Combining probabilistic and non-deterministic choice via weak distributive laws

Alexandre Goy
Université Paris-Saclay, CentraleSupélec, MICS
Gif-sur-Yvette, France
alexandre.goy@centralesupelec.fr

Daniela Petrişan
Université de Paris, IRIF
Paris, France
petrisan@irif.fr

Abstract

Combining probabilistic choice and non-determinism is a long standing problem in denotational semantics. From a category theory perspective, the problem stems from the absence of a distributive law of the powerset monad over the distribution monad. In this paper we prove the existence of a *weak* distributive law of the powerset monad over the finite distribution monad. As a consequence, we retrieve the well-known convex powerset monad as a weak lifting of the powerset monad to the category of convex algebras. We provide applications to the study of trace semantics and behavioral equivalences of systems with an interplay between probability and non-determinism.

Keywords: weak distributive law, weak lifting, finite distribution monad, convex powerset monad, trace semantics, behavioral equivalence, probability, non-determinism, probabilistic automata

1 Introduction

The question of combining non-deterministic and probabilistic choice has a long history in the semantics of programming languages. In domain theory there has been a lot of work in this direction by Tix, Kleimel and Plotkin [16, 29], Mislove [21] and Varacca and Winskel [30, 31]. More recently, the same notorious difficulties appear in the coalgebraic semantics of Segala systems [6, 7].

We follow the well established tradition of modeling various effects using the category theoretic concept of a monad. Non-determinism can be modeled via powerdomains [23], or, more simplistically, if we restrain ourselves to working in the category of sets and functions, by the powerset monad $\mathcal{P}\colon \mathsf{Set} \to \mathsf{Set}$. For probabilistic choice, there is also a variety of probabilistic powerdomains that have been considered [14, 15, 25]. Again, for the sake of simplicity, in this paper we restrict our attention to the finite distribution monad $\mathcal{D}\colon \mathsf{Set} \to \mathsf{Set}$ mapping a set X to the set of the finitely supported distributions on X.

Modular semantics for computational effects could be achieved if the corresponding monads compose well. However, composing two monads is not straightforward, and in general, additional conditions are required. For example, if we consider two monads (T, η^T, μ^T) and (S, η^S, μ^S) with their respective units and multiplications, how can we define a

monad structure on the composed functor ST ? Defining the multiplication $\mu^{ST}\colon STST \Rightarrow ST$ from the multiplication of S and that of T would be rather easy should we have a way of 'swapping' S and T, via a natural transformation $\gamma\colon TS \Rightarrow ST$, for then we could define the desired multiplication as the composite: $STST \xrightarrow{S\gamma T} SSTT \xrightarrow{\mu^S TT} STT \xrightarrow{S\mu^T} ST$. This approach to composing monads is due to the seminal work of Jon Beck [2]. The natural transformation $\gamma\colon TS \Rightarrow ST$ should satisfy four additional axioms, recalled below in Definition 2.5, and is called a distributive law of the monad S over the monad T. The four additional axioms essentially state that γ interacts well with the multiplications, respectively with the units of the two monads. Beck showed that distributive laws of the monad S over the monad T are in one-to-one correspondence with liftings of S to the category of Eilenberg-Moore algebras for T and with extensions of T to the Kleisli category for the monad S.

Unfortunately, as we know from Varacca's thesis [30], there is no distributive law of the monad $\mathcal{P}$, modeling nondeterminism, over the monad $\mathcal{D}$, modeling probabilistic choice. The proof of this fact, attributed to Plotkin, is by contradiction and relies on building a counterexample by smartly manipulating the naturality of the assumed law and the axiom corresponding to the unit $\eta^{\mathcal{D}}$ of the monad $\mathcal{D}$.

More negative results have been proved since: In the same spirit, Klin and Salamanca [17] showed that there is no distributive law of the powerset monad over itself, and that furthermore, there is no monad structure on the iterated powerset functor. Zwart and Marsden [33] generalized Plotkin's example and provided a fine analysis of non-existence of distributive laws of various monads.

Coming back to combining non-deterministic and probabilistic effects, various way of circumventing the problem have been proposed: Varacca [30] uses instead of $\mathcal{D}$ the monad of indexed valuations and shows the existence of a distributive law in this setting. A different approach, coming from domain theory, is to define directly a powerset-like monad on the category of Eilenberg-Moore algebras for the monad $\mathcal{D}$. The objects of this category are called convex algebras in [27], or barycentric algebras in [16], and the analogous of the powerset monad is the convex powerset monad $\mathcal{P}_c$ mapping a convex algebra to the convex algebra of its convex subsets. This monad was considered, for example in [6]. In domain theory the analogous construction is called

the power Kegelspitzen [16]. A composite monad $\mathcal{P}_c\mathcal{D}$ of convex sets of distributions was considered in [12, 20].

A question that arises naturally is whether the convex powerset monad $\mathcal{P}_c$ and the composite monad $\mathcal{P}_c\mathcal{D}$ described in [20] can be obtained in a systematic way. How far away are we from having a distributive law $\mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}$? Can we at least obtain a natural transformation? If yes, which of the four axioms from the definition of a distributive law does it satisfy?

The first observation is that we have a canonical extension of the functor $\mathcal{D}$ to the category Rel of sets and relations, which is precisely the Kleisli category of the powerset monad. This is due to a result of Barr which states that weak pullback preserving functors can be extended to locally monotone functors on Rel. Indeed, the functor $\mathcal{D}$ preserves weak pullbacks, see [8, 22], hence it can be extended to Rel. We thus obtain a canonical distributive law

$$\delta \colon \mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}$$

of the monad $\mathcal{P}$ over the *functor* $\mathcal{D}$. This already entails that the natural transformation δ interacts well with the unit and multiplication of the monad $\mathcal{P}$. So which of the remaining two axioms fail? We will see that the interaction with the unit $\eta^\mathcal{D}$ of $\mathcal{D}$ is problematic, but the diagram corresponding to the interaction with the multiplication $\mu^\mathcal{D}$ does commute.

Weaker notions of distributive laws have been studied by Street [28] and Böhm [3]. One such notion – precisely the one we need – involves dropping one of the unit axioms is the centerpiece in the work of Garner [11]. He simply calls this notion a weak distributive law and hereafter we adopt his terminology. Garner proves that the powerset monad $\mathcal{P}$ weakly distributes over the Stone-Čech compactification monad β. Just as in the case of distributive laws, weak distributive laws $\delta \colon TS \Rightarrow ST$ correspond to a suitable notion of weak liftings of S to the Eilenberg-Moore category of T. In the particular case of the weak distributive law $\beta\mathcal{P} \Rightarrow \mathcal{P}\beta$ exhibited by Garner, the corresponding weak lifting of $\mathcal{P}$ to the category of Eilenberg-Moore algebras for β – which is known to be the category of compact Hausdorff spaces – is exactly the Vietoris monad.

A first contribution of our paper is the following theorem

Theorem 1.1. *The natural transformation $\delta \colon \mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}$ is a weak distributive law of the powerset monad $\mathcal{P}$ over the finite distribution monad $\mathcal{D}$. The corresponding weak lifting of $\mathcal{P}$ to the category of convex algebras is the convex powerset monad $\mathcal{P}_c$.*

The proof of this theorem relies on proving a certain property of the multiplication of the monad $\mathcal{D}$, namely that it is a weakly cartesian natural transformation. We achieve this using a well known result from mathematical optimization, Farkas' Lemma.

Furthermore, just as a distributive law $TS \Rightarrow ST$ induces a monad structure on the composite functor ST, a weak distributive law $TS \Rightarrow ST$ induces a composite monad $\widetilde{ST}$ whose Eilenberg-Moore algebras are the algebras for the weak lifting of S. In our case, the composite monad is exactly the monad $\mathcal{P}_c\mathcal{D}$ of convex subsets of probability distributions, considered in [12, 20].

The existence of a weak distributive law sheds new light on the compositional aspect of combining non-deterministic and probabilistic choice. There are several instances in the literature where the absence of a distributive law between the two monads is deplored. The second part of the paper is devoted to such applications:

First, we revisit the results of [6] on behavioral equivalences for probabilistic automata. In *loc. cit.* Bonchi et al. provide a coalgebraic account of obtaining from a probabilistic automata a belief-state transformer, i.e., a labelled transition system whose states are probability distributions. Had a distributive law $\mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}$ existed, this procedure would have been yet another example of the so-called generalized powerset construction [26]. In the absence of a distributive law [6] redevelops the theory by hand and in particular constructs the monad $\mathcal{P}_c$ directly on the category of convex algebras – the Eilenberg-Moore category for $\mathcal{D}$. In Section 5 we extend the generalized powerset construction of [26] to the setting of weak distributive laws and, as a corollary, we obtain a systematic way of turning a probabilistic automaton into a belief-state transformer.

A second application concerns trace semantics of systems combining probabilistic and non-deterministic features – as studied in [7] – using tools akin to the generalized determinization. The lack of a proper distributive law means that the authors of [7] had to build from scratch an algebraic presentation for the algebras of the composite monad $\mathcal{P}_c\mathcal{D}$ and to mend the mechanism of the generalized determinization accordingly. Applying the results of [3], we automatically obtain the presentation of $\mathcal{P}_c\mathcal{D}$: its algebras can be alternatively described as δ-algebras, that is, sets carrying simultaneously $\mathcal{P}$- and $\mathcal{D}$- algebraic structures related via an additional distributivity axiom.

Synopsis. After recalling the necessary preliminaries on weak distributive laws in Section 2, we exhibit the canonical weak distributive law of the powerset monad over the finite distribution monad in Section 3. The corresponding weak lifting of the powerset monad is the convex powerset monad. We compute this weak lifting and discuss a presentation for its algebras in Section 4. In Section 5 we provide a coalgebraic generalized powerset construction for weak distributive laws and we apply it to non-deterministic probabilistic systems.

2 Monads and Weak Distributive Laws

This section presents some of the preliminary notions needed in the sequel. After recalling the monads of interest, the second part of the section is dedicated to the framework of weak distributive laws [3], adopting Garner's terminology [11].

2.1 Monads

To fix the notation we recall the basic notions of monad, Eilenberg-Moore algebras and the Kleisli category for a monad, see [19].

Definition 2.1 (Monad). A *monad* $\mathbb{T} = (T, \eta, \mu)$ consists of a functor $T : \mathsf{C} \to \mathsf{C}$ along with two natural transformations, the unit $\eta : Id \Rightarrow T$ and the multiplication $\mu : TT \Rightarrow T$ such that the following diagrams commute:

$$
\begin{array}{ccc}
T & \xrightarrow{\eta T} & TT \\
{\scriptstyle T\eta}\downarrow & \parallel \;\; \downarrow{\scriptstyle \mu} & \\
TT & \xrightarrow{\mu} & T
\end{array}
\qquad
\begin{array}{ccc}
TTT & \xrightarrow{\mu T} & TT \\
{\scriptstyle T\mu}\downarrow & & \downarrow{\scriptstyle \mu} \\
TT & \xrightarrow{\mu} & T
\end{array}
$$

An *Eilenberg-Moore algebra* (X, a) for the monad $\mathbb{T}$ is a morphism $a \colon TX \to X$ satisfying $a \circ \mu_X = a \circ Ta$ and $a \circ \eta_X = id_X$. A morphism between algebras (X, a) and (Y, b) is a morphism $f \colon X \to Y$ so that $f \circ a = b \circ Tf$. The category of Eilenberg-Moore algebras for $\mathbb{T}$ is denoted by $\mathbf{EM}(\mathbb{T})$. The *Kleisli category* of $\mathbb{T}$, denoted by $\mathbf{Kl}(\mathbb{T})$ is isomorphic to the category of free $\mathbb{T}$-algebras and can be described as follows. Its objects are the objects of C and a morphism $X \twoheadrightarrow Y$ in $\mathbf{Kl}(\mathbb{T})$ is a morphism $X \to \mathbb{T}Y$ in C.

We recall the definitions of the monads of interest in this paper and describe their categories of Eilenberg-Moore algebras.

Definition 2.2 (Powerset monad). The *powerset monad*, denoted by $(\mathcal{P}, \eta^{\mathcal{P}}, \mu^{\mathcal{P}})$, consists of the functor $\mathcal{P} : \mathsf{Set} \to \mathsf{Set}$ which sends a set X to the set $\mathcal{P}X$ of all subsets of X and a function $f : X \to Y$ to its direct image $\mathcal{P}f : \mathcal{P}X \to \mathcal{P}Y$, the unit $\eta^{\mathcal{P}}_X : X \to \mathcal{P}X$ given by $\eta^{\mathcal{P}}_X(x) = \{x\}$ and the multiplication $\mu^{\mathcal{P}}_X : \mathcal{P}\mathcal{P}X \to \mathcal{P}X$ given by $\mu^{\mathcal{P}}_X(\mathcal{A}) = \bigcup \mathcal{A}$.

An Eilenberg-Moore $\mathcal{P}$-algebra $a : \mathcal{P}X \to X$ can be seen as a complete lattice with the operation a describing either the sup or the inf.

Given a function $f : X \to Y$, define its *graph* $f_* : X \times Y \to \{0, 1\}$ by $f_*(x, y) = 1$ if and only if $y = f(x)$.

Definition 2.3 (Finite distribution monad). The *finite distribution monad* $(\mathcal{D}, \eta^{\mathcal{D}}, \mu^{\mathcal{D}})$ consists in the following data:

- The functor $\mathcal{D} : \mathsf{Set} \to \mathsf{Set}$ which maps a set X to the set $\mathcal{D}X$ of all finitely supported probability distributions on X, i.e.,

$$\mathcal{D}X = \left\{ \varphi : X \to [0, 1] : supp(\varphi) \text{ is finite and } \sum_{x \in X} \varphi(x) = 1 \right\}.$$

  It sends a function $f : X \to Y$ to the function $\mathcal{D}f : \mathcal{D}X \to \mathcal{D}Y$ defined by $\mathcal{D}f(\varphi)(y) = \sum_{x \in X} \varphi(x) f_*(x, y)$.

- The unit $\eta^{\mathcal{D}}_X : X \to \mathcal{D}X$ sends an element $x \in X$ to the Dirac distribution $\eta^{\mathcal{D}}_X(x)(y) = id_{X*}(x, y)$.
- The multiplication $\mu^{\mathcal{D}}_X : \mathcal{D}\mathcal{D}X \to \mathcal{D}X$ is defined by $\mu^{\mathcal{D}}_X(\Phi)(x) = \sum_{\varphi \in \mathcal{D}X} \Phi(\varphi)\varphi(x)$.

The category of Eilenberg-Moore algebras for $\mathcal{D}$ is denoted by $\mathbf{EM}(\mathcal{D})$. An Eilenberg-Moore $\mathcal{D}$-algebra $a : \mathcal{D}X \to X$ is called a *convex algebra* or a barycentric algebra, see for example [6, 16, 27, 32]. There are various concrete presentations by operations and equations of these algebras. We recall here the one from [16, Definition 2.2]. A convex algebra is a set X equipped with a binary operation $+_r$ for each real number $r \in [0, 1]$ satisfying the following equations

$$
\begin{aligned}
a +_1 b &= a \\
a +_r a &= a \\
a +_r b &= b +_{1-r} a \\
(a +_p b) +_r c &= a +_{pr} \left(b +_{\frac{r-pr}{1-pr}} c\right) \qquad \text{if } p, r \neq 1
\end{aligned}
$$

A morphism of convex algebras is an affine map, that is, a function $f : X \to Y$ so that $f(a +_r b) = f(a) +_r f(b)$. The fact that convex algebras $(X, (+_r)_{r \in [0,1]})$ and their morphisms are precisely the Eilenberg-Moore algebras for the finite distribution monad is well-known, see e.g., [9, 32].

Given a convex algebra $(X, (+_r)_{r \in [0,1]})$, a *convex subset* of X is a subalgebra of X, that is, a subset $A \in X$ so that for all $r \in [0, 1]$, $a, b \in A$ implies $a +_r b \in A$.

We give next a variant of the convex powerset monad, as defined in [6], the sole difference being that we also include the empty set. Given a convex algebra $(X, (+_r)_{r \in [0,1]})$ let $\mathcal{P}_c(X)$ be the set of convex subsets of X. Given two convex subsets A and B of X, we define

$$
\begin{aligned}
A +_r B &= \{a +_r b \mid a \in A, b \in B\} \qquad \text{if } r \in (0, 1) \\
A +_0 B &= B \\
A +_1 B &= A
\end{aligned}
$$

One can verify that $(\mathcal{P}_c(X), (+_r)_{r \in [0,1]})$ is indeed a convex algebra. On morphisms $\mathcal{P}_c$ is defined by direct images.

Definition 2.4 (Convex powerset monad). The *convex powerset monad* $(\mathcal{P}_c, \eta^{\mathcal{P}_c}, \mu^{\mathcal{P}_c})$ consists of the functor

$$\mathcal{P}_c : \mathbf{EM}(\mathcal{D}) \to \mathbf{EM}(\mathcal{D})$$
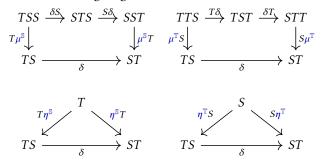
defined above, with the unit and multiplication defined, just as for the powerset monad: $\eta^{\mathcal{P}_c}_X(x) = \{x\}$, respectively $\mu^{\mathcal{P}_c}_X(\mathcal{A}) = \bigcup \mathcal{A}$.

Notice that [6] considers only the non-empty convex powerset and uses a different concrete presentation for convex algebra involving n-ary operations $(p_i)_{i=1}^n$ with $\sum p_i = 1$ for any natural number n. The definition of the full convex powerset, including the empty set, is nevertheless possible but requires a careful case analysis for the situations where some of the p_i-s are zero.
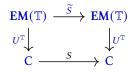
2.2 (Weak) Distributive Laws

In this section we recall the definitions of a distributive law [2] and of a weak distributive law [11] and their corresponding notions of liftings and extensions. Let $\mathbb{S} = (S, \eta^{\mathbb{S}}, \mu^{\mathbb{S}})$ and $\mathbb{T} = (T, \eta^{\mathbb{T}}, \mu^{\mathbb{T}})$ be two monads on a category C.

Definition 2.5 (Distributive law). A *distributive law* of $\mathbb{S}$ over $\mathbb{T}$ is a natural transformation $\delta : TS \Rightarrow ST$ such that the four following diagrams commute:

$$TSS \xrightarrow{\delta S} STS \xrightarrow{S\delta} SST \qquad TTS \xrightarrow{T\delta} TST \xrightarrow{\delta T} STT$$
$$T\mu^{\mathbb{S}} \downarrow \qquad\qquad \downarrow \mu^{\mathbb{S}}T \qquad \mu^{\mathbb{T}}S \downarrow \qquad\qquad \downarrow S\mu^{\mathbb{T}}$$
$$TS \xrightarrow{\quad\delta\quad} ST \qquad\qquad TS \xrightarrow{\quad\delta\quad} ST$$

$$T\eta^{\mathbb{S}} \swarrow \quad^{T}\quad \searrow \eta^{\mathbb{S}}T \qquad \eta^{\mathbb{T}}S \swarrow \quad^{S}\quad \searrow S\eta^{\mathbb{T}}$$
$$TS \xrightarrow{\quad\delta\quad} ST \qquad\qquad TS \xrightarrow{\quad\delta\quad} ST$$

Beck showed that distributive laws $\delta : TS \Rightarrow ST$ are in one-to-one correspondence with liftings of $\mathbb{S}$ to $\mathbf{EM}(\mathbb{T})$ and extensions of $\mathbb{T}$ to $\mathbf{Kl}(\mathbb{S})$, notions that we recall next.

Definition 2.6 (Lifting, extension). Let $U^{\mathbb{T}} : \mathbf{EM}(\mathbb{T}) \to \mathsf{C}$ be the forgetful functor from the category of $\mathbb{T}$-algebras. A *lifting* of $\mathbb{S}$ to $\mathbf{EM}(\mathbb{T})$ is a monad $\widetilde{\mathbb{S}} = (\widetilde{S}, \widetilde{\eta}^{\mathbb{S}}, \widetilde{\mu}^{\mathbb{S}})$ on $\mathbf{EM}(\mathbb{T})$ such that the following diagram commutes

$$\mathbf{EM}(\mathbb{T}) \xrightarrow{\widetilde{S}} \mathbf{EM}(\mathbb{T})$$
$$U^{\mathbb{T}} \downarrow \qquad\qquad \downarrow U^{\mathbb{T}}$$
$$\mathsf{C} \xrightarrow{\quad S\quad} \mathsf{C}$$

and, additionally, $U^{\mathbb{T}}\widetilde{\eta}^{\mathbb{S}} = \eta^{\mathbb{S}}U^{\mathbb{T}}$ and $U^{\mathbb{T}}\widetilde{\mu}^{\mathbb{S}} = \mu^{\mathbb{S}}U^{\mathbb{T}}$.

Let also $F_{\mathbb{S}} : \mathsf{C} \to \mathbf{Kl}(\mathbb{S})$ be the free functor into the Kleisli category of $\mathbb{S}$. An *extension* of $\mathbb{T}$ to $\mathbf{Kl}(\mathbb{S})$ is a monad $\widetilde{\mathbb{T}} = (\widetilde{T}, \widetilde{\eta}^{\mathbb{T}}, \widetilde{\mu}^{\mathbb{T}})$ on $\mathbf{Kl}(\mathbb{S})$ such that the following diagram commutes

$$\mathbf{Kl}(\mathbb{S}) \xrightarrow{\widetilde{T}} \mathbf{Kl}(\mathbb{S})$$
$$F_{\mathbb{S}} \uparrow \qquad\qquad \uparrow F_{\mathbb{S}}$$
$$\mathsf{C} \xrightarrow{\quad T\quad} \mathsf{C}$$

and, additionally, $\widetilde{\eta}^{\mathbb{T}}F_{\mathbb{S}} = F_{\mathbb{S}}\eta^{\mathbb{T}}$ and $\widetilde{\mu}^{\mathbb{T}}F_{\mathbb{S}} = F_{\mathbb{S}}\mu^{\mathbb{T}}$.

When things *almost* work well, it is possible to drop the last diagram of Definition 2.5 to define a weaker notion of distributive law.

Among the various possible ways of weakeaning the notion of distributive law considered in [3, 28], we consider the one which drops the axiom involving the unit $\eta^{\mathbb{T}}$ of $\mathbb{T}$ from Definition 2.5, and which Garner simply calls weak distributive law in [11].

Definition 2.7 (Weak distributive law). A *weak distributive law* of $\mathbb{S}$ over $\mathbb{T}$ is a natural transformation $\delta : TS \Rightarrow ST$

such that the first three diagrams of Definition 2.5 commute, that is, the diagrams involving the multiplications $\mu^{\mathbb{T}}, \mu^{\mathbb{S}}$ and the unit $\eta^{\mathbb{S}}$.

Example 2.8 (from [11]). There is a weak distributive law of the powerset monad $\mathcal{P}$ over the ultrafilter monad β given by $\delta_X : \beta\mathcal{P}X \to \mathcal{P}\beta X$ where
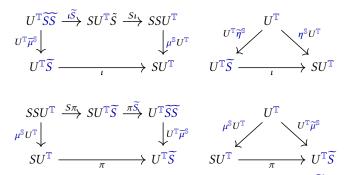
$$\delta_X(\mathfrak{F}) = \left\{ \mathcal{F} \in \beta X : \bigcup \mathcal{A} \in \mathcal{F} \text{ for all } \mathcal{A} \in \mathfrak{F} \right\} \qquad (1)$$

The notions of lifting and extension can be weakened accordingly:

Definition 2.9 (Weak lifting, weak extension). A *weak lifting* of $\mathbb{S}$ to $\mathbf{EM}(\mathbb{T})$ consists of a monad $\widetilde{\mathbb{S}}$ on $\mathbf{EM}(\mathbb{T})$ and two natural transformations

$$\iota : U^{\mathbb{T}}\widetilde{S} \Rightarrow SU^{\mathbb{T}} \text{ and } \pi : SU^{\mathbb{T}} \Rightarrow U^{\mathbb{T}}\widetilde{S}$$

such that $\pi\iota = 1$ and the four following diagrams commute:

$$U^{\mathbb{T}}\widetilde{SS} \xrightarrow{\iota\widetilde{S}} SU^{\mathbb{T}}\widetilde{S} \xrightarrow{S\iota} SSU^{\mathbb{T}} \qquad \qquad U^{\mathbb{T}}$$
$$U^{\mathbb{T}}\widetilde{\mu}^{\mathbb{S}} \downarrow \qquad\qquad \downarrow \mu^{\mathbb{S}}U^{\mathbb{T}} \qquad U^{\mathbb{T}}\widetilde{\eta}^{\mathbb{S}} \swarrow \quad \searrow \eta^{\mathbb{S}}U^{\mathbb{T}}$$
$$U^{\mathbb{T}}\widetilde{S} \xrightarrow{\quad\iota\quad} SU^{\mathbb{T}} \qquad\qquad U^{\mathbb{T}}\widetilde{S} \xrightarrow{\quad\iota\quad} SU^{\mathbb{T}}$$

$$SSU^{\mathbb{T}} \xrightarrow{S\pi} SU^{\mathbb{T}}\widetilde{S} \xrightarrow{\pi\widetilde{S}} U^{\mathbb{T}}\widetilde{SS} \qquad\qquad U^{\mathbb{T}}$$
$$\mu^{\mathbb{S}}U^{\mathbb{T}} \downarrow \qquad\qquad \downarrow U^{\mathbb{T}}\widetilde{\mu}^{\mathbb{S}} \qquad \mu^{\mathbb{S}}U^{\mathbb{T}} \swarrow \quad \searrow U^{\mathbb{T}}\widetilde{\mu}^{\mathbb{S}}$$
$$SU^{\mathbb{T}} \xrightarrow{\quad\pi\quad} U^{\mathbb{T}}\widetilde{S} \qquad\qquad SU^{\mathbb{T}} \xrightarrow{\quad\pi\quad} U^{\mathbb{T}}\widetilde{S}$$

A *weak extension* of $\mathbb{T}$ to $\mathbf{Kl}(\mathbb{S})$ consists of a functor $\widetilde{T} : \mathbf{Kl}(\mathbb{S}) \to \mathbf{Kl}(\mathbb{S})$ and a natural transformation $\widetilde{\mu}^{\mathbb{T}} : \widetilde{T}\widetilde{T} \Rightarrow \widetilde{T}$ such that $\widetilde{T}F_{\mathbb{S}} = F_{\mathbb{S}}T$ and $\widetilde{\mu}^{\mathbb{T}}F_{\mathbb{S}} = F_{\mathbb{S}}\mu^{\mathbb{T}}$.

The prominent correspondence between distributive laws, liftings and extensions extends to their weak variants, as proved in [3, 11]:

Theorem 2.10 (from [11]). *There is a bijection between weak distributive laws of $\mathbb{S}$ over $\mathbb{T}$ and weak extensions of $\mathbb{T}$ to $\mathbf{Kl}(\mathbb{S})$. Moreover, if idempotents split in C, these are also in bijection with weak liftings of $\mathbb{S}$ to $\mathbf{EM}(\mathbb{T})$.*

Let us elaborate a bit more on how to transform a weak extension into a weak distributive law and then into a weak lifting. Given a weak extension $\widetilde{T}$ of $\mathbb{T}$ to $\mathbf{Kl}(\mathbb{S})$, the weak distributive law $\delta : TS \Rightarrow ST$ is defined by $\delta_X = \widetilde{T}(\ni)$, where $\ni : SX \nrightarrow X$ is the Set-morphism id_{SX} viewed as a $\mathbf{Kl}(\mathbb{S})$-morphism. What is less usual is the next step, explained in more detail in [11]: the weak lifting $\widetilde{\mathbb{S}}$ is given on an T-algebra (X, a) by splitting the idempotent $Sa \circ \delta_X \circ \eta^{\mathbb{T}}_{SX}$ on $(SX, Sa \circ \delta_X)$ in the category of T-semialgebras, i.e., algebras that do not necessarily satisfy the unit axiom. This computation is performed thoroughly in the proof of Theorem 4.1.

Example 2.11 (from [11]). The weak lifting of the powerset monad to the category of Eilenberg-Moore algebras for β

corresponding to the distributive law of Example 2.8 is the Vietoris monad.

Note that there also exists a notion of distributive law of a monad $\mathbb{S}$ over a functor T, along with the expected correspondence with functors $\widehat{T}$ extending T to $\mathbf{Kl}(\mathbb{S})$ and with monads $\widetilde{\mathbb{S}}$ lifting $\mathbb{S}$ to $\mathbf{Alg}(T)$, the category of algebras for the *functor* T. The *weak* setting is actually intermediate: it takes into account the multiplication of $\mathbb{T}$ but not its unit.

In the rest of this section, we focus on the case where $\mathbb{S} = \mathcal{P}$, the powerset monad on Set. In this specific case, the Kleisli category of $\mathcal{P}$ is the category Rel of sets and relations. There is a wealth of results (going back to Barr [1] and extensively used in [11]) pertaining to extending Set-functors to Rel. In this context weak pullbacks play a crucial role.

Definition 2.12. A functor is *weakly cartesian* if it preserves weak pullback squares. A natural transformation is *weakly cartesian* if the naturality squares are weak pullbacks.

Barr [1] showed that a Set-functor T can be extended to a 2-functor $\widetilde{T}$ on the category Rel if and only if T is weakly cartesian. That $\widetilde{T}$ is a 2-functor simply means that for two relations R, R' with $R \subseteq R'$ we have $\widetilde{T}R \subseteq \widetilde{T}R'$. Recall, for example from [18], the computation of the extension functor $\widetilde{T}$. On objects, i.e. for a set X, we have $\widetilde{T}X = TX$. Given a relation $R: X \nrightarrow Y$, we regard it as a span with projections denoted by $p: R \to X$ and $q: R \to Y$ and we define

$$\widetilde{T}(R) = \{(x, y) \in TX \times TY \mid \exists z \in TR. \qquad (2)$$
$$x = Tp(z) \text{ and } y = Tq(z)\}$$

Similarly, a natural transformation $\alpha: T \Rightarrow T'$ can be extended to a 2-natural transformation $\widetilde{\alpha}: \widetilde{T} \Rightarrow \widetilde{T}'$ if and only if α is weakly cartesian. Furthermore, in both cases the extension is unique. An immediate consequence of these facts is the following result, see [11, Corollary 16]:

Proposition 2.13 ([11]). *Let* $\mathbb{T} = (T, \eta^{\mathbb{T}}, \mu^{\mathbb{T}})$ *be a monad on* Set. *If* T *and* $\mu^{\mathbb{T}}$ *are weakly cartesian, then there exists a unique weak extension of* $\mathbb{T}$ *to* Rel *whose functor is a 2-functor and whose multiplication is 2-natural.*

Combining the above proposition with Theorem 2.10, we can define the notions of canonical weak lifting and canonical weak distributive law.

Definition 2.14. Assume $\mathbb{T}$ is a monad on Set such that T and $\mu^{\mathbb{T}}$ are weakly cartesian. The *canonical weak extension* of $\mathbb{T}$ to Rel is its unique extension given by Proposition 2.13. It corresponds to a weak lifting of $\mathcal{P}$ to $\mathbf{EM}(\mathbb{T})$, hereafter called the *canonical weak lifting* of $\mathbb{T}$, and to a weak distributive law $\delta : T\mathcal{P} \Rightarrow \mathcal{P}T$, hereafter called the *canonical weak distributive law* of $\mathcal{P}$ over $\mathbb{T}$.

3 The Canonical Weak Distributive Law $\mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}$

In this section we present the canonical weak distributive law $\mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}$. The route map we adopt is the following.

- extend the *functor* $\mathcal{D}$ to Rel
- compute the canonical distributive law of the powerset monad over the *functor* $\mathcal{D}$
- show that this law is a weak distributive law of the powerset monad over the *monad* $\mathcal{D}$ via Proposition 2.13 and a technical lemma that proves that $\mu^{\mathcal{D}}$ is weakly cartesian.

It is well known that the functor $\mathcal{D}$ is weakly cartesian. This was proved independently by Moss [22] and by de Vink and Rutten [8] using the max flow-min cut theorem of network theory. Hence, using Barr's results [1], we know that the functor $\mathcal{D}$ extends in a canonical way to a functor $\widetilde{\mathcal{D}}$ on the category Rel of sets and relations. On objects, $\widetilde{\mathcal{D}}X = \mathcal{D}X$. Given a relation $R: X \nrightarrow Y$, the relation $\widetilde{\mathcal{D}}R: \mathcal{D}X \nrightarrow \mathcal{D}Y$ is the relation whose graph is given by

$$\{(\varphi, \psi) \in \mathcal{D}X \times \mathcal{D}Y \mid \exists \Theta \in \mathcal{D}R.$$
$$\mathcal{D}\pi_1(\Theta) = \varphi \text{ and } \mathcal{D}\pi_2(\Theta) = \psi\}$$

Identifying Rel with the Kleisli category of the powerset monad $\mathcal{P}$, corresponding to the extension $\widetilde{\mathcal{D}}$, we obtain a distributive law δ of the monad $\mathcal{P}$ over the functor $\mathcal{D}$

$$\delta: \mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}.$$

This is computed by applying the extension $\widetilde{\mathcal{D}}$ to the identity on $\mathcal{P}X$ seen as a relation $\ni: \mathcal{P}X \nrightarrow X$, whose graph is $\{(A, x) \mid A \subseteq X, x \in A\}$. We can thus compute the relation $\widetilde{\mathcal{D}}(\ni): \mathcal{D}\mathcal{P}X \nrightarrow \mathcal{D}X$. This corresponds in turn to the map $\delta_X: \mathcal{D}\mathcal{P}X \to \mathcal{P}\mathcal{D}X$ defined by

$$\delta_X(\Phi) = \{\varphi \in \mathcal{D}X \mid \exists \Theta \in \mathcal{D}(\ni).$$
$$\forall A \in \mathcal{P}X, \Phi(A) = \sum_{x \in A} \Theta(A, x) \qquad (3)$$
$$\forall x \in X, \varphi(x) = \sum_{A \ni x} \Theta(A, x)\}$$

Lemma 3.1. *The natural transformation* δ *is defined on a distribution* $\Phi \in \mathcal{D}\mathcal{P}X$ *written as a formal series* $\Phi = \sum_{i=1}^{n} p_i A_i$ *with* $A_i \in \mathcal{P}X$ *and* $p_i \in (0, 1]$ *by*

$$\delta_X(\Phi) = \{\sum_{i=1}^{n} p_i \varphi_i \in \mathcal{D}X \mid \forall i.\varphi_i \in \mathcal{D}X \qquad (4)$$
$$\text{and } supp(\varphi_i) \subseteq A_i\}$$

Proof. Assume $\varphi \in \delta_X(\Phi)$ and $\Theta \in \mathcal{D}(\ni)$ satisfies the conditions from (3). For $x \in X$ we define $\varphi_i(x) = \Theta(A_i, x)/p_i$. One can verify that $\varphi_i \in \mathcal{D}X$, $supp(\varphi_i) \subseteq A_i$ and $\varphi = \sum_{i=1}^{n} p_i \varphi_i$. Conversely, assume $(\varphi_i)_{1 \leq i \leq n}$ are distributions on X satisfying the conditions from (4). For $1 \leq i \leq n$ and $x \in X$

we define $\Theta(A_i, x) = p_i\varphi_i(x)$. For $A \in \mathcal{P}X$ such that $A \notin \{A_1, \ldots, A_n\}$ and for $x \in X$ we set $\Theta(A, x) = 0$. One can verify that $\Theta \in \mathcal{D}(\ni)$ and that it satisfies the conditions from (3). $\square$

The remainder of this section is devoted to proving the next theorem:

Theorem 3.2. *The natural transformation* $\delta\colon \mathcal{DP} \Rightarrow \mathcal{PD}$ *is a* weak distributive law *of the monad* $\mathcal{P}$ *over the monad* $\mathcal{D}$.

Proof. Recall that δ was obtained via the extension $\widetilde{\mathcal{D}}$ of the *functor* $\mathcal{D}$ to Rel. Therefore, we already know that δ interacts well with the multiplication and unit of $\mathcal{P}$, since it comes from the canonical extension of the functor $\mathcal{D}$ to the Kleisli category of $\mathcal{P}$. It only remains to prove that the diagram corresponding to the interaction with $\mu^{\mathcal{D}}$ commutes.

Using the correspondence between weak extensions and weak distributive laws given in Theorem 2.10, it is enough to show that $\widetilde{\mathcal{D}}$ is a weak extension of the *monad* $\mathcal{D}$ to Rel. Using Proposition 2.13 we show that $\widetilde{\mathcal{D}}$ is in fact the canonical weak extension of $\mathcal{D}$ by proving that the multiplication $\mu^{\mathcal{D}}$ is weakly cartesian. To this end, we apply Lemma 3.3 below, in which we reduce the problem of showing that the naturality squares of $\mu^{\mathcal{D}}$ are weak pullbacks to finding a non-negative solution to a linear system. We show its existence by employing Farkas' Lemma, a key result from linear programming duality. $\square$

Lemma 3.3. *The multiplication* $\mu^{\mathcal{D}}$ *of* $\mathcal{D}$ *is* weakly cartesian.

Proof. For probability distributions we will use the formal series notation, e.g. given $\varphi \in \mathcal{D}X$ we will write $\varphi = \sum_{i=1}^{n} a_i x_i$ where the x_i are distinct, $supp(\varphi) = \{x_i : 1 \le i \le n\}$ and $\varphi(x_i) = a_i \in (0, 1]$. Let $f : X \to Y$ and consider the diagram:

$$\begin{array}{ccc} \mathcal{DD}X & \xrightarrow{\ \mathcal{DD}f\ } & \mathcal{DD}Y \\ {\scriptstyle \mu^{\mathcal{D}}_X}\downarrow & & \downarrow{\scriptstyle \mu^{\mathcal{D}}_Y} \\ \mathcal{D}X & \xrightarrow[\ \mathcal{D}f\ ]{} & \mathcal{D}Y \end{array}$$

As we are in the category Set, proving that the above diagram is a weak pullback amounts to show that for all $\Psi \in \mathcal{DD}Y$ and all $\varphi \in \mathcal{D}X$ such that $\mu^{\mathcal{D}}_Y(\Psi) = \mathcal{D}f(\varphi)$, there is a $\Phi \in \mathcal{DD}X$ such that $\mu^{\mathcal{D}}_X(\Phi) = \varphi$ and $\mathcal{DD}f(\Phi) = \Psi$. Let $\Psi = \sum_{j=1}^{m} B_j\psi_j \in \mathcal{DD}Y$ and $\varphi \in \mathcal{D}X$. We will provide $(\varphi_j)_{1 \le j \le m}$ in $\mathcal{D}X$ such that

$$\forall j \in \{1, \ldots, m\}, \mathcal{D}f(\varphi_j) = \psi_j \tag{5}$$

$$\mu^{\mathcal{D}}_X(\sum_{j=1}^{m} B_j\varphi_j) = \varphi \tag{6}$$

Assuming the existence of φ_j satisfying (5) and (6) above, we can define $\Phi = \sum_{j=1}^{m} B_j\varphi_j$. The notation is correct because if $\varphi_j = \varphi_i$, apply $\mathcal{D}f$ to get $\psi_j = \psi_i$ and so $i = j$. In addition,

we obtain

$$\mathcal{DD}f(\Phi) = \sum_{j=1}^{m} B_j\mathcal{D}f(\varphi_j) = \sum_{j=1}^{m} B_j\psi_j = \Psi$$

by condition (5), and $\mu^{\mathcal{D}}_X(\Phi) = \varphi$ by (6). So this would achieve the proof. It remains to prove the existence of the distributions φ_j.

The conditions (5) and (6) can be rewritten as a linear system with variables $\varphi_j(x)$ for $1 \le j \le m$ and $x \in supp(\varphi)$, for which we need to exhibit a non-negative solution.

$$\forall y \in Y, \forall j \in \{1, \ldots, m\}, \sum_{x \in X} \varphi_j(x)f_*(x, y) = \psi_j(y) \tag{7}$$

$$\forall x \in X, \sum_{j=1}^{m} B_j\varphi_j(x) = \varphi(x) \tag{8}$$

One can solve this system one y at a time. Indeed, let $y \in Y$ and consider only $f^{-1}(\{y\})$ to get the system

$$\forall j \in \{1, \ldots, m\}, \sum_{x \in f^{-1}(\{y\})} \varphi_j^y(x) = \psi_j(y) \tag{9}$$

$$\forall x \in f^{-1}(\{y\}), \sum_{j=1}^{m} B_j\varphi_j^y(x) = \varphi(x) \tag{10}$$

If this system has a non-negative solution $(\varphi_j^y)_{1 \le j \le m}$ for every $y \in Y$, then there is a global solution defined by $\varphi_j(x) = \varphi_j^y(x)$ if $x \in f^{-1}(\{y\})$. One must justify that $\varphi_j \in \mathcal{D}X$. First, the equation (10) above entails that $supp(\varphi_j) \subseteq supp(\varphi)$, hence $supp(\varphi_j)$ is finite. Moreover:

$$\sum_{x \in X} \varphi_j(x) = \sum_{y \in Y} \sum_{x \in f^{-1}(\{y\})} \varphi_j^y(x) = \sum_{y \in Y} \psi_j(y) = 1 \tag{11}$$

Let us fix the $y \in Y$ and work on the system above. Write $f^{-1}(\{y\}) \cap supp(\varphi) = \{x_1, \ldots, x_k\}$. The system can be rewritten

$$\forall j \in \{1, \ldots, m\}, \sum_{i=1}^{k} \varphi_j^y(x_i) = \psi_j(y) \tag{12}$$

$$\forall i \in \{1, \ldots, k\}, \sum_{j=1}^{m} B_j\varphi_j^y(x_i) = \varphi(x_i) \tag{13}$$

Consider the vectors

$$u = (u_{1,1}, \ldots, u_{m,1}, \ldots, u_{1,k}, \ldots, u_{m,k}) \tag{14}$$

$$= (\varphi_1^y(x_1), \ldots, \varphi_m^y(x_1), \ldots, \varphi_1^y(x_k), \ldots, \varphi_m^y(x_k)) \tag{15}$$

$$v = (v_1, \ldots v_{m+k}) = (\psi_1(y), \ldots, \psi_m(y), \varphi(x_1), \ldots \varphi(x_k)) \tag{16}$$

$$B = (B_1, \ldots, B_m) \tag{17}$$

Our aim is to prove the existence of a non-negative solution of the system $M_k u = v$, where

$$M_k = \begin{pmatrix} I_m & I_m & \dots & I_m \\ B & 0 & \dots & 0 \\ 0 & B & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & B \end{pmatrix} \in \mathcal{M}_{m+k,mk}(\mathbb{R}) \qquad (18)$$

Recall that equation $\mu_Y^{\mathcal{D}}(\Psi) = \mathcal{D}f(\varphi)$ holds. In particular on y we have $\sum_{j=1}^m B_j \psi_j(y) = \sum_{i=1}^k \varphi(x_i)$, which is the same as $\langle B_k^- \mid v \rangle = 0$ where $B_k^- = (B_1, ..., B_m, -1, ... -1)$ is of length $m+k$. The notation $\langle u \mid v \rangle$ stands for the usual scalar product of vectors.

To solve this efficiently we will invoke a lemma from Farkas [10].

Lemma 3.4. *Let $A \in \mathcal{M}_{p,q}(\mathbb{R})$ and $b \in \mathbb{R}^p$. Then exactly one of the following statements is true:*

- *There exists an $x \in \mathbb{R}^q$ such that $Ax = b$ and $x \geq 0$.*
- *There exists an $y \in \mathbb{R}^p$ such that $A^T y \geq 0$ and $b^T y < 0$.*

We use Farkas' lemma with $p = m + k$, $q = mk$, $A = M$ and $b = v$. Assume, towards a contradiction, that there exists an $y \in \mathbb{R}^{m+k}$ such that $M^T y \geq 0$ and $\langle v \mid y \rangle < 0$. Along with the fact that $\langle B_k^- \mid v \rangle = 0$, this yields the following system

$$\forall i \in \{1, ..., m\}, \forall j \in \{1, ..., k\}, y_i + B_i y_{m+j} \geq 0 \qquad (19)$$

$$\sum_{i=1}^{m+k} v_i y_i < 0 \qquad (20)$$

$$\sum_{i=1}^m B_i v_i = \sum_{j=1}^k v_{m+j} \qquad (21)$$

Recall that all B_i and v_i are positive. For a fixed $j \in \{1, ..., k\}$, sum equation $(19)_i$ times v_i for all $i \in \{1, ..., m\}$ to get

$$\sum_{i=1}^m v_i y_i + y_{m+j} \left[\sum_{i=1}^m B_i v_i \right] \geq 0 \qquad (22)$$

Using (21) and then (20) we obtain

$$y_{m+j} \left[\sum_{l=1}^k v_{m+l} \right] \geq -\sum_{i=1}^m v_i y_i > \sum_{l=1}^k v_{m+l} y_{m+l}$$

Let $S = \left(\sum_{l=1}^k v_{m+l} y_{m+l} \right) / \left(\sum_{l=1}^k v_{m+l} \right)$. We just proved that for all j we have $y_{m+j} > S$. Hence

$$\frac{v_{m+j} y_{m+j}}{\sum_{l=1}^k v_{m+l}} > S \frac{v_{m+j}}{\sum_{l=1}^k v_{m+l}} \qquad (23)$$

and summing these inequalities for all $j \in \{1, ..., k\}$ produces the contradiction $S > S$. Therefore, the first property of Farkas' lemma is true. This gives us a solution u to the equation $M_k u = v$ such that $u \geq 0$, i.e., this gives rise to a proper distribution. Hence $\mu^{\mathcal{D}}$ is weakly cartesian. $\square$

We end up this section with the remark concerning the troublesome unit $\eta^{\mathcal{D}}$. Indeed, this natural transformation is not weakly cartesian, as the next example shows. This is the

reason why the canonical extension of the monad $\mathcal{D}$ to Rel is only a weak extension.

Example 3.5. Consider the unique map $! : \{0, 1\} \to \{0\}$. Then the naturality square

$$\begin{array}{ccc} \{0, 1\} & \xrightarrow{\eta_{\{0,1\}}^{\mathcal{D}}} & \mathcal{D}(\{0, 1\}) \\ \downarrow{!} & & \downarrow{\mathcal{D}!} \\ \{0\} & \xrightarrow{\eta_{\{0\}}^{\mathcal{D}}} & \mathcal{D}(\{0\}) \end{array}$$

is not a weak pullback. Indeed, consider $\varphi = \frac{1}{2}0 + \frac{1}{2}1 \in \mathcal{D}(\{0, 1\})$. Then $\eta_{\{0\}}^{\mathcal{D}}(0) = \mathcal{D}!(\varphi)$, but φ is not in the image of $\eta_{\{0,1\}}^{\mathcal{D}}$.

4 The Weak Lifting of $\mathcal{P}$ is $\mathcal{P}_c$

This section aims at identifying the convex powerset monad $\mathcal{P}_c$ of Definition 2.4 as the canonical weak lifting of the powerset monad $\mathcal{P}$ to the category of convex algebras.

Using the correspondence between weak distributive laws, weak liftings and weak extensions (see [11]), we know that the canonical weak distributive law $\delta : \mathcal{DP} \Rightarrow \mathcal{PD}$ exhibited in Theorem 3.2 induces a weak lifting of the powerset monad $\mathcal{P}$ to $\mathbf{EM}(\mathcal{D})$. Garner [11, Proposition 13] also provides the recipe for computing this weak lifting when idempotents split in the base category and we will apply it in our situation.

Theorem 4.1. *The canonical weak lifting of the powerset monad to the category of convex algebras is the convex powerset monad.*

Proof. We will also consider the category of semi-algebras for the monad $\mathcal{D}$, that is, maps $a : \mathcal{D}X \to X$ which interact well with the multiplication $\mu^{\mathcal{D}}$, but not necessarily with $\eta^{\mathcal{D}}$. It is easy to see that the weak distributive law entails the existence of a strict lifting of $\mathcal{P}$ to the category of $\mathcal{D}$-semialgebras. For such a semi-algebra $a : \mathcal{D}X \to X$ the lifting is computed as in Beck's situation as the composite

$$\mathcal{DP}X \xrightarrow{\delta_X} \mathcal{PD}X \xrightarrow{\mathcal{P}a} \mathcal{P}X$$

We consider next a $\mathcal{D}$-algebra (X, a), that is, an object of $\mathbf{EM}(\mathcal{D})$. Of course this can also be seen as a $\mathcal{D}$-semialgebra. Following the proof of [11, Proposition 13] we know that the composite map

$$f = \mathcal{P}X \xrightarrow{\eta_{\mathcal{P}X}^{\mathcal{D}}} \mathcal{DP}X \xrightarrow{\delta_X} \mathcal{PD}X \xrightarrow{\mathcal{P}a} \mathcal{P}X$$

is an idempotent in the category of $\mathcal{D}$-semialgebras and that the weak lifting $\widetilde{\mathcal{P}}(X, a)$ can be computed by splitting this idempotent in the category of $\mathcal{D}$-semialgebras. In *loc. cit.* Garner shows that in this fashion we obtain an algebra.

To finish the proof it remains to compute the idempotent f and to split it. Given $A \in \mathcal{P}X$ we can write $\eta_{\mathcal{P}X}^{\mathcal{D}}(A)$ as the

formal series $1 \cdot A$. Hence, by Lemma 3.1 we obtain

$$(\delta_X \circ \eta^{\mathcal{D}}_{\mathcal{P}X})(A) = \{\varphi \in \mathcal{D}X \mid supp(\varphi) \subseteq A\}$$

Henceforth, our idempotent is the function $f : \mathcal{P}X \to \mathcal{P}X$ defined by

$$f(A) = \{a(\varphi) \mid \varphi \in \mathcal{D}X \text{ such that } supp(\varphi) \subseteq A\}.$$

One gets the $\mathcal{D}$-algebra $\widetilde{\mathcal{P}}(X, a)$ by splitting the idempotent f. Note that a fixpoint of f is an $A \in \mathcal{P}X$ that satisfies

$$\forall \sum_{i=1}^{n} p_i = 1, \forall x_1 \ldots x_n \in A, a\left(\sum_{i=1}^{n} p_i x_i\right) \in A \qquad (24)$$

$$\forall x \in A, \exists \varphi \in \mathcal{D}X, supp(\varphi) \subseteq A \text{ and } x = a(\varphi) \qquad (25)$$
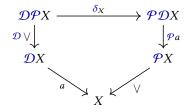
The second condition is trivial using $\eta^{\mathcal{D}}_X(x)$ and the projection axiom. Hence fixpoints of f are convex subsets of X w.r.t. (X, a) i.e. the carrier set of $\widetilde{\mathcal{P}}(X, a)$ is $\mathcal{P}_c X$. $\square$

Let us now consider the category of Eilenberg-Moore algebras for the monad $\mathcal{P}_c$ with the respective forgetful and free functors. The composite adjunction induces a monad structure on the functor $U^{\mathcal{D}} \mathcal{P}_c F^{\mathcal{D}}$.

$$\mathbf{EM}(\mathcal{P}_c) \underset{F^{\mathcal{P}_c}}{\overset{U^{\mathcal{P}_c}}{\rightleftarrows}} \mathbf{EM}(\mathcal{D}) \underset{F^{\mathcal{D}}}{\overset{U^{\mathcal{D}}}{\rightleftarrows}} \mathbf{Set}$$

We thus obtain, in a systematic way, exactly the monad $\mathcal{P}_c \mathcal{D}$ of convex subsets of finite distributions which was used in [20] and with mild variations in [6, 7, 12].

Using [11, Lemma 14] or [3, Proposition 3.7] we know that the category of Eilenberg-Moore algebras for the monad $\mathcal{P}_c$ is isomorphic to the category of Eilenberg-Moore algebras for $\mathcal{P}_c \mathcal{D}$ and that these algebras can be alternatively described as δ-algebras, that is sets X equipped with both a $\mathcal{P}$-algebra structure $\bigvee : \mathcal{P}X \to X$ and a $\mathcal{D}$-algebra structure $a : \mathcal{D}X \to X$, which are related by the distributive law δ:

This characterization gives an immediate concrete presentation for the $\mathcal{P}_c \mathcal{D}$-algebras. These are algebras

$$(X, \bigvee, (+_r)_{r \in [0,1]})$$

so that $(X, \bigvee)$ is a complete sup-semilattice, $(X, (+_r)_{r \in [0,1]})$ is a convex algebra, and the following distributivity axiom holds for $r \in (0, 1)$:

$$(\bigvee x_i) +_r y = \bigvee(x_i +_r y)$$

This is very similar to the algebraic theory of convex semilattices presented in [7]. However, the monad C considered in *loc. cit.* consists only in the *finitely generated* non-empty

convex sets of distributions, hence only a binary sup is used instead of $\bigvee$.

5 Generalized determinization via weak distributive law

If we forget about initial states, deterministic automata over a finite alphabet A can be identified with coalgebras for the functor $2 \times (-)^A$, while their non-deterministic counterpart can be seen as coalgebras for the functor $2 \times \mathcal{P}(-)^A$. The procedure of turning a non-deterministic automaton with states Q into a deterministic automaton with states $\mathcal{P}(Q)$ – also known as the powerset construction – can thus be seen as a way of transforming a coalgebra $c : Q \to 2 \times (\mathcal{P}Q)^A$ into a coalgebra $c^\# : \mathcal{P}Q \to 2 \times (\mathcal{P}Q)^A$. In [26] this was generalized to wider setting, considering on a given category $\mathbf{C}$ a functor F and a monad $\mathbb{T} = (T, \eta^{\mathbb{T}}, \mu^{\mathbb{T}})$, the former specifying the branching behavior of a given system. Whenever the functor F can be lifted to a functor $\widehat{F}$ on the Eilenberg-Moore category of $\mathbb{T}$, i.e., in the presence of a distributive law

$$TF \Rightarrow FT$$

of the functor F over the monad $\mathbb{T}$, then one can transform FT-coalgebras with carrier X into $\widehat{F}$-coalgebras with carrier TX. This was nicely summarized in [13] by stating that the free functor $F^{\mathbb{T}} : \mathbf{C} \to \mathbf{EM}(\mathbb{T})$ can be lifted to a functor between the categories of coalgebras. We thus have the following commuting diagram, where the vertical arrows denote the forgetful functors mapping an FT-coalgebra, respectively an $\widehat{F}$-coalgebra, to its carrier set, respectively $\mathbb{T}$-algebra.

$$(26)$$

The generalized powerset construction has many applications including 'determinization' of partial Mealy machines, automata with exceptions or with side effects, total subsequential transducers, see e.g., [26]. However, as stated in [6], the translation of non-deterministic probabilistic automata into belief-state transformers cannot be seen as an instance of the generalized powerset construction for the lack of a suitable distributive law $\mathcal{D}\mathcal{P} \Rightarrow \mathcal{P}\mathcal{D}$.
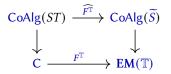
So we turn our attention to adapting the generalized powerset construction of [26] to the setting of weak distributive laws and then we will apply this framework to probabilistic automata. We will assume that idempotents split in $\mathbf{C}$ and we will consider a weak distributive law

$$\delta : TS \Rightarrow ST$$

of a monad $\mathbb{S} = (S, \eta^{\mathbb{S}}, \mu^{\mathbb{S}})$ over a monad $\mathbb{T} = (T, \eta^{\mathbb{T}}, \mu^{\mathbb{T}})$, but we could develop the same theory assuming just a functor S, instead of a monad $\mathbb{S}$. By [11], we know that such a weak distributive law corresponds to a weak lifting of $\mathbb{S}$ to the

category $\mathbf{EM}(\mathbb{T})$ of Eilenberg-Moore algebras for $\mathbb{T}$. Let us recall the adjunction $F^{\mathbb{T}} \dashv U^{\mathbb{T}} \colon \mathbf{EM}(\mathbb{T}) \to \mathbf{C}$.

Lemma 5.1. *Consider a weak distributive law $\delta \colon TS \Rightarrow ST$ of $\mathbb{S}$ over $\mathbb{T}$ and let $\widetilde{\mathbb{S}}$ be the corresponding weak lifting of $\mathbb{S}$ to $\mathbf{EM}(\mathbb{T})$.*

  i) Then, the free functor $F^{\mathbb{T}}$ lifts to a functor $\widehat{F^{\mathbb{T}}}$ transforming an ST-coalgebra into an $\widetilde{S}$-coalgebra.

$$\begin{array}{ccc} \mathsf{CoAlg}(ST) & \xrightarrow{\widehat{F^{\mathbb{T}}}} & \mathsf{CoAlg}(\widetilde{S}) \\ \downarrow & & \downarrow \\ \mathbf{C} & \xrightarrow{F^{\mathbb{T}}} & \mathbf{EM}(\mathbb{T}) \end{array}$$

  ii) The forgetful functor $U^{\mathbb{T}}$ lifts to a functor $\widehat{U^{\mathbb{T}}}$ transforming an $\widetilde{S}$-coalgebra into an S-coalgebra.

$$\begin{array}{ccc} \mathsf{CoAlg}(\widetilde{S}) & \xrightarrow{\widehat{U^{\mathbb{T}}}} & \mathsf{CoAlg}(S) \\ \downarrow & & \downarrow \\ \mathbf{EM}(\mathbb{T}) & \xrightarrow{U^{\mathbb{T}}} & \mathbf{C} \end{array}$$

Proof. i) Recall that we have a natural transformation

$$\pi \colon SU^{\mathbb{T}} \Rightarrow U^{\mathbb{T}}\widetilde{S}$$

as in Definition 2.9. Consider a coalgebra $c \colon X \to STX$ in $\mathsf{CoAlg}(ST)$. Upon writing $T = U^{\mathbb{T}}F^{\mathbb{T}}$ and composing with π, we obtain the composite mapping

$$X \xrightarrow{c} SU^{\mathbb{T}}F^{\mathbb{T}}X \xrightarrow{\pi_{F^{\mathbb{T}}X}} U^{\mathbb{T}}\widetilde{S}F^{\mathbb{T}}X \,.$$

Now consider the adjoint transpose $c^{\#}$ of $\pi_{F^{\mathbb{T}}X} \circ c$:

$$F^{\mathbb{T}}X \xrightarrow{c^{\#}} \widetilde{S}F^{\mathbb{T}}X$$

This is a coalgebra for the functor $\widetilde{S}$ with carrier the free $\mathbb{T}$-algebra on X. Clearly this construction is functorial, so we have obtained the desired lifting of $F^{\mathbb{T}}$.

ii) Recall from Definition 2.9 the natural transformation

$$\iota \colon U^{\mathbb{T}}\widetilde{S} \Rightarrow SU^{\mathbb{T}}$$

Consider a $\widetilde{S}$-coalgebra $c \colon X \to \widetilde{S}X$. Applying the forgetful functor $U^{\mathbb{T}}$ and composing with ι_X, we obtain the S-coalgebra:

$$U^{\mathbb{T}}X \xrightarrow{U^{\mathbb{T}}c} U^{\mathbb{T}}\widetilde{S}X \xrightarrow{\iota_X} SU^{\mathbb{T}}X$$

This construction is functorial and yields the lifting $\widehat{U^{\mathbb{T}}}$ of $U^{\mathbb{T}}$. □
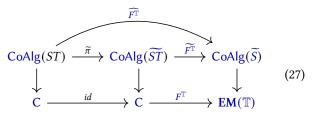
Notice that the lifting $\widehat{F^{\mathbb{T}}}$ defined above extends the construction in (26) to the setting of weak distributive laws, and when the weak distributive law is a strict distributive law in the sense of Beck, then we retrieve exactly the generalized powerset construction.

Analyzing the proof of Lemma 5.1, we notice that the functor $\widehat{F^{\mathbb{T}}}$ factors through the category of $\widetilde{ST}$-coalgebras.

Recall from [11] that $\widetilde{ST}$ is obtained as the monad induced by the composite adjunction

$$\mathbf{EM}(\widehat{\mathbb{S}}) \underset{F^{\widehat{\mathbb{S}}}}{\overset{U^{\widehat{\mathbb{S}}}}{\rightleftarrows}} \mathbf{EM}(\mathbb{T}) \underset{F^{\mathbb{T}}}{\overset{U^{\mathbb{T}}}{\rightleftarrows}} \mathbf{C}$$

hence $\widetilde{ST} = U^{\mathbb{T}}\widetilde{S}F^{\mathbb{T}}$. So the map $\pi_{F^{\mathbb{T}}X} \circ c$ obtained in the proof of Lemma 5.1 is a $\widetilde{ST}$-coalgebra. Thus the functor $\widehat{F^{\mathbb{T}}}$ factors through $\mathsf{CoAlg}(\widetilde{ST})$ as in the following diagram

$$\begin{array}{ccccc} & & \overset{\widehat{F^{\mathbb{T}}}}{\overgroup{\hspace{6cm}}} & & \\ \mathsf{CoAlg}(ST) & \xrightarrow{\widetilde{\pi}} & \mathsf{CoAlg}(\widetilde{ST}) & \xrightarrow{\widehat{F^{\mathbb{T}}}} & \mathsf{CoAlg}(\widetilde{S}) \\ \downarrow & & \downarrow & & \downarrow \\ \mathbf{C} & \xrightarrow{id} & \mathbf{C} & \xrightarrow{F^{\mathbb{T}}} & \mathbf{EM}(\mathbb{T}) \end{array} \quad (27)$$

Application to probabilistic automata. Let us instantiate this generic framework to the setting of probabilistic automata. For the sake of simplicity, we will assume that the input alphabet consists in only one letter, so we will identify probabilistic automata with coalgebras $c \colon X \to \mathcal{P}\mathcal{D}X$. The data in the first part of this section then instantiates as follows:

$\mathbb{T}$	finite distribution monad $\mathcal{D}$
$\mathbb{S}$	powerset monad $\mathcal{P}$
$\mathbf{EM}(\mathbb{T})$	convex algebras $(X, (+_r)_{r \in [0,1]})$
$\widetilde{\mathbb{S}}$	convex powerset monad $\mathcal{P}_c$
$\widetilde{ST}$	convex sets of distributions monad $\mathcal{P}_c\mathcal{D}$
π	taking the convex hull of a subset of a convex algebra
ι	inclusion, a convex subset of a convex algebra is, in particular, a subset of its carrier set

The functor $\widehat{F^{\mathbb{T}}}$ provided in Lemma 5.1 transforms a $\mathcal{P}\mathcal{D}$-coalgebra with a set of states X into a $\mathcal{P}_c$-coalgebra whose carrier is the convex algebra $\mathcal{D}X$ of finite distributions over X. This corresponds precisely to transforming a probabilistic automata into a belief-state transformer.

This transformation can be decomposed in two steps:

- First, for each state q of the automaton one takes the convex hull of the set of distributions $c(q)$. This corresponds to composing with the natural transformation π and one obtains a coalgebra for the functor $\mathcal{P}_c\mathcal{D}$.
- The second step consists in determinizing this coalgebra into a $\widetilde{\mathbb{S}}$-coalgebra. This is rather easy since $\mathcal{P}_c\mathcal{D}X$ carries a convex algebra structure.

Let us see how this construction compares to the generalized powerset construction of [6, Section 6]. The functor $\overline{\overline{\mathcal{F}}}$ which performs the generalized determinization in *loc. cit.* corresponds to the functor $\widehat{F^{\mathbb{T}}}$ in (27), hence only to the second step of our transformation. In [6] the convexification of a $\mathcal{P}\mathcal{D}$-coalgebra into a $\mathcal{P}_c\mathcal{D}$-coalgebra is performed by hand, via the functor $\mathcal{T}_{conv}$ of [6, Example 12]. The convex

hull natural transformation π, which is part of the definition of a weak lifting in our setting is called conv in *loc. cit.* and is not featured in the definition of a quasi-lax lifting [6, Definition 24]. So one advantage of our framework is that on the one hand, we identify in which sense the monad $\mathcal{P}_c$ constructed in [6] is canonical, and, on the other, we fully explain the determinization of a probabilistic automaton in one-go.

6 Conclusions

Summary of results. The counterexample provided in [30] for the existence of a distributive law $\mathcal{DP} \Rightarrow \mathcal{PD}$ already hints at the fact that the trouble lies with the unit $\eta^{\mathcal{D}}$ of the monad $\mathcal{D}$. In this paper we proved that the axiom involving this unit is indeed the only one which poses problems. We proved that we have a canonical weak extension of the monad $\mathcal{D}$ to the category Rel of sets and relations. This extension induces a canonical weak distributive law $\mathcal{DP} \Rightarrow \mathcal{PD}$, which in turn corresponds to a canonical weak lifting of the powerset monad to the category of Eilenberg-Moore algebras for $\mathcal{D}$ – that is– the category of convex algebras. We identify in this fashion the convex powerset monad $\mathcal{P}_c$ as the *canonical* weak lifting of the powerset monad to convex algebras.

The canonical weak distributive law also induces the composite monad $\mathcal{P}_c\mathcal{D}$: Set → Set mapping a set to the convex subsets of distributions on X. This monad and some of its variations were previously considered in the literature to model systems displaying both non-deterministic and probabilistic behavior [7, 12, 20]–and we exhibit here as a canonical monad for composing non-determinism and probabilistic choice.

Furthermore, we recover almost automatically a concrete presentation for the $\mathcal{P}_c\mathcal{D}$-algebras. This is very similar to the presentation provided in [7] for the monad of non-empty finitely generated convex subsets of distributions, which plays an important role in the trace semantics of nondeterministic probabilistic systems.

A more theoretic contribution is the extension of the generalized powerset construction of [26] to weak distributive laws. Applying this construction to the canonical weak distributive law $\mathcal{DP} \Rightarrow \mathcal{PD}$, we obtain a systematic way of translating a non-deterministic probabilistic system into a belief state transformer. We thus streamline and emphasize the canonicity of the results presented in [6].

Future research directions. A way of optimizing proofs of behavioral equivalences of systems and, more generally, of various coinductive properties is to employ the so-called up-to techniques [24]. These have been understood in a category theoretic setting via various forms of distributive laws in [4, 5]. This theory could be extended to the setting of weak distributive laws, and we can instantiate it to prove

the compatibility of the *up-to convex hull* technique of [6, Section 7].

Another question is whether we can obtain a weak distributive law of the powerset monad over the *countable* distribution monad similar to the one presented here.

Also, can we obtain the composite monad considered in [7] of *finitely generated* convex subsets of distributions from a weak distributive law?

Another line of research involves stepping away from the category of sets to various categories of domains. In particular we plan to investigate whether the power Kegelspitzen [16] can be obtained as a canonical weak lifting of a powerdomain.

We mentioned that the literature abounds of negative results of non-existence of strict distributive laws, see for example the 'no-go' theorems of [33]. Can we revisit these theorems? In which of these cases do we have various forms of weak distributive laws?

References

[1] Michael Barr. 1970. Relational algebras. In *Reports of the Midwest Category Seminar IV*. Springer, 39–55.

[2] Jon Beck. 1969. Distributive laws. In *Seminar on Triples and Categorical Homology Theory*, B. Eckmann (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 119–140.

[3] Gabriella Böhm. 2010. The weak theory of monads. *Advances in Mathematics* 225, 1 (2010), 1 – 32. https://doi.org/10.1016/j.aim.2010.02.015

[4] Filippo Bonchi, Daniela Petrisan, Damien Pous, and Jurriaan Rot. 2014. Coinduction up-to in a fibrational setting. In *CSL-LICS*. ACM, 20:1–20:9.

[5] Filippo Bonchi, Daniela Petrisan, Damien Pous, and Jurriaan Rot. 2017. A general account of coinduction up-to. *Acta Inf.* 54, 2 (2017), 127–190.

[6] Filippo Bonchi, Alexandra Silva, and Ana Sokolova. 2017. The Power of Convex Algebras. In *28th International Conference on Concurrency Theory, CONCUR 2017, September 5-8, 2017, Berlin, Germany*. 23:1–23:18. https://doi.org/10.4230/LIPIcs.CONCUR.2017.23

[7] Filippo Bonchi, Ana Sokolova, and Valeria Vignudelli. 2019. The Theory of Traces for Systems with Nondeterminism and Probability. In *34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019, Vancouver, BC, Canada, June 24-27, 2019*. 1–14. https://doi.org/10.1109/LICS.2019.8785673

[8] E. P. de Vink and J. J. M. M. Rutten. 1997. Bisimulation for probabilistic transition systems: A coalgebraic approach. In *Automata, Languages and Programming*, Pierpaolo Degano, Roberto Gorrieri, and Alberto Marchetti-Spaccamela (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 460–470.

[9] Ernst-Erich Doberkat. 2006. Eilenberg–Moore algebras for stochastic relations. *Information and Computation* 204, 12 (2006), 1756 – 1781. https://doi.org/10.1016/j.ic.2006.09.001

[10] J Farkas. 1902. Ober die Theorie der einfachen Ungleichungen. *J. Reine Angew. Math* 124 (1902), 1–24.

[11] Richard Garner. 2019. The Vietoris Monad and Weak Distributive Laws. *Applied Categorical Structures* (16 Oct 2019). https://doi.org/10.1007/s10485-019-09582-w

[12] Bart Jacobs. 2008. Coalgebraic Trace Semantics for Combined Possibilistic and Probabilistic Systems. *Electr. Notes Theor. Comput. Sci.* 203, 5 (2008), 131–152.

[13] Bart Jacobs, Alexandra Silva, and Ana Sokolova. 2015. Trace semantics via determinization. *J. Comput. System Sci.* 81, 5 (2015), 859 – 879.

https://doi.org/10.1016/j.jcss.2014.12.005 11th International Workshop on Coalgebraic Methods in Computer Science, CMCS 2012 (Selected Papers).

[14] C. Jones and Gordon D. Plotkin. 1989. A Probabilistic Powerdomain of Evaluations. In *LICS*. IEEE Computer Society, 186–195.

[15] Achim Jung and Regina Tix. 1998. The Troublesome Probabilistic Powerdomain. *Electronic Notes in Theoretical Computer Science* 13 (1998), 70 – 91. https://doi.org/10.1016/S1571-0661(05)80216-6 Comprox III, Third Workshop on Computation and Approximation.

[16] Klaus Keimel and Gordon D. Plotkin. 2017. Mixed powerdomains for probability and nondeterminism. *Logical Methods in Computer Science* Volume 13, Issue 1 (Jan. 2017). https://doi.org/10.23638/LMCS-13(1:2)2017

[17] Bartek Klin and Julian Salamanca. 2018. Iterated Covariant Powerset is not a Monad. *Electronic Notes in Theoretical Computer Science* 341 (2018), 261 – 276. https://doi.org/10.1016/j.entcs.2018.11.013 Proceedings of the Thirty-Fourth Conference on the Mathematical Foundations of Programming Semantics (MFPS XXXIV).

[18] Alexander Kurz and Jiri Velebil. 2016. Relation lifting, a survey. *J. Log. Algebr. Meth. Program.* 85, 4 (2016), 475–499.

[19] Saunders MacLane. 1971. *Categories for the Working Mathematician*. Springer-Verlag, New York. Graduate Texts in Mathematics, Vol. 5.

[20] Matteo Mio. 2014. Upper-Expectation Bisimilarity and Łukasiewicz μ-Calculus. In *Foundations of Software Science and Computation Structures*, Anca Muscholl (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 335–350.

[21] Michael Mislove. 2000. Nondeterminism and Probabilistic Choice: Obeying the Laws. In *CONCUR 2000 — Concurrency Theory*, Catuscia Palamidessi (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 350–365.

[22] Lawrence S. Moss. 1999. Coalgebraic logic. *Annals of Pure and Applied Logic* 96, 1 (1999), 277 – 317. https://doi.org/10.1016/S0168-0072(98)00042-6

[23] G. D. Plotkin. 1976. A powerdomain construction. *SIAM J. OF COMPUTING* (1976).

[24] Damien Pous and Davide Sangiorgi. 2011. *Enhancements of the bisimulation proof method*. Cambridge University Press, 233–289. https://doi.org/10.1017/CBO9780511792588.007

[25] N. Saheb-Djahromi. 1980. Cpo's of measures for nondeterminism. *Theoretical Computer Science* 12, 1 (1980), 19 – 37. https://doi.org/10.1016/0304-3975(80)90003-1

[26] Alexandra Silva, Filippo Bonchi, Marcello M. Bonsangue, and Jan J. M. M. Rutten. 2010. Generalizing the powerset construction, coalgebraically. In *FSTTCS (LIPIcs)*, Vol. 8. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 272–283.

[27] Ana Sokolova and Harald Woracek. 2015. Congruences of convex algebras. *Journal of Pure and Applied Algebra* 219, 8 (2015), 3110 – 3148. https://doi.org/10.1016/j.jpaa.2014.10.005

[28] Ross Street. 2009. Weak distributive laws. *Theory and Applications of Categories* 22, 12 (2009), 313–320.

[29] Regina Tix, Klaus Keimel, and Gordon Plotkin. 2009. Semantic Domains for Combining Probability and Non-Determinism. *Electronic Notes in Theoretical Computer Science* 222 (2009), 3 – 99. https://doi.org/10.1016/j.entcs.2009.01.002

[30] Daniele Varacca. 2003. *Probability, Nondeterminism and Concurrency: Two Denotational Models for Probabilistic Computation*. Technical Report. PhD thesis, Univ. Aarhus, 2003. BRICS Dissertation Series.

[31] Daniele Varacca and Glynn Winskel. 2006. Distributing probability over non-determinism. *Mathematical Structures in Computer Science* 16, 1 (2006), 87–113. https://doi.org/10.1017/S0960129505005074

[32] T. Šwirszcz. 1974. Monadic functors and convexity. *Bulletin de l'Académie Polonaise des Sciences, Ser. Sci. Math. Astronom. Phys.* 22 (1974), 39–42. https://www.fuw.edu.pl/~kostecki/scans/swirszcz1974.pdf

[33] M. Zwart and D. Marsden. 2019. No-Go Theorems for Distributive Laws. In *2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE Computer Society, Los Alamitos, CA, USA, 1–13. https://doi.org/10.1109/LICS.2019.8785707