic and Algebraic Programming*, 60-61:287, 2004.
53. A. Sokolova. Probabilistic Systems Coalgebraically: A survey. *Theoretical Computer Science*, 412(38), 2011.
54. L. Song, L. Zhang, and J. C. Godskesen. Bisimulations Meet PCTL Equivalences for Probabilistic Automata. In *CONCUR*, 2011.
55. C. Stirling. *Modal and temporal logics for processes*. Springer, 2001.
56. R. Tix, K. Keimel, and G. D. Plotkin. Semantic domains for combining probability and non-determinism. *Electronic Notes in Theoretical Computer Science*, pages 1–104, 2005.
57. P. Walley. *Statistical reasoning with imprecise probabilities*. Chapman and Hall, 1991.