The Metric Monad for Probabilistic Nondeterminism

Franck van Breugel [a,1]

[a] *York University, Department of Computer Science*
4700 Keele Street, Toronto, M3J 1P3, Canada

Abstract

A metric on the set of Borel probability measures on a metric space, introduced by Kantorovich in the early fourties, is shown to be *the* metric to model the computational effect of probabilistic nondeterminism. This metric gives rise to robust models, since small changes in the probabilities result in small changes in the distances.

Key words: probabilistic nondeterminism, monad, metric space, robust model

1 Introduction

It is well known that there is a close correspondence between *computational effects* and *monads* (see, for example, [6,44,47]). In this paper, we focus on monads related to the computational effect of *probabilistic nondeterminism*. Such a monad was introduced for the category $\mathbb{M}es$ of measurable spaces and measurable functions by Lawvere [40] [2] and Giry [29]. Jones and Plotkin [34,35] presented such a monad for the category $\mathbb{C}ont$ of domains (also known as continuous directed complete partial orders) and Scott continuous functions. To see these monads in action, we refer the reader to, for example, [35,50]. In this paper, we introduce such a monad for the category $\mathbb{CM}et_1$ of 1-bounded complete metric spaces and nonexpansive functions.

[1] Supported by Natural Sciences and Engineering Research Council of Canada.

[2] This paper dealt with the fact that the functor $\mathcal{P} : \mathbb{M}es \to \mathbb{M}es$ forms a monad, although of course Lawvere did not use the term monad then. In [40], the functor $\mathcal{P}$ is shown to be the composite of two adjoints, which gives a "standard construction." [42]

1.1 Why metrics?

As far as we know, Giacalone, Jou and Smolka [26] were the first to advocate *metric spaces* as a basis to model probabilistic nondeterminism. They consider a probabilistic variant on labelled transition systems. Their model induces distances on the states of the system. The distance between states, a real number between 0 and 1, is used to express the similarity of the behaviour of those states. The smaller the distance, the more the states behave alike. In particular, the distance between states is 0 if they are behaviourally indistinguishable.

The main advantage of such a metric model over behavioural equivalences like probabilistic bisimilarity [39] is its *robustness*. Behavioural equivalences are not robust, since they are too sensitive to the exact probabilities of the various transitions. Two states are either behaviourally equivalent or they are not. A slight change in the probabilities associated to the transitions may cause behaviourally equivalent states to become inequivalent and vice versa. However, in the proposed metric model of Giacalone et al. slight changes of the probabilities will only result in slight changes of the distances. For a more detailed discussion of the merits of such a metric model, we refer the reader to, for example, [26].

Recently, there has been a renewed interest in metric models a la Giacalone et al. Desharnais, Gupta, Jagadeesan and Panangaden [18–21,30] introduced metric analogues of strong and weak bisimilarity for probabilistic transition systems. Van Breugel and Worrell also presented a metric analogue of strong probabilistic bisimilarity in [13,14]. De Alfaro, Henzinger and Majumdar [15,16] and Deng, Chothia, Palamidessi and Pang [17] studied very similar metrics. Ying [56] introduced a continuous spectrum of behavioural equivalences which induces a metric model for probabilistic systems. Also Baier and Kwiatkowska [3], Den Hartog [31], Kwiatkowska and Norman [38] and De Vink and Rutten [54] exploited metrics to model probabilistic nondeterminism. However, they used their metrics mainly as a tool to model infinite behaviour, rather than to provide a robust model.

1.2 How to relate monads?

Let us compare the monad of Lawvere and Giry and the monad of Jones and Plotkin. The monad of Lawvere and Giry consists of a functor $\mathcal{P} : \mathbb{M}es \to \mathbb{M}es$ and two natural transformations. The functor $\mathcal{P}$ maps a measurable space to the set of probability measures on the space (provided with a suitable σ-field that we will define later). For now, we restrict the monad of Jones and Plotkin to the subcategory $\omega\mathbb{C}oh$ of ω-coherent domains and Scott continuous functions. This restricted monad consists of a functor $\mathcal{V} : \omega\mathbb{C}oh \to \omega\mathbb{C}oh$ and two natural transformations. The functor $\mathcal{V}$ is known as the *probabilistic powerdomain*. It maps each ω-coherent domain to the set of continuous valuations on the Scott topology of the domain (equipped with an appropriate order that we introduce later). As we will see, valuations bear a close resemblance to subprobability measures. Since each subprobability measure on a measurable space X can be viewed as a probability measure on $\mathbf{1} + X$, we consider the composition of the monads $\mathbf{1} + -$ and $\mathcal{P}$. Because there exists a distributive law [5] of $\mathbf{1} + -$ over $\mathcal{P}$, this composition is a monad as well. We denote this monad by $\mathcal{P}'$.

The forgetful functor $\mathcal{U} : \omega\mathbb{C}oh \to \mathbb{M}es$ provides the obvious way to mediate between ω-coherent domains and measurable spaces. This functor maps an ω-coherent domain to the Borel measurable space generated by the Scott topology of the domain. As Van Breugel, Mislove, Ouaknine and Worrell have already shown in [11, Proposition 19], the diagram

$$\begin{array}{ccc} \omega\mathbb{C}oh & \xrightarrow{\mathcal{V}} & \omega\mathbb{C}oh \\ {\scriptstyle \mathcal{U}}\downarrow & & \downarrow{\scriptstyle \mathcal{U}} \\ \mathbb{M}es & \xrightarrow[\mathcal{P}']{} & \mathbb{M}es \end{array}$$

commutes (when identifying Borel subprobability measures and continuous valuations). As we will see, also the two natural transformations, the unit and the multiplication, of both monads "coincide." The commutativity of the above diagram and the "coincidence" of the units and multiplications can be captured in categorical terms. The functor $\mathcal{U} : \omega\mathbb{C}oh \to \mathbb{M}es$ and the natural transformation $\mathrm{id} : \mathcal{P}'\mathcal{U} \stackrel{\cdot}{\to} \mathcal{U}\mathcal{V}$, which restricts a Borel subprobability measure on the Scott topology of an ω-coherent domain to its open sets, form a morphism in the category of monads [52]. This captures that the monad (with functor) $\mathcal{V}$ extends the monad (with functor) $\mathcal{P}'$.

The monad $\mathcal{V}$ can be characterized as the monad on $\omega\mathbb{C}oh$ that extends the monad $\mathcal{P}'$ with the following universal property. Consider another monad on $\omega\mathbb{C}oh$, say $\mathcal{F}$, that extends $\mathcal{P}'$. That is, the functor $\mathcal{U} : \omega\mathbb{C}oh \to \mathbb{M}es$ and the natural transformation $\mathrm{id} : \mathcal{P}'\mathcal{U} \stackrel{\cdot}{\to} \mathcal{U}\mathcal{F}$ also form a monad morphism from monad $\mathcal{F}$ to monad $\mathcal{P}'$. If $\mathcal{F}\mathbf{1}$ is isomorphic to $[0, 1]$ then the natural transformation $\mathrm{id} : \mathcal{F} \stackrel{\cdot}{\to} \mathcal{V}$ is a morphism in the category of monads on $\omega\mathbb{C}oh$ from monad $\mathcal{F}$ to monad $\mathcal{V}$.

1.3 Which metric?

One may wonder if the monad $\mathcal{P}'$ can also be extended to a monad on a category of metric spaces. For now, let us restrict our attention to the category $\mathbb{KM}et_1$ of 1-bounded compact metric spaces and nonexpansive functions. This time we use the forgetful functor $\mathcal{U} : \mathbb{KM}et_1 \to \mathbb{M}es$ to mediate between 1-bounded compact metric spaces and measurable spaces. This functor maps each 1-bounded compact metric space to the Borel measurable space generated by the ϵ-ball topology of the metric space. Now we can formalize our question as follows.

> Does there exist a monad $\mathcal{B}'$ on $\mathbb{KM}et_1$ that extends the monad $\mathcal{P}'$ with the same universal property as the monad $\mathcal{V}$?

If such a monad exists, then the diagram

$$\begin{array}{ccc} \mathbb{KM}et_1 & \xrightarrow{\mathcal{B}'} & \mathbb{KM}et_1 \\ {\scriptstyle \mathcal{U}}\downarrow & & \downarrow{\scriptstyle \mathcal{U}} \\ \mathbb{M}es & \xrightarrow[\mathcal{P}']{} & \mathbb{M}es \end{array}$$

should commute. As a consequence, the functor $\mathcal{B}'$ should map a 1-bounded compact metric space X to the set of Borel subprobability measures on X. These Borel subprobability measures on X can be seen as Borel probability measures on $\mathbf{1} + X$. As we will see, the monad $\mathcal{B}'$ is the composition of the monad $\mathbf{1}+-$ and a monad $\mathcal{B}$. The functor $\mathcal{B}$ maps a 1-bounded compact metric space X to the set of Borel probability measures on X. Many different metrics on the set of Borel probability measures have been proposed in the literature (see, for example, [49] for an overview). In this paper, we will consider a metric on Borel probability measures independently proposed by Dobrushin [23], Hutchinson [33], Kantorovich [37] and Vaserstein [53], and known under many different names including the Hutchinson metric, the Kantorovich metric, the Vaserstein metric and the Wasserstein metric (we will name the metric after Kantorovich who was the first, as far as we know, to discover it). We will show that the Kantorovich metric is *the* metric with the desired property. That is, the monad consisting of the functor $\mathcal{B} : \mathbb{KM}et_1 \to \mathbb{KM}et_1$, which maps each 1-bounded compact metric space to the set of Borel probability measures on the space endowed with the Kantorovich metric, and corresponding unit and multiplication is the monad on $\mathbb{KM}et_1$ such that the monad $\mathcal{B}'$ extends the monad $\mathcal{P}'$ with the same universal property as the monad $\mathcal{V}$. Therefore, one may view this monad as *the* metric monad of probabilistic nondeterminism.

All the metrics presented in [10,13–21,30] are either directly or indirectly based on the Kantorovich metric. Although numerous results in op. cit. already suggest that the Kantorovich metric is a suitable candidate for modelling probabilistic nondeterminism, our results show that it is *the* metric to model probabilistic nondeterminism. Furthermore, we establish a relationship between the Kantorovich functor and the probabilistic powerdomain.

1.4 The complete picture

So far, we have restricted our attention to monads on the categories $\omega\mathbb{C}oh$ and $\mathbb{KM}et_1$. Now let us discuss how these monads can be extended to monads on the categories $\mathbb{C}ont$ and $\mathbb{CM}et_1$. The monad $\mathcal{V}$ can be extended to a monad on $\mathbb{C}ont$ in the obvious way. We denote the extended monad also by $\mathcal{V}$. Since $\omega\mathbb{C}oh$ is a subcategory of $\mathbb{C}ont$ and the monad on $\mathbb{C}ont$ restricted to $\omega\mathbb{C}oh$ coincides with the monad on $\omega\mathbb{C}oh$, we have that the diagram

$$\begin{array}{ccc} \omega\mathbb{C}oh & \xrightarrow{\mathcal{V}} & \omega\mathbb{C}oh \\ \big\downarrow & & \big\downarrow \\ \mathbb{C}ont & \xrightarrow[\mathcal{V}]{} & \mathbb{C}ont \end{array}$$

commutes. One can also easily verify that the inclusion functor and the identity natural transformation form a monad morphism between the monads. A related monad morphism is considered in [35].

Also the monad $\mathcal{B}$ on $\mathbb{KM}et_1$ can be extended to a monad $\mathcal{B}$ on $\mathbb{CM}et_1$. To preserve completeness, we restrict ourselves to *tight* Borel probability measures. Since every measure on a compact metric space is tight (see, for example, [46, Section II.3]), the diagram

$$\begin{array}{ccc} \mathbb{KM}et_1 & \xrightarrow{\mathcal{B}} & \mathbb{KM}et_1 \\ \big\downarrow & & \big\downarrow \\ \mathbb{CM}et_1 & \xrightarrow[\mathcal{B}]{} & \mathbb{CM}et_1 \end{array}$$

commutes. Also in this case, the inclusion functor and the identity natural transformation form a monad morphism between the monads.

Note that the forgetful functors from $\mathbb{C}ont$ to $\mathbb{M}es$ and from $\mathbb{CM}et_1$ to $\mathbb{M}es$ do not give rise to monad morphisms.

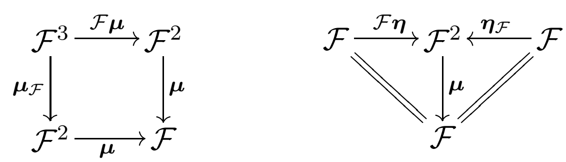
In [29], Giry introduced also a monad on the category $\mathbb{P}ol$ of Polish spaces and continuous functions (see, also [22]). The corresponding functor maps a Polish space to the set of Borel probability measures on the space equipped with the weak topology. For this monad, we can prove a characterization similar to the characterizations of $\mathcal{V}$ and $\mathcal{B}$ described above. In this paper, we will not study this monad in any detail as we focus on metric spaces.

2 Monads

We assume that the reader is familiar with the basics of category theory. Those basics can be found in, for example, [43]. Here we only review the notion of a monad, that will play a key role in our development, two categories of monads, and a distributive law that allows us to compose monads.

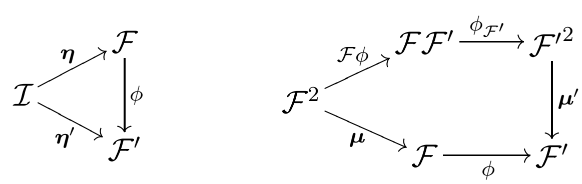
Below, we use $\mathcal{I}$ to denote the identity functor.

Definition 1 *A monad on a category* $\mathbb{C}$ *is a triple* $\langle \mathcal{F}, \boldsymbol{\eta}, \boldsymbol{\mu} \rangle$ *where* $\mathcal{F} : \mathbb{C} \to \mathbb{C}$ *is a functor, and the* unit $\boldsymbol{\eta} : \mathcal{I} \stackrel{.}{\to} \mathcal{F}$ *and the* multiplication $\boldsymbol{\mu} : \mathcal{F}^2 \stackrel{.}{\to} \mathcal{F}$ *are natural transformations which make the following diagrams commute.*



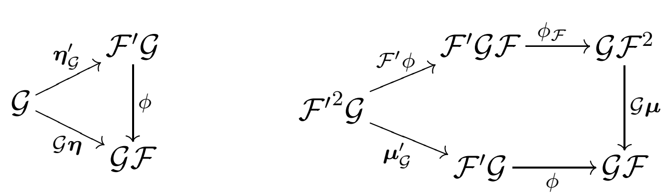
Given a category $\mathbb{C}$, the monads on $\mathbb{C}$ form a category. The morphisms between monads on $\mathbb{C}$ are defined as follows.

Definition 2 *A morphism from monad* $\langle \mathcal{F}, \boldsymbol{\eta}, \boldsymbol{\mu} \rangle$ *on category* $\mathbb{C}$ *to monad* $\langle \mathcal{F}', \boldsymbol{\eta}', \boldsymbol{\mu}' \rangle$ *on* $\mathbb{C}$ *consists of a natural transformation* $\phi : \mathcal{F} \stackrel{.}{\to} \mathcal{F}'$ *which makes the following diagrams commute.*



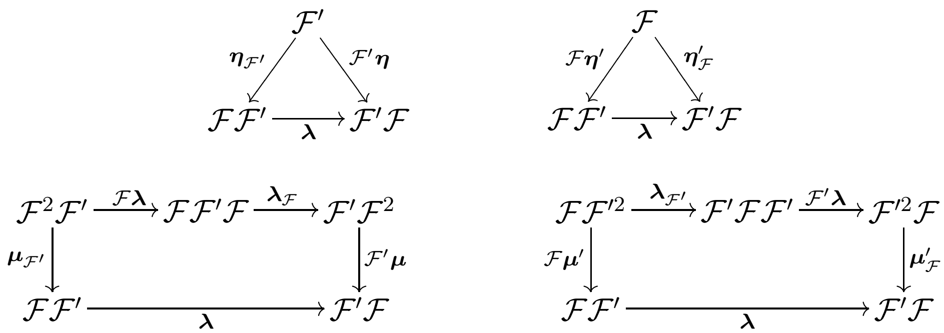
Also all monads form a category. In this case, the morphisms are defined as follows.

Definition 3 *A morphism from monad* $\langle \mathcal{F}, \boldsymbol{\eta}, \boldsymbol{\mu} \rangle$ *on category* $\mathbb{C}$ *to monad* $\langle \mathcal{F}', \boldsymbol{\eta}', \boldsymbol{\mu}' \rangle$ *on category* $\mathbb{C}'$ *consists of a tuple* $\langle \mathcal{G}, \phi \rangle$ *where* $\mathcal{G} : \mathbb{C} \to \mathbb{C}'$ *is a functor and* $\phi : \mathcal{F}'\mathcal{G} \stackrel{.}{\to} \mathcal{G}\mathcal{F}$ *is a natural transformation which make the following diagrams commute.*



Monads on a category $\mathbb{C}$ can be composed if there is a distributive law.

Definition 4 *A distributive law of a monad $\langle \mathcal{F}, \boldsymbol{\eta}, \boldsymbol{\mu} \rangle$ on category $\mathbb{C}$ over a monad $\langle \mathcal{F}', \boldsymbol{\eta}', \boldsymbol{\mu}' \rangle$ on $\mathbb{C}$ is a natural transformation $\boldsymbol{\lambda} : \mathcal{F}\mathcal{F}' \dot{\rightarrow} \mathcal{F}'\mathcal{F}$ which makes the following diagrams commute.*



Proposition 5 *Let $\langle \mathcal{F}, \boldsymbol{\eta}, \boldsymbol{\mu} \rangle$ and $\langle \mathcal{F}', \boldsymbol{\eta}', \boldsymbol{\mu}' \rangle$ be monads on a category $\mathbb{C}$. If $\boldsymbol{\lambda} : \mathcal{F}\mathcal{F}' \dot{\rightarrow} \mathcal{F}'\mathcal{F}$ is a distributive law, then $\langle \mathcal{F}'\mathcal{F}, (\mathcal{F}\boldsymbol{\eta}')\boldsymbol{\eta}, (\mathcal{F}'\boldsymbol{\mu})\boldsymbol{\mu}'_{\mathcal{F}^2}(\mathcal{F}'\boldsymbol{\lambda}_{\mathcal{F}}) \rangle$ is a monad.*

3 The monads $\mathcal{P}$ and $\mathcal{P}'$

In [29], Giry presented a monad $\mathcal{P}$ on the category $\mathbb{M}es$ of measurable spaces and measurable functions. This monad was also studied by Lawvere in [40]. We assume that the reader is familiar with basic notions of probability theory as can be found in, for example, [7]. The functor $\mathcal{P} : \mathbb{M}es \rightarrow \mathbb{M}es$ maps a measurable space to the set of probability measures on the space. To equip this set with a σ-field we introduce the following evaluation functions.

Definition 6 *Let X be a measurable space. Let B be a measurable subset of X. The function $\varepsilon_B : \mathcal{P}X \rightarrow [0, 1]$ is defined by*

$$\varepsilon_B(\mu) = \mu(B)$$

for μ a probability measure on X.

We turn the set of probability measures into a measurable space by giving it the smallest σ-field Σ such that ε_B is Σ-measurable for each measurable subset B of the space. The functor $\mathcal{P}$ acts on morphisms as follows.

Definition 7 *Let X and Y be measurable spaces. Let $f : X \rightarrow Y$ be a measurable function. The measurable function $\mathcal{P}f : \mathcal{P}X \rightarrow \mathcal{P}Y$ is defined by*

$$(\mathcal{P}f)(\mu)(B) = \mu(f^{-1}(B))$$

for μ a probability measure on X and B a measurable subset of Y.

Next, we present the unit $\boldsymbol{\eta}$ and the multiplication $\boldsymbol{\mu}$ of the monad $\mathcal{P}$.

Definition 8 *Let X be a measurable space. The measurable function $\boldsymbol{\eta}_X : X \to \mathcal{P}X$ is defined by*

$$\boldsymbol{\eta}_X(x)(B) = \begin{cases} 1 \text{ if } x \in B \\ 0 \text{ otherwise} \end{cases}$$

for $x \in X$ and B a measurable subset of X. The measurable function $\boldsymbol{\mu}_X : \mathcal{P}^2 X \to \mathcal{P}X$ is defined by

$$\boldsymbol{\mu}_X(\mu)(B) = \int_{\mathcal{P}X} \varepsilon_B \, d\mu$$

for μ a probability measure on $\mathcal{P}X$ and B a measurable subset of X.

Note that $\boldsymbol{\eta}_X(x)$ is the Dirac measure at x. For a proof that $\langle \mathcal{P}, \boldsymbol{\eta}, \boldsymbol{\mu} \rangle$ is indeed a monad we refer the reader to [29, Theorem 1].

Proposition 9 *Let X be a measurable space. Let $f : X \to [0,1]$ be a measurable function. For each probability measure μ on $\mathcal{P}X$,*

$$\int_X f \, d\boldsymbol{\mu}_X(\mu) = \int_{\mathcal{P}X} i_f \, d\mu$$

where the measurable function $i_f : \mathcal{P}X \to [0,1]$ is defined by

$$i_f(\nu) = \int_X f \, d\nu$$

for ν a probability measure on X.

PROOF. See, for example, the proof of [29, Theorem 1]. □

Next, we present the monad $\mathcal{P}'$. The functor $\mathcal{P}' : \mathbb{M}es \to \mathbb{M}es$ can be viewed as mapping a measurable space to the set of subprobability measures on the space. To define this functor, we exploit the terminal object and the coproduct in $\mathbb{M}es$. The category $\mathbb{M}es$ has a terminal object $\mathbf{1}$ consisting of a singleton set, whose single element we denote by $\mathbf{0}$, and the unique σ-field on this singleton set. Given measurable spaces $\langle X, \Sigma_X \rangle$ and $\langle Y, \Sigma_Y \rangle$, the coproduct $\langle X, \Sigma_X \rangle + \langle Y, \Sigma_Y \rangle$ consists of the disjoint union of the sets X and Y and the σ-field generated by the disjoint union of the σ-fields Σ_X and Σ_Y. Since $\mathbb{M}es$ has a terminal object and a coproduct, we have the functor $\mathbf{1} + - : \mathbb{M}es \to \mathbb{M}es$. This functor can be extended in a straightforward way to a monad (see, for example, [4, Section 3.1]). To combine the monads $\mathcal{P}$ and $\mathbf{1} + -$, we exploit a distributive law.

Definition 10 *Let X be a measurable space. The measurable function $\boldsymbol{\lambda}_X : \mathbf{1} + \mathcal{P}X \to \mathcal{P}(\mathbf{1} + X)$ is defined by*

$$\boldsymbol{\lambda}_X(\mathbf{0})(B) = \begin{cases} 1 \text{ if } \mathbf{0} \in B \\ 0 \text{ otherwise} \end{cases}$$
$$\boldsymbol{\lambda}_X(\mu)(B) = \mu(B \cap X)$$

for μ a probability measure on X and B a measurable subset of $\mathbf{1} + X$.

Note that $\boldsymbol{\lambda}_X(\mathbf{0}) = \boldsymbol{\eta}_{\mathbf{1}+X}(\mathbf{0})$ and $\boldsymbol{\lambda}_X(\mu) = \mathcal{P}(\iota_X)(\mu)$, where ι_X is the injection from X to $\mathbf{1} + X$.

For the natural transformation $\boldsymbol{\lambda}$, we can prove the following result.

Proposition 11 *$\boldsymbol{\lambda}$ is a distributive law of $\mathbf{1} + -$ over $\mathcal{P}$.*

PROOF. We have to prove that the four diagrams of Definition 4 commute. We only consider the last one. Proving the commutativity of the other diagrams is fairly straightforward and left to the reader. Let X be a measurable space. Then, the last diagram amounts to following.

$$\begin{array}{ccccc} \mathbf{1} + \mathcal{P}^2 X & \xrightarrow{\boldsymbol{\lambda}_{\mathcal{P}X}} & \mathcal{P}(\mathbf{1} + \mathcal{P}X) & \xrightarrow{\mathcal{P}\boldsymbol{\lambda}_X} & \mathcal{P}^2(\mathbf{1} + X) \\ {\scriptstyle \mathbf{1}+\boldsymbol{\mu}_X}\downarrow & & & & \downarrow{\scriptstyle \boldsymbol{\mu}_{\mathbf{1}+X}} \\ \mathbf{1} + \mathcal{P}X & & \xrightarrow[\boldsymbol{\lambda}_X]{} & & \mathcal{P}(\mathbf{1} + X) \end{array}$$

where $\boldsymbol{\mu}$ is the multiplication of the monad $\mathcal{P}$. Let B be a measurable subset of $\mathbf{1} + X$. We distinguish two cases. First of all,

$$\begin{aligned} & \boldsymbol{\mu}_{\mathbf{1}+X}(\mathcal{P}(\boldsymbol{\lambda}_X)(\boldsymbol{\lambda}_{\mathcal{P}X}(\mathbf{0})))(B) \\ = & \int_{\mathcal{P}(\mathbf{1}+X)} \varepsilon_B \, d(\mathcal{P}(\boldsymbol{\lambda}_X)(\boldsymbol{\lambda}_{\mathcal{P}X}(\mathbf{0}))) \\ = & \int_{\mathbf{1}+\mathcal{P}X} (\varepsilon_B \circ \boldsymbol{\lambda}_X) \, d(\boldsymbol{\lambda}_{\mathcal{P}X}(\mathbf{0})) \end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbf{1}+\mathcal{P}X} (\varepsilon_B \circ \boldsymbol{\lambda}_X)\, d(\boldsymbol{\eta}_{\mathbf{1}+\mathcal{P}X}(\mathbf{0})) \\
&= \varepsilon_B(\boldsymbol{\lambda}_X(\mathbf{0})) \\
&= \boldsymbol{\lambda}_X(\mathbf{0})(B) \\
&= \boldsymbol{\lambda}_X((\mathbf{1}+\boldsymbol{\mu}_X)(\mathbf{0}))(B).
\end{aligned}
$$

Secondly, let μ be a probability measure on $\mathcal{P}X$. Since for each probability measure ν on X,

$$
\begin{aligned}
&\varepsilon_B(\boldsymbol{\lambda}_X(\iota_{\mathcal{P}X}(\nu))) \\
&= \boldsymbol{\lambda}_X(\iota_{\mathcal{P}X}(\nu))(B) \\
&= \iota_{\mathcal{P}X}(\nu)(B \cap X) \\
&= \nu(B \cap X) \\
&= \varepsilon_{B\cap X}(\nu),
\end{aligned}
$$

we have that

$$
\begin{aligned}
&\boldsymbol{\mu}_{\mathbf{1}+X}(\mathcal{P}(\boldsymbol{\lambda}_X)(\boldsymbol{\lambda}_{\mathcal{P}X}(\mu)))(B) \\
&= \int_{\mathcal{P}(\mathbf{1}+X)} \varepsilon_B\, d(\mathcal{P}(\boldsymbol{\lambda}_X)(\boldsymbol{\lambda}_{\mathcal{P}X}(\mu))) \\
&= \int_{\mathbf{1}+\mathcal{P}X} (\varepsilon_B \circ \boldsymbol{\lambda}_X)\, d(\boldsymbol{\lambda}_{\mathcal{P}X}(\mu)) \\
&= \int_{\mathbf{1}+\mathcal{P}X} (\varepsilon_B \circ \boldsymbol{\lambda}_X)\, d(\mathcal{P}(\iota_{\mathcal{P}X})(\mu)) \\
&= \int_{\mathcal{P}X} (\varepsilon_B \circ \boldsymbol{\lambda}_X \circ \iota_{\mathcal{P}X})\, d\mu \\
&= \int_{\mathcal{P}X} \varepsilon_{B\cap X}\, d\mu \\
&= \boldsymbol{\mu}_X(\mu)(B \cap X) \\
&= \boldsymbol{\lambda}_X(\boldsymbol{\mu}_X(\mu))(B) \\
&= \boldsymbol{\lambda}_X((\mathbf{1}+\boldsymbol{\mu}_X)(\mu))(B).
\end{aligned}
$$

□

From Proposition 5 we can conclude that the functor $\mathcal{P}(\mathbf{1}+-)$ forms a monad as well. We denote this monad by $\mathcal{P}'$. The functor $\mathcal{P}'$ maps a measurable space X to the set of probability measures on $\mathbf{1}+X$. These probability measures on $\mathbf{1}+X$ can be viewed as subprobability measures on X.

4 The monad $\mathcal{V}$

In [51], Saheb-Djahromi introduced a functor on the category of domains and Scott continuous functions, mapping a domain X to the Borel probability measures on the Scott topology of X. We assume that the reader is familiar with the basics of domain theory. For more details on domain theory, we refer the reader to, for example, [28]. Instead of Borel probability measures, Jones and Plotkin [34,35] considered continuous valuations. Recall that a continuous valuation on a topological space X is a function $\nu : \mathcal{O}_X \to [0, 1]$, where $\mathcal{O}_X$ denotes the set of open sets of the topology X, satisfying

- $\nu(\emptyset) = 0$,
- if $U \subseteq V$ then $\nu(U) \leq \nu(V)$,
- $\nu(U) + \nu(V) = \nu(U \cup V) + \nu(U \cap V)$, and
- for every directed (with respect to the inclusion relation) subset $\mathcal{U}$ of $\mathcal{O}_X$, $\nu(\bigcup \mathcal{U}) = \sup_{U \in \mathcal{U}} \nu(U)$.

Jones and Plotkin introduced a monad $\mathcal{V}$ on the category $\mathbb{C}ont$ of domains and Scott continuous functions. Recall that a domain is a directed complete partial order which is continuous, that is, which has a basis. The functor $\mathcal{V} : \mathbb{C}ont \to \mathbb{C}ont$ maps a continuous domain X to the set of continuous valuations on (the Scott topology of) X with the partial order $\sqsubseteq_{\mathcal{V}X}$ defined by

$$\mu \sqsubseteq_{\mathcal{V}X} \nu \text{ if for all Scott open subsets } U \text{ of } X,\ \mu(U) \leq \nu(U)$$

for μ and ν continuous valuations on X. This functor acts on morphisms similar to the functor $\mathcal{P}$. Let X and Y be domains and let $f : X \to Y$ be a Scott continuous function. The Scott continuous function $\mathcal{V}f : \mathcal{V}X \to \mathcal{V}Y$ is defined by

$$(\mathcal{V}f)(\nu)(U) = \nu(f^{-1}(U))$$

for ν a continuous valuation on X and U a Scott open subset of X. Also the unit $\boldsymbol{\eta}$ and the multiplication $\boldsymbol{\mu}$ of the monad $\mathcal{V}$ are defined analogous to those of the monad $\mathcal{P}$.

Jones [34, Corollary 5.5] has proved that the functor $\mathcal{V}$ preserves ω-continuity: if the domain X is ω-continuous, that is, it has a countable basis, then $\mathcal{V}X$ is ω-continuous as well. Furthermore, Jung and Tix [36, Theorem 4.2] have shown that $\mathcal{V}$ preserves Lawson compactness: if the domain X is Lawson compact, that is, its Lawson topology is compact, then $\mathcal{V}X$ is Lawson compact as well. Hence, we can restrict the functor $\mathcal{V} : \mathbb{C}ont \to \mathbb{C}ont$ to the category $\omega\mathbb{C}oh$ of ω-coherent domains, that is, ω-continuous domains whose Lawson topology is compact, and Scott continuous functions. In this way we obtain a functor $\mathcal{V} : \omega\mathbb{C}oh \to \omega\mathbb{C}oh$ and a monad $\mathcal{V}$ on $\omega\mathbb{C}oh$.

Valuations bear a close resemblance to Borel subprobability measures. In fact, any valuation on a domain can be uniquely extended to a measure on the Borel σ-field generated by the Scott topology of the domain [1, Corollary 4.3]. Conversely, any Borel subprobability measure on the Scott topology of an ω-continuous domain defines a continuous valuation when restricted to the open sets [1, Lemma 2.5]. These results allow us to relate the monad $\mathcal{V}$ on $\omega\mathbb{C}oh$ to the monad $\mathcal{P}'$ on $\mathbb{M}es$ in the following way.

Proposition 12 *The forgetful functor $\mathcal{U} : \omega\mathbb{C}oh \to \mathbb{M}es$ and the natural transformation* $\mathrm{id} : \mathcal{P}'\mathcal{U} \dot{\to} \mathcal{U}\mathcal{V}$ *form a monad morphism from the monad $\mathcal{V}$ to the monad $\mathcal{P}'$ and $\varepsilon_{\{\mathbf{0}\}} : \mathcal{V}\mathbf{1} \to [0,1]$ is an isomorphism.*

Furthermore, we have the following universal property.

Proposition 13 *Let $\mathcal{F}$ be a monad on $\omega\mathbb{C}oh$ such that the forgetful functor $\mathcal{U} : \omega\mathbb{C}oh \to \mathbb{M}es$ and the natural transformation* $\mathrm{id} : \mathcal{P}'\mathcal{U} \dot{\to} \mathcal{U}\mathcal{F}$ *form a monad morphism from the monad $\mathcal{F}$ to the monad $\mathcal{P}'$. If $\varepsilon_{\{\mathbf{0}\}} : \mathcal{F}\mathbf{1} \to [0,1]$ is an isomorphism then the natural transformation* $\mathrm{id} : \mathcal{F} \dot{\to} \mathcal{V}$ *forms a monad morphism from the monad $\mathcal{F}$ to the monad $\mathcal{V}$.*

Since our focus is on metric spaces, we refrain from presenting proofs of Proposition 12 and 13 and we refer the reader to the very similar proofs of their metric counterparts Proposition 18 and 23. We will discuss the condition that $\varepsilon_{\{\mathbf{0}\}} : \mathcal{F}\mathbf{1} \to [0,1]$ is an isomorphism in Section 6.

5 The monads $\mathcal{B}$ and $\mathcal{B}'$

In [37], Kantorovich introduced a metric on the set of Borel probability measures on a metric space. Below, we present the Kantorovich metric and the corresponding monads on $\mathbb{KM}et_1$ and $\mathbb{CM}et_1$. We assume that the reader is familiar with basic metric topology. Those basics can be found in, for example, [25]. Given a 1-bounded metric space X, we use $\mathcal{B}X$ to denote the set of Borel probability measures on X and $X \underset{1}{\to} [0,1]$ to denote the set of nonexpansive functions from X to $[0,1]$. Recall that a metric space X is 1-bounded if $d_X(x,y) \leq 1$ for all x, $y \in X$, and a function $f : X \to Y$ on metric spaces is nonexpansive if $d_Y(f(x), f(y)) \leq d_X(x,y)$ for all x, $y \in X$. To turn the set $\mathcal{B}X$ into a metric space we introduce the following distance function.

Definition 14 *Let X be a 1-bounded metric space. The distance function*

$d_{\mathcal{B}X} : \mathcal{B}X \times \mathcal{B}X \to [0,1]$ is defined by

$$d_{\mathcal{B}X}(\mu,\nu) = \sup \left\{ \left| \int_X f d\mu - \int_X f d\nu \right| : f \in X \underset{1}{\to} [0,1] \right\}$$

for μ and ν Borel probability measures on X.

For a proof that the distance function $d_{\mathcal{B}X}$ is a metric we refer the reader to, for example, [24, Proposition 2.5.14]. As shown in, for example, [46, Theorem II.6.4], if the metric space X is compact, then the metric space $\mathcal{B}X$ is compact as well. For a proof that $\mathcal{B}$ preserves completeness if we restrict ourselves to tight measures, we refer the reader to, for example, [24, Theorem 2.5.25].

5.1 Compact metric spaces

The functor $\mathcal{B} : \mathbb{KM}et_1 \to \mathbb{KM}et_1$ on the category $\mathbb{KM}et_1$ of 1-bounded compact metric spaces and nonexpansive functions maps each 1-bounded compact metric space to the set of Borel probability measures on the space endowed with the Kantorovich metric. This functor acts the same way on morphisms as the functor $\mathcal{P}$. In [14, Proposition 16], it is shown that if the function f is nonexpansive then the function $\mathcal{B}f$ is nonexpansive as well.

The unit $\boldsymbol{\eta}$ and the multiplication $\boldsymbol{\mu}$ of the monad $\mathcal{B}$ on $\mathbb{KM}et_1$ are defined analogous to those of the monad $\mathcal{P}$. For example, the component of the unit $\boldsymbol{\eta}$ for the 1-bounded compact metric space X is defined by

$$\boldsymbol{\eta}_X(x)(B) = \begin{cases} 1 \text{ if } x \in B \\ 0 \text{ otherwise} \end{cases}$$

for $x \in X$ and B a Borel subset of X. As is shown in, for example, [46, Lemma II.6.1], $\boldsymbol{\eta}_X$ is an isometric embedding of X into $\mathcal{B}X$. Hence, each component of the unit is nonexpansive. Also the components of the multiplication are nonexpansive.

Proposition 15 *For each 1-bounded compact metric space X, the function $\boldsymbol{\mu}_X$ is nonexpansive.*

PROOF. Let $g : X \to [0,1]$ be a nonexpansive function. First, we show that the function i_g, introduced in Proposition 9, is nonexpansive. Let μ and ν be Borel probability measures on X. Then

$$
\begin{aligned}
&|i_g(\mu) - i_g(\nu)| \\
&= \left| \int_X g \, d\mu - \int_X g \, d\nu \right| \\
&\leq \sup \left\{ \left| \int_X f \, d\mu - \int_X f \, d\nu \right| : f \in X \underset{1}{\rightarrow} [0,1] \right\} \quad [g \text{ is nonexpansive}] \\
&= d_{\mathcal{B}X}(\mu, \nu).
\end{aligned}
$$

As a consequence, for μ and ν Borel probability measures on $\mathcal{B}X$,

$$
\begin{aligned}
&d_{\mathcal{B}X}\left(\boldsymbol{\mu}_X(\mu), \boldsymbol{\mu}_X(\nu)\right) \\
&= \sup \left\{ \left| \int_X g \, d\boldsymbol{\mu}_X(\mu) - \int_X g \, d\boldsymbol{\mu}_X(\nu) \right| : g \in X \underset{1}{\rightarrow} [0,1] \right\} \\
&= \sup \left\{ \left| \int_{\mathcal{B}X} i_g \, d\mu - \int_{\mathcal{B}X} i_g \, d\nu \right| : g \in X \underset{1}{\rightarrow} [0,1] \right\} \quad [\text{Proposition 9}] \\
&\leq \sup \left\{ \left| \int_{\mathcal{B}X} f \, d\mu - \int_{\mathcal{B}X} f \, d\nu \right| : f \in \mathcal{B}X \underset{1}{\rightarrow} [0,1] \right\} \quad [\text{each } i_g \text{ is nonexpansive}] \\
&= d_{\mathcal{B}^2X}(\mu, \nu).
\end{aligned}
$$

Therefore, the function $\boldsymbol{\mu}_X$ is nonexpansive as well. $\square$

Hence, $\langle \mathcal{B}, \boldsymbol{\eta}, \boldsymbol{\mu} \rangle$ forms a monad on $\mathbb{KM}et_1$.

5.2 Complete metric spaces

When we consider the category $\mathbb{CM}et_1$ of 1-bounded complete metric spaces and nonexpansive functions, we restrict ourselves to tight measures. That is, the functor $\mathcal{B} : \mathbb{CM}et_1 \to \mathbb{CM}et_1$ maps each 1-bounded complete metric space to the set of tight Borel probability measures on the space endowed with the Kantorovich metric. Recall that a Borel probability measure μ on a metric space X is tight if for every $\epsilon > 0$ there exists a compact subset K of X such that $\mu(K) > 1 - \epsilon$.

Since the Dirac measures are obviously tight, each component of the unit maps an element of the space to a tight Borel probability measure on the space. Also the components of the multiplication give rise to tight measures.

Proposition 16 *For each 1-bounded complete metric space X and tight Borel* *probability measure μ on the tight Borel probability measures on X, the Borel probability measure $\boldsymbol{\mu}_X(\mu)$ is tight.*

PROOF. Let $\epsilon > 0$. By the definition of tightness, it suffices to show that there a compact subset K of X such that $\boldsymbol{\mu}_X(\mu)(K) > 1 - \epsilon$. Since the measure μ is tight, by definition there exists a compact subset K_μ of $\mathcal{B}X$ such that

$$\mu(K_\mu) > 1 - \tfrac{\epsilon}{2}. \tag{1}$$

According to Prohorov's theorem (see, for example, [8, Section 1.5]), a compact set of tight measures on a complete space is uniformly tight. Hence, there exists a compact subset K of X such that

$$\text{for all } \nu \in K_\mu, \nu(K) > 1 - \tfrac{\epsilon}{2}. \tag{2}$$

Therefore,

$$\begin{aligned}
\boldsymbol{\mu}_X(\mu)(K) &= \int_{\mathcal{B}X} \varepsilon_K \, d\mu \\
&\geq \int_{K_\mu} \varepsilon_K \, d\mu \\
&> \int_{K_\mu} (1 - \tfrac{\epsilon}{2}) \, d\mu \quad [(2)] \\
&= \int_{\mathcal{B}X} (1 - \tfrac{\epsilon}{2}) \, d\mu - \int_{\mathcal{B}X \setminus K_\mu} (1 - \tfrac{\epsilon}{2}) \, d\mu \\
&\geq 1 - \tfrac{\epsilon}{2} - \int_{\mathcal{B}X \setminus K_\mu} 1 \, d\mu \\
&\geq 1 - \tfrac{\epsilon}{2} - \tfrac{\epsilon}{2} \quad [(1)]
\end{aligned}$$

□

Hence, also the functor $\mathcal{B}$ on $\mathbb{CM}et_1$ gives rise to a monad.

5.3 Subprobability measures

We can extend the monad $\mathcal{B}$ to a monad $\mathcal{B}'$ of subprobability measures in the same way we extended $\mathcal{P}$ to $\mathcal{P}'$. In the categories $\mathbb{KM}et_1$ and $\mathbb{CM}et_1$, the terminal object $\mathbf{1}$ is a singleton set. Given metric spaces $\langle X, d_X\rangle$ and $\langle Y, d_Y\rangle$, the coproduct $\langle X, d_X\rangle + \langle Y, d_Y\rangle$ consists of the disjoint union of the sets X and Y endowed with the metric d_{X+Y} defined by

$$d_{X+Y}(v,w) = \begin{cases} d_X(v,w) & \text{if } v,\, w \in X \\ d_Y(v,w) & \text{if } v,\, w \in Y \\ 1 & \text{otherwise.} \end{cases}$$

To compose the monads $\mathcal{B}$ and $\mathbf{1} + -$, we introduce a distributive law similar to the one introduced in Definition 10.

Proposition 17 *For each 1-bounded complete metric space X, the function $\boldsymbol{\lambda}_X$ is nonexpansive.*

PROOF. It suffices to show that for all tight Borel probability measures μ and ν on X,

$$d_{\mathcal{B}(\mathbf{1}+X)}(\boldsymbol{\lambda}_X(\mu), \boldsymbol{\lambda}_X(\nu)) \leq d_{\mathcal{B}X}(\mu,\nu).$$

Let μ and ν be tight Borel probability measures on X. Then

$$\begin{aligned}
& d_{\mathcal{B}(\mathbf{1}+X)}(\boldsymbol{\lambda}_X(\mu), \boldsymbol{\lambda}_X(\nu)) \\
&= \sup\left\{ \left| \int_{\mathbf{1}+X} f\, d(\boldsymbol{\lambda}_X(\mu)) - \int_{\mathbf{1}+X} f\, d(\boldsymbol{\lambda}_X(\nu)) \right| : f \in \mathbf{1}+X \xrightarrow[1]{} [0,1] \right\} \\
&= \sup\left\{ \left| \int_{\mathbf{1}+X} f\, d(\mathcal{P}(\iota_X)(\mu)) - \int_{\mathbf{1}+X} f\, d(\mathcal{P}(\iota_X)(\nu)) \right| : f \in \mathbf{1}+X \xrightarrow[1]{} [0,1] \right\} \\
&= \sup\left\{ \left| \int_X (f \circ \iota_X)\, d\mu - \int_X (f \circ \iota_X)\, d\nu \right| : f \in \mathbf{1}+X \xrightarrow[1]{} [0,1] \right\} \\
&\leq \sup\left\{ \left| \int_X g\, d\mu - \int_X g\, d\nu \right| : g \in X \xrightarrow[1]{} [0,1] \right\} \quad [f \circ \iota_X \text{ is nonexpansive}] \\
&= d_{\mathcal{B}X}(\mu,\nu).
\end{aligned}$$

□

The rest of the proof that $\boldsymbol{\lambda}$ is a distributive law is very similar to the proof of Proposition 11. We denote the composition of the monads $\mathbf{1} + -$ and $\mathcal{B}$ by

$\mathcal{B}'$. The functor $\mathcal{B}'$ can be viewed as mapping a space X to the set of Borel subprobability measures on X.

6 Relating $\mathcal{P}$ and $\mathcal{B}$

First, we will show that the monad $\mathcal{B}$ on $\mathbb{KM}et_1$ extends the monad $\mathcal{P}$ on $\mathbb{M}es$. That is, we will prove that the forgetful functor $\mathcal{U} : \mathbb{KM}et_1 \to \mathbb{M}es$ and the natural transformation id : $\mathcal{PU} \dot{\to} \mathcal{UB}$ form a monad morphism from the monad $\mathcal{B}$ on $\mathbb{KM}et_1$ to the monad $\mathcal{P}$ on $\mathbb{M}es$. In [12, Lemma 1], Van Breugel, Shalit and Worrell already sketched a proof of

Proposition 18 $\mathcal{PU} = \mathcal{UB}$.

Here, we provide some more details. Let X be a 1-bounded compact metric space. Clearly, both $\mathcal{PU}X$ and $\mathcal{UB}X$ are the set of Borel probability measures on X equipped with a σ-field. Next, we prove that their σ-fields, denoted Σ_B and Σ_K below, coincide. According to, for example, [25, Theorem 4.1.15], a compact metrizable space is second-countable. That is, the topology induced by X has a countable basis, say $\mathcal{A}$. Without loss of generality, we may assume that $\mathcal{A}$ is closed under finite intersections. The set of Borel probability measures on X can be provided with a σ-field in the following ways.

- Σ_B is the smallest σ-field Σ such that ε_B is Σ-measurable for each Borel subset B of X.
- Σ_O is the smallest σ-field Σ such that ε_A is Σ-measurable for each open subset A in $\mathcal{A}$.
- Σ_K is the Borel σ-field of the metric space $\mathcal{B}X$.

Next, we show that these three σ-fields coincide. The proof of this result is split into following parts.

Proposition 19 $\Sigma_B \subseteq \Sigma_O$.

PROOF. By minimality of Σ_B, it suffices to show that ε_B is Σ_O-measurable for each Borel subset B of X.

Consider

$$\mathcal{L} = \{\, B \text{ is a Borel subset of } X : \varepsilon_B \text{ is } \Sigma_O\text{-measurable}\,\}.$$

We only need to prove that each Borel subset of X is an element of the set $\mathcal{L}$.

One can easily verify that $\emptyset \in \mathcal{L}$ and that $\mathcal{L}$ is closed under complement, and finite and countable disjoint unions. Hence, $\mathcal{L}$ is a λ-system. Since the collection $\mathcal{A}$ is closed under finite intersections, $\mathcal{A}$ is a π-system. By the definition of Σ_O, we have that $\mathcal{A} \subseteq \mathcal{L}$. According to the $\lambda - \pi$ theorem (see, for example, [7, Theorem 3.2]), if a π-system is a subset of a λ-system then the σ-field generated by the π-system is also a subset of the λ-system. Hence, the σ-field generated by $\mathcal{A}$ is a subset of $\mathcal{L}$.

Since $\mathcal{A}$ is a countable basis of the topology induced by X, the σ-field generated by $\mathcal{A}$ equals the Borel σ-field of X according to [46, Theorem I.1.8]. Hence, each Borel subset of X is an element of $\mathcal{L}$. $\square$

Proposition 20 $\Sigma_O \subseteq \Sigma_B$.

PROOF. Since each open subset A in $\mathcal{A}$ is a Borel subset of X, we can conclude that $\Sigma_O \subseteq \Sigma_B$ by minimality of Σ_O. $\square$

Proposition 21 $\Sigma_O \subseteq \Sigma_K$.

PROOF. By minimality of Σ_O, it suffices to prove that ε_A is Σ_K-measurable for each open subset A in $\mathcal{A}$.

Let A be an open subset in $\mathcal{A}$. To show that a function is measurable, it suffices to prove that the inverse images of those sets that generate the σ-field (rather than all sets of the σ-field) are measurable sets. Since the Borel σ-field of $[0, 1]$ is generated by the sets $(q, 1]$ where q is a rational in $[0, 1]$, we have left to prove that

$$\varepsilon_A^{-1}(q, 1] = \{\, \mu \in \mathcal{B}X : \mu(A) > q \,\}$$

is in Σ_K.

Let q be a rational in $[0, 1]$. Since the Kantorovich metric metrizes the weak topology, as is proved in, for example, [46, Theorem 6.2], it suffices to show that the set $\varepsilon_A^{-1}(q, 1]$ is open in the weak topology. We conclude this proof by showing that the complement of $\varepsilon_A^{-1}(q, 1]$ is closed in the weak topology.

According to, for example, [25, Corollary 1.6.4], a set is closed if and only if the set contains all limits of each net in the set. Let $\{\mu_\alpha\}$ be a net consisting of measures not in $\varepsilon_A^{-1}(q, 1]$ with limit μ. We have left to show that μ is not in $\varepsilon_A^{-1}(q, 1]$ either, that is, $\mu(A) \leq q$. This follows immediately from the fact that the net $\{\mu_\alpha\}$ converges to μ in the weak topology if and only if $\liminf_\alpha \mu_\alpha(B) \geq \mu(B)$ for all open subsets B of X as shown in [46, Theorem II.6.1]. $\square$

Proposition 22 $\Sigma_O = \Sigma_K$.

PROOF. According to the unique structure theorem (see, for example, [2, Theorem 3.3.5]), on a compact metric space any countably generated σ-field of Borel sets which separates points is in fact equal to the whole Borel σ-field.

As we already mentioned, $\mathcal{B}$ preserves compactness. Since X is compact, the metric space $\mathcal{B}X$ is compact as well. Σ_K is the Borel σ-field of this compact metric space. As we have shown in the proof of Proposition 21, Σ_O is a subfield of Σ_K. Therefore, we have left to prove that Σ_O is countably generated and separates points.

By definition, Σ_O is the smallest σ-field Σ such that ε_A is Σ-measurable for all $A \in \mathcal{A}$. As we have seen in the proof of Proposition 21 ε_A is Σ-measurable if $\varepsilon_A^{-1}(q, 1] \in \Sigma$ for each rational q in $[0, 1]$. Therefore, Σ_O is generated by the countable collection of sets $\varepsilon_A^{-1}(q, 1]$ where $A \in \mathcal{A}$ and q is a rational in $[0, 1]$.

Finally, we show that Σ_O separates points. Let μ and ν be different measures. Without loss of generality we may assume that $\mu(B) \leq q < \nu(B)$ for some Borel set B and rational q. Clearly, $\mu \notin \varepsilon_B^{-1}(q, 1]$ and $\nu \in \varepsilon_B^{-1}(q, 1]$. Hence, $\varepsilon_B^{-1}(q, 1]$ separates μ and ν. $\square$

We can easily check that $\mathrm{id} : \mathcal{PU} \dot{\rightarrow} \mathcal{UB}$ is a natural transformation. Furthermore, we can easily verify that the diagrams of Definition 3 commute for the forgetful functor $\mathcal{U} : \mathbb{KM}et_1 \rightarrow \mathbb{M}es$ and the natural transformation $\mathrm{id} : \mathcal{PU} \dot{\rightarrow} \mathcal{UB}$. Hence, $\mathcal{U}$ and id form a monad morphism from the monad $\mathcal{B}$ on $\mathbb{KM}et_1$ to the monad $\mathcal{P}$ on $\mathbb{M}es$. Note that $\varepsilon_{\{\mathbf{0}\}} : \mathcal{B}\mathbf{2} \rightarrow [0, 1]$ is an isomorphism.

Second, we will show that the monad $\mathcal{B}$ can be characterized as the monad on $\mathbb{KM}et_1$ that extends the monad $\mathcal{P}$ with the following universal property.

Proposition 23 *Let $\mathcal{F}$ be a monad on $\mathbb{KM}et_1$ such that the forgetful functor $\mathcal{U} : \mathbb{KM}et_1 \rightarrow \mathbb{M}es$ and the natural transformation* $\mathrm{id} : \mathcal{PU} \dot{\rightarrow} \mathcal{UF}$ *form a monad morphism from the monad $\mathcal{F}$ to the monad $\mathcal{P}$. If $\varepsilon_{\{\mathbf{0}\}} : \mathcal{F}\mathbf{2} \rightarrow [0, 1]$ is an isomorphism then the natural transformation* $\mathrm{id} : \mathcal{F} \dot{\rightarrow} \mathcal{B}$ *forms a monad morphism from the monad $\mathcal{F}$ to the monad $\mathcal{B}$.*

PROOF. Since the forgetful functor $\mathcal{U} : \mathbb{KM}et_1 \rightarrow \mathbb{M}es$ and the natural transformation $\mathrm{id} : \mathcal{PU} \dot{\rightarrow} \mathcal{UF}$ form a monad morphism from the monad $\mathcal{F}$

to the monad $\mathcal{P}$, the diagram

$$\begin{array}{ccc} \mathbb{KM}et_1 & \xrightarrow{\mathcal{F}} & \mathbb{KM}et_1 \\ {\scriptstyle\mathcal{U}}\downarrow & & \downarrow{\scriptstyle\mathcal{U}} \\ \mathbb{M}es & \xrightarrow[\mathcal{P}]{} & \mathbb{M}es \end{array}$$

commutes. As a consequence, the functor $\mathcal{F}$ maps a 1-bounded compact metric space X to the set of Borel probability measures on X endowed with a metric $d_{\mathcal{F}X}$.

Next, we will prove that $d_{\mathcal{B}X} \leq d_{\mathcal{F}X}$. By definition, $d_{\mathcal{B}X}$ is the smallest distance function for which integration of nonexpansive functions is nonexpansive. That is, for each nonexpansive function $f : X \to [0,1]$, the function $i_f : \mathcal{B}X \to [0,1]$, as introduced in Proposition 9, is nonexpansive. (A similar characterization of the probabilistic powerdomain can be found in [34, Theorem 4.2].) Hence, it suffices to show that integration of nonexpansive functions is nonexpansive with respect to $d_{\mathcal{F}X}$, that is, $i_f : \mathcal{F}X \to [0,1]$ is nonexpansive.

Let $f : X \to [0,1]$ be a nonexpansive function. Let μ be a probability measure on X. Then

$$\begin{aligned} & \varepsilon_{\{\mathbf{0}\}}(\boldsymbol{\mu_2}(\mathcal{F}(\varepsilon^{-1}_{\{\mathbf{0}\}} \circ f)(\mu))) \\ &= \boldsymbol{\mu_2}(\mathcal{F}(\varepsilon^{-1}_{\{\mathbf{0}\}} \circ f)(\mu))(\{\mathbf{0}\}) \\ &= \int_{\mathcal{F}X} \varepsilon_{\{\mathbf{0}\}} \, d(\mathcal{F}(\varepsilon^{-1}_{\{\mathbf{0}\}} \circ f)(\mu)) \\ &= \int_{\mathcal{F}X} \varepsilon_{\{\mathbf{0}\}} \circ \varepsilon^{-1}_{\{\mathbf{0}\}} \circ f \, d\mu \\ &= i_f. \end{aligned}$$

Since $\varepsilon_{\{\mathbf{0}\}}$ is an isomorphism, both $\varepsilon_{\{\mathbf{0}\}}$ and $\varepsilon^{-1}_{\{\mathbf{0}\}}$ are nonexpansive. Since $\boldsymbol{\mu_2}$ and f are also nonexpansive, i_f is nonexpansive as well. $\square$

Let us briefly discuss the condition that $\mathcal{F}\mathbf{2}$ and $[0,1]$ are isomorphic via $\varepsilon_{\{\mathbf{0}\}}$. Notice that if $\varepsilon_{\{\mathbf{0}\}}$ is an isomorphism, then $\varepsilon_{\{\mathbf{1}\}}$ is an isomorphism as well. Obviously, the sets underlying $\mathcal{F}\mathbf{2}$ and $[0,1]$ are isomorphic via $\varepsilon_{\{\mathbf{0}\}}$. Hence, it is natural to require that $\varepsilon_{\{\mathbf{0}\}}$ is an isomorphism. One encounters a similar condition when contrasting categories of deterministic and stochastic models. In that case, $\mathcal{F}\mathbf{2}$ is the "unit interval" [42]. Integration of a "real-valued" function $f : X \to \mathcal{F}\mathbf{2}$ can be defined as $\boldsymbol{\mu_2} \circ \mathcal{F}f : \mathcal{F}X \to \mathcal{F}\mathbf{2}$.

Since every measure on a compact metric space is tight (see, for example, [46, Section II.3]), the diagram

$$\begin{array}{ccc} \mathbb{KM}et_1 & \xrightarrow{\mathcal{B}} & \mathbb{KM}et_1 \\ \big\downarrow & & \big\downarrow \\ \mathbb{CM}et_1 & \xrightarrow[\mathcal{B}]{} & \mathbb{CM}et_1 \end{array}$$

commutes. Also in this case, the inclusion functor and the identity natural transformation form a monad morphism between the monads. Note, however, that the diagram

$$\begin{array}{ccc} \mathbb{CM}et_1 & \xrightarrow{\mathcal{B}} & \mathbb{CM}et_1 \\ {\scriptstyle\mathcal{U}}\big\downarrow & & \big\downarrow{\scriptstyle\mathcal{U}} \\ \mathbb{M}es & \xrightarrow[\mathcal{P}]{} & \mathbb{M}es \end{array}$$

does not commute. Consider, for example, the set $\mathbb{R}$ endowed with the discrete metric (different real numbers have distance 1). The Borel probability measure μ defined by

$$\mu(B) = \begin{cases} 0 \text{ if } B \text{ is countable} \\ 1 \text{ otherwise} \end{cases}$$

is not tight, since $\mu(B) = 0$ for each compact (in this case, finite) subset B of $\mathbb{R}$.

To show that the monad $\mathcal{B}'$ on $\mathbb{KM}et_1$ extends the monad $\mathcal{P}'$ on $\mathbb{M}es$, it suffices to show that the monad $\mathbf{1} + -$ on $\mathbb{KM}et_1$ extends the monad $\mathbf{1} + -$ on $\mathbb{M}es$. Obviously, $(\mathbf{1} + -)\mathcal{U} = \mathcal{U}(\mathbf{1} + -)$. The forgetful functor $\mathcal{U} : \mathbb{KM}et_1 \to \mathbb{M}es$ and the natural transformation $\mathrm{id} : (\mathbf{1} + -)\mathcal{U} \stackrel{\cdot}{\to} \mathcal{U}(\mathbf{1} + -)$ form a monad morphism from the monad $\mathbf{1} + -$ on $\mathbb{KM}et_1$ to the monad $\mathbf{1} + -$ on $\mathbb{M}es$. Composing this monad morphism with the monad morphism from $\mathcal{B}$ to $\mathcal{P}$ gives us a monad morphism from $\mathcal{B}'$ to $\mathcal{P}'$. Note that $\mathcal{B}'\mathbf{1} = \mathcal{B}(\mathbf{1} + \mathbf{1}) = \mathcal{B}\mathbf{2}$. Therefore, $\varepsilon_{\{\mathbf{0}\}} : \mathcal{B}'\mathbf{1} \to [0, 1]$ is an isomorphism. For the monad $\mathcal{B}'$ we can prove a universal characterization similar to Proposition 13.

7 Conclusion

Let us first summarize our main contributions, before discussing some related and future work. We have shown that the functor $\mathcal{B}$ can be extended to a monad on the categories $\mathbb{KM}et_1$ and $\mathbb{CM}et_1$. Furthermore, we have demonstrated that the monad $\mathcal{B}'$ extends the monad $\mathcal{P}'$ in the same way as the monad $\mathcal{V}$ extends $\mathcal{P}'$.

As advocated in, for example, [47], computational effects like probabilistic nondeterminism determine monads, where such a monad is generated by a family of operations and their equational theory. Therefore, one may try to relate the monads $\mathcal{P}$, $\mathcal{V}$ and $\mathcal{B}$ by comparing the operations and equations determining $\mathcal{P}$, $\mathcal{V}$ and $\mathcal{B}$. However, there are many different families of operations (and corresponding equational theories) that determine one and the same monad (see, for example, [32] for a number of different axiomatizations of the probabilistic powerdomain). Furthermore, the equational theory often also captures properties of the underlying category. The monad $\mathcal{B}$ can, however, be captured by a $\frac{1}{2}$-contractive binary operation, say $\oplus$, and the equations

$$\begin{aligned} x \oplus x &= x \\ x \oplus y &= y \oplus x \\ (v \oplus w) \oplus (x \oplus y) &= (x \oplus w) \oplus (v \oplus y) \end{aligned}$$

The operation $\oplus$ can be viewed as a probabilistic choice. That is, in $x \oplus y$, the x is chosen with probability $\frac{1}{2}$ and so is the y. Note that we require $\oplus$ to be $\frac{1}{2}$-contractive (rather than nonexpansive). That is,

$$d_X(v \oplus w, x \oplus y) \leq \tfrac{1}{2}(d_X(v, x) + d_X(w, y)).$$

Now consider

$$\begin{aligned} y_0 &= y & z_0 &= z \\ y_{n+1} &= x \oplus y_n & z_{n+1} &= x \oplus z_n \end{aligned}$$

In y_n, the x is chosen with probability $1 - 2^{-n}$ and the y is chosen with probability 2^{-n}. Since $\oplus$ is $\frac{1}{2}$-contractive, we have that

$$d_X(y_n, z_n) \leq 2^{-n} d_X(y, z).$$

Hence, the metric d_X takes the probabilities into account. If $\oplus$ were nonexpansive, then we would have that

$$d_X(y_n, z_n) \leq d_X(y, z).$$

In that case, the probabilities would not be reflected in the metric d_X. Some more details can be found in [10]. A very similar characterization of the probabilistic powerdomain can be found in [32].

An alternative approach to relate the monads $\mathcal{V}$ and $\mathcal{B}'$ would be to consider generalized metric spaces. These spaces are a common generalization of partial orders and metric spaces (see, for example, [41]). Generalized metric spaces have been successfully exploited to reconcile numerous fundamental notions and constructions for partial orders and metric spaces. For example, in [9], a completion, two topologies, and three powerdomains are studied. In [55, Section 4.8], the functor $\mathcal{B}'$ is defined for the category $\mathbb{GM}et$ of generalized metric spaces and nonexpansive functions. This functor can be extended to a monad on $\mathbb{GM}et$ alike the construction given in Section 5. We conjecture that this monad reconciles the monads $\mathcal{V}$ and $\mathcal{B}'$.

Let $\mathbb{C}$ be a category and $\mathcal{F} : \mathbb{C} \to \mathbb{C}$ a functor. An $\mathcal{F}$-coalgebra is a pair $\langle C, f \rangle$ consisting of an object C in $\mathbb{C}$ and a morphism $f : C \to \mathcal{F}C$ in $\mathbb{C}$. Let $\mathcal{F}$ be one of the functors studied in this paper. Then the $\mathcal{F}$-coalgebras can be seen as probabilistic transition systems (see, for example, [54]). The object C represents the state space and the morphism $f : C \to \mathcal{F}C$ captures the transitions. As shown in [12,45], the category of $\mathcal{P}'$-coalgebras has a terminal object. This terminal $\mathcal{P}'$-coalgebra captures probabilistic bisimilarity [39] in the following way. The kernel of the unique morphism from a $\mathcal{P}'$-coalgebra to the terminal $\mathcal{P}'$-coalgebra is probabilistic bisimilarity. Since the functor $\mathcal{V}$ is locally continuous, also the category of $\mathcal{V}$-coalgebras has a terminal object (see, for example, [11]). Van Breugel, Hermida, Makkai and Worrell [10] have recently proved that the category of $\mathcal{B}'$-coalgebras also has a terminal object. Both the terminal $\mathcal{V}$-coalgebra and the terminal $\mathcal{B}'$-coalgebra capture probabilistic bisimilarity.

Recall that the monad $\mathcal{B}$ satisfies the following universal property: for every monad $\mathcal{F}$ on $\mathbb{KM}et_1$ that extends the monad $\mathcal{P}$, the natural transformation id : $\mathcal{F} \stackrel{.}{\to} \mathcal{B}$ is a morphism in the category of monads on $\mathbb{KM}et_1$. Since $\mathcal{F}$ extends $\mathcal{P}$, it has to map a 1-bounded compact metric space X to the set of Borel probability measures on X endowed with a metric $d_{\mathcal{F}X}$. Because id : $\mathcal{F} \stackrel{.}{\to} \mathcal{B}$ is a natural transformation in $\mathbb{KM}et_1$, each component $\text{id}_X : \mathcal{F}X \to \mathcal{B}X$ is nonexpansive, that is, $d_{\mathcal{B}X} \leq d_{\mathcal{F}X}$. Hence, the Kantorovich metric is the smallest among all those that extend the monad $\mathcal{P}$. Since the unit of the monad $\mathcal{F}$ is a natural transformation in $\mathbb{KM}et_1$, each component, mapping an element to its Dirac measure, has to be nonexpansive. Because each component of the unit of $\mathcal{B}$ is an isometric embedding, we can conclude that $d_{\mathcal{B}X} \geq d_{\mathcal{F}X}$ when restricted to the Dirac measures. Therefore, the Kantorovich metric restricted to the Dirac measures is the largest among all those that extend the monad $\mathcal{P}$. Hence, one may wonder whether the Kantorovich metric is the unique extension. We have not been able to prove this yet. Neither have we been able to find a metric different from the Kantorovich metric that extends the monad $\mathcal{P}$. The Prokhorov metric [48], which like the Kantorovich metric metrizes the weak topology, seems a candidate. However, if we were able to prove that each component of the multiplication is nonexpansive with respect to the Prokhorov metric, then we would not only show that the Kantorovich metric is not the unique extension. At the same time we would also improve the known bound between the Kantorovich metric d_K and the Prokhorov metric d_P, since it is only known that ${d_P}^2 \leq d_K \leq 2d_P$ (see, for example, [27]).

References

[1] M. Alvarez-Manilla, A. Edalat, and N. Saheb-Djahromi. An extension result for continuous valuations. *Journal of the London Mathematical Society*, 61(2):629–640, April 2000.

[2] W. Arveson. *An Invitation to C^*-Algebras*, volume 39 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1976.

[3] C. Baier and M. Kwiatkowska. Domain equations for probabilistic processes. *Mathematical Structures in Computer Science*, 10(6):665–717, December 2000.

[4] M. Barr and C. Wells. *Toposes, Triples and Theories*, volume 278 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, 1985.

[5] J. Beck. Distributive laws. In A. Hold and B. Eckmann, editors, *Seminar on Triples and Categorical Homology Theory*, volume 80 of *Lecture Notes in Mathematics*, pages 119–140, Zürich, 1966–1967. Springer-Verlag.

[6] N. Benton, J. Hughes, and E. Moggi. Monads and effects. In G. Barthe, P. Dybjer, L. Pinto, and J. Saraiva, editors, *Advanced Lectures of the International Summer School on Applied Semantics*, volume 2395 of *Lecture Notes in Computer Science*, pages 42–122, Caminha, September 2000. Springer-Verlag.

[7] P. Billingsley. *Probability and Measure*. John Wiley & Sons, New York, third edition, 1995.

[8] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, New York, second edition, 1999.

[9] M.M. Bonsangue, F. van Breugel, and J.J.M.M. Rutten. Generalized metric spaces: completion, topology, and powerdomains via the Yoneda embedding. *Theoretical Computer Science*, 193(1/2):1–51, February 1998.

[10] F. van Breugel, C. Hermida, M. Makkai, and J. Worrell. An accessible approach to behavioural pseudometrics. In *Proceedings of the 32nd International Colloquium on Automata, Languages, and Programming*, Lecture Notes in Computer Science, Lisbon, July 2005. Springer-Verlag.

[11] F. van Breugel, M. Mislove, J. Ouaknine, and J. Worrell. Domain theory, testing and simulation for labelled Markov processes. *Theoretical Computer Science*, 333(1/2):171–197, March 2005.

[12] F. van Breugel, S. Shalit, and J. Worrell. Testing labelled Markov processes. In P. Widmayer, F. Triguero, R. Morales, M. Hennessy, S. Eidenbenz, and R. Conejo, editors, *Proceedings of the 29th International Colloquium on Automata, Languages, and Programming*, volume 2380 of *Lecture Notes in Computer Science*, pages 537–548, Malaga, July 2002. Springer-Verlag.

[13] F. van Breugel and J. Worrell. Towards quantitative verification of probabilistic systems. In F. Orejas, P.G. Spirakis, and J. van Leeuwen, editors, *Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, volume 2076 of *Lecture Notes in Computer Science*, pages 421–432, Crete, July 2001. Springer-Verlag.

[14] F. van Breugel and J. Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331(1):115–142, February 2005.

[15] L. de Alfaro. Quantitative verification and control via the mu-calculus. In R.M. Amadio and D. Lugiez, editors, *Proceedings of the 14th International Conference on Concurrency Theory*, volume 2761 of *Lecture Notes in Computer Science*, pages 102–126, Marseille, September 2003. Springer-Verlag.

[16] L. de Alfaro, T.A. Henzinger, and R. Majumdar. Discounting the future in systems theory. In J.C.M. Baeten, J.K. Lenstra, J. Parrow, and G.J. Woeginger, editors, *Proceedings of the 30th International Colloquium on Automata, Languages and Programming*, volume 2719 of *Lecture Notes in Computer Science*, pages 1022–1037, Eindhoven, June/July 2003. Springer-Verlag.

[17] Y. Deng, T. Chothia, C. Palamidessi, and J. Pang. Metrics for action-labelled quantitative transition systems. In *Proceedings of 3rd Workshop on Quantitative Aspects of Programming Languages*, Electronic Notes in Theoretical Computer Science, Edinburgh, April 2005. Elsevier.

[18] J. Desharnais. *Labelled Markov Processes.* PhD thesis, McGill University, Montreal, November 1999.

[19] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labeled Markov systems. In J.C.M. Baeten and S. Mauw, editors, *Proceedings of the 10th International Conference on Concurrency Theory*, volume 1664 of *Lecture Notes in Computer Science*, pages 258–273, Eindhoven, August 1999. Springer-Verlag.

[20] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proceedings of the 17th Annual IEEE Symposium on Logic in Computer Science*, pages 413–422, Copenhagen, July 2002. IEEE.

[21] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, June 2004.

[22] E.-E. Doberkat. Derandomizing probabilistic semantics through Eilenberg-Moore algebras for the Giry monad. Technical Report 152, University of Dortmund, August 2004.

[23] R.L. Dobrushin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory of Probability and its Applications*, 13(2):197–224, 1968.

[24] G.A. Edgar. *Integral, Probability, and Fractal Measures*. Springer-Verlag, New York, 1998.

[25] R. Engelking. *General Topology*, volume 6 of *Sigma Series in Pure Mathematics*. Heldermann Verlag, Berlin, revised and completed edition, 1989.

[26] A. Giacalone, C.-C. Jou, and S.A. Smolka. Algebraic reasoning for probabilistic concurrent systems. In *Proceedings of the IFIP WG 2.2/2.3 Working Conference on Programming Concepts and Methods*, pages 443–458, Sea of Gallilee, April 1990. North-Holland.

[27] A.L. Gibbs and F.E. Su. On choosing and bounding probability metrics. *International Statistical Review*, 70(3):419–435, December 2002.

[28] G. Gierz, K.H. Hofmann, K. Keimel, J.D. Lawson, M. Mislove, and D.S. Scott. *Continuous Lattices and Domains*. Cambridge University Press, 2003.

[29] M. Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Proceedings of the International Conference on Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85, Ottawa, August 1981. Springer-Verlag.

[30] V. Gupta, R. Jagadeesan, and P. Panangaden. Approximate reasoning for real-time probabilistic processes. In *Proceedings of the 1st International Conference on the Quantitative Evaluation of Systems*, pages 304–313, Enschede, September 2004. IEEE.

[31] J.I. den Hartog. *Probabilistic Extensions of Semantical Models*. PhD thesis, Free University, Amsterdam, October 2002.

[32] R. Heckmann. Probabilistic domains. In S. Tison, editor, *Proceedings of the 19th International Colloquium on Trees in Algebra and Programming*, volume 787 of *Lecture Notes in Computer Science*, pages 142–156, Edinburgh, April 1994. Springer-Verlag.

[33] J.E. Hutchinson. Fractals and self similarity. *Indiana University Mathematics Journal*, 30(5):713–747, 1981.

[34] C. Jones. *Probabilistic Non-Determinism*. PhD thesis, University of Edinburgh, August 1989.

[35] C. Jones and G.D. Plotkin. A probabilistic powerdomain of evaluations. In *Proceedings of the 4th Annual IEEE Symposium on Logic in Computer Science*, pages 186–195, Pacific Grove, CA, June 1989. IEEE.

[36] A. Jung and R. Tix. The troublesome probabilistic powerdomain. In A. Edalat, A. Jung, K. Keimel, and M. Kwiatkowska, editors, *Proceedings of 2nd Workshop on Computation and Approximation*, volume 13 of *Electronic Notes in Theoretical Computer Science*, Birmingham, September 1997. Elsevier.

[37] L. Kantorovich. On the transfer of masses (in Russian). *Doklady Akademii Nauk*, 5(1):1–4, 1942. Translated in *Management Science*, 5(1):1–4, October 1958.

[38] M. Kwiatkowska and G.J. Norman. Probabilistic metric semantics for a simple language with recursion. In W. Penczek and A. Szalas, editors, *Proceedings of the 21st International Symposium on Mathematical Foundations of Computer Science*, volume 1113 of *Lecture Notes in Computer Science*, pages 419–430, Cracow, September 1996. Springer-Verlag.

[39] K.G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94(1):1–28, September 1991.

[40] F.W. Lawvere. The category of probabilistic mappings. 1962.

[41] F.W. Lawvere. Metric spaces, generalized logic, and closed categories. *Rendiconti del Seminario Matematico e Fisico di Milano*, 43:135–166, 1973.

[42] F.W. Lawvere. Personal communication. September 2004.

[43] S. Mac Lane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1998.

[44] E. Moggi. Notions of computation and monads. *Information and Computation*, 93(1):55–92, July 1991.

[45] L.S. Moss and I.D. Viglizzo. Harsanyi type spaces and final coalgebras constructed from satisfied theories. In J. Adámek and S. Milius, editors, *Proceedings of the 7th Workshop on Coalgebraic Methods in Computer Science*, volume 106 of *Electronic Notes in Theoretical Computer Science*, pages 279–295, Barcelona, March 2004. Elsevier.

[46] K.R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, New York, 1967.

[47] G.D. Plotkin and J. Power. Notions of computation determine monads. In M. Nielsen and U. Engberg, editors, *Proceedings of the 5th International Conference on Foundations of Software Science and Computation Structures*, volume 2303 of *Lecture Notes in Computer Science*, pages 342–356, Grenoble, April 2002. Springer-Verlag.

[48] Yu.V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability and its Applications*, 1:157–214, 1956.

[49] S.T. Rachev. *Probability Metrics and the Stability of Stochastic Models*. John Wiley & Sons, Chichester, NY, 1991.

[50] N. Ramsey and A. Pfeffer. Stochastic lambda calculus and monads of probability distributions. In *Proceedings of the 29th SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 154–165, Portland, OR, January 2002. ACM.

[51] N. Saheb-Djahromi. CPO's of measures for nondeterminism. *Theoretical Computer Science*, 12(1):19–37, September 1980.

[52] R. Street. The formal theory of monads. *Journal of Pure and Applied Algebra*, 2(2):149–168, July 1972.

[53] L. Vaserstein. Markov processes over denumerable products of spaces, describing large systems of automata. *Problems in Information Transmission*, 5(3):47–52, July/September 1969.

[54] E.P. de Vink and J.J.M.M. Rutten. Bisimulation for probabilistic transition systems: a coalgebraic approach. *Theoretical Computer Science*, 221(1/2):271–293, June 1999.

[55] J. Worrell. *On Coalgebras and Final Semantics*. PhD thesis, Oxford University, October 2000.

[56] M.S. Ying. Additive models of probabilistic processes. *Theoretical Computer Science*, 275(1/2):481–519, March 2002.