States of Convex Sets

Bart Jacobs, Bas Westerbaan, and Bram Westerbaan

Institute for Computing and Information Sciences,
Radboud Universiteit Nijmegen, The Netherlands
{bart,bwesterb,awesterb}@cs.ru.nl

Abstract. State spaces in probabilistic and quantum computation are convex sets, that is, Eilenberg–Moore algebras of the distribution monad. This article studies some computationally relevant properties of convex sets. We introduce the term effectus for a category with suitable coproducts (so that predicates, as arrows of the shape $X \to 1 + 1$, form effect modules, and states, arrows of the shape $1 \to X$, form convex sets). One main result is that the category of *cancellative* convex sets is such an effectus. A second result says that the state functor is a "map of effecti". We also define 'normalisation of states' and show how this property is closed related to conditional probability. This is elaborated in an example of probabilistic Bayesian inference.

1 Introduction

The defining property of a convex set X is its closure under convex combinations. This means that for $x, y \in X$ and $\lambda \in [0, 1]$ the convex combination $\lambda x + (1-\lambda)y$ is also in X. There are some subtle properties that these convex combinations should satisfy, going back to Stone [Sto49]. Here we shall use a more abstract — but equivalent — categorical approach and call an Eilenberg–Moore algebra of the distribution monad $\mathcal{D}$ a convex set.

It is a basic fact that state spaces (*i.e.* sets of states) in probabilistic computation (both discrete and continuous) and in quantum computation are convex sets. Any serious model of such forms of computation will thus involve convex structures. It is within this line of research that the present paper contributes by clarifying several issues in the (computational) theory of convex sets. On a technical level the paper pinpoints (1) the relevance of a property of convex sets called 'cancellation', and (2) a 'normalisation' condition that is crucial for conditional probability and (Bayesian) inference.

These two points may seem strange and obscure. However, they play an important role in an ongoing project [Jac14] to determine the appropriate categorical axiomatisation for probabilistic and quantum logic and computation. Here we introduce the term 'effectus' for such a category. The main technical results of the paper can then be summarised as: the category **CConv** of **cancellative** convex sets is an effectus, and: the state functor $\mathrm{Stat}\colon \mathbf{B} \to \mathbf{CConv}$ from an arbitrary effectus $\mathbf{B}$ to **CConv** is a map of effecti. We illustrate how these results solidify the notion of effectus, and its associated state-and-effect triangle. We further show that conditional probability and (Bayesian) inference can be described both succinctly and generally via the idea of normalisation of stages.

Convex structures play an important role in mathematics (esp. functional analysis, see *e.g.* [AE80]), and in many application areas like economics. In the context of the axiomatisation of quantum (and probability) theory they are used systematically in for instance [Gud73] or [Fri09,BW11]. This paper fits in the latter line of research. It continues and refines [Jac14], by concentrating on the role of state spaces and their structure as convex sets.

The paper starts by describing background information on (discrete probability) distributions and convex sets. Coproducts $+$ of convex sets play an important role in the sequel, and are analysed in some detail. Subsequently, Section 3 concentrates on a well-known property of convex sets, known as cancellation. We recall how cancellation can be formulated in various ways, and show the equivalence with a joint monicity property that occurs in earlier work on categorical quantum axiomatisation [Jac14]. Section 4 introduces a categorical description of the well-known phenomenon of normalisation in probability. Finally, the resulting abstract description of conditional state in Section 6 is illustrated in a concrete example in Bayesian inference, using probability distributions as states.

2 Preliminaries on distributions and convex sets

For an arbitrary set X we write $\mathcal{D}(X)$ for the set of formal finite convex combinations of elements from X. These elements of $\mathcal{D}(X)$ will be represented in two equivalent ways.

- As formal convex sums $\lambda_1 |x_1\rangle + \cdots + \lambda_n |x_n\rangle$, for $x_i \in X$ and $\lambda_i \in [0,1]$ with $\sum_i \lambda_i = 1$. We use the 'ket' notation $|x\rangle$ in such formal sums to prevent confusion with elements $x \in X$.
- As functions $\varphi\colon X \to [0,1]$ with finite support and $\sum_x \varphi(x) = 1$. The support of φ is the set $\{\, x \in X \colon \varphi(x) \neq 0 \,\}$.

Elements of $\mathcal{D}(X)$ are also called (discrete probability) distributions over X.

The mapping $X \mapsto \mathcal{D}(X)$ can be made functorial: for $f\colon X \to Y$ we get a function $\mathcal{D}(f)\colon \mathcal{D}(X) \to \mathcal{D}(Y)$ which may be described in two equivalent ways:

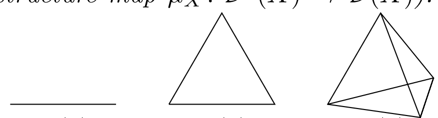
$$\mathcal{D}(f)\big(\textstyle\sum_i \lambda_i |x_i\rangle \big) = \sum_i \lambda_i |f(x_i)\rangle \qquad \text{or} \qquad \mathcal{D}(f)(\varphi)(y) = \sum_{x \in f^{-1}(y)} \varphi(x).$$

Moreover, $\mathcal{D}$ is a monad, with unit $\eta\colon X \to \mathcal{D}(X)$ given by $\eta(x) = 1|x\rangle$, and multiplication $\mu\colon \mathcal{D}^2(X) \to \mathcal{D}(X)$ by $\mu\big(\sum_i \lambda_i |\varphi\rangle \big)(x) = \sum_i \lambda_i \cdot \varphi_i(x)$. This monad is *monoidal* (or sometimes called *commutative*) from which the following result follows by general categorical reasoning (see [Koc71a,Koc71b]).

Proposition 1 *The category* **Conv** $= \mathcal{EM}(\mathcal{D})$ *of Eilenberg–Moore algebras is both complete and cocomplete, and it is symmetric monoidal closed. The tensor unit is the final singleton set* 1*, since* $\mathcal{D}(1) \cong 1$. $\square$

We recall that an Eilenberg–Moore algebra (of the monad $\mathcal{D}$) is a map of the form $\gamma\colon \mathcal{D}(X) \to X$ satisfying $\gamma \circ \eta = \mathrm{id}$ and $\gamma \circ \mu = \gamma \circ \mathcal{D}(\gamma)$. A morphism $\big(\mathcal{D}(X) \stackrel{\gamma}{\to} X\big) \longrightarrow \big(\mathcal{D}(X') \stackrel{\gamma'}{\to} X'\big)$ in $\mathcal{EM}(\mathcal{D})$ is a map $f\colon X \to X'$ with $f \circ \gamma = \gamma' \circ \mathcal{D}(f)$. An important point is that we identify an algebra with a convex set: the map $\gamma\colon \mathcal{D}(X) \to X$ turns a formal convex combination into an actual element in X. Maps of algebras preserve such convex sums and are commonly called *affine* functions. Therefore we often write **Conv** for the category $\mathcal{EM}(\mathcal{D})$.

Examples 2 *1. Let X be a set. The space $\mathcal{D}(X)$ of formal convex combinations over X is itself a convex set (with structure map $\mu_X\colon \mathcal{D}^2(X) \to \mathcal{D}(X)$). Given a natural number n the space $\mathcal{D}(n+1)$ is (isomorphic to) the n-th* simplex. *E.g., $\mathcal{D}(1)$ contains a single point, and $\mathcal{D}(2)$, $\mathcal{D}(3)$ and $\mathcal{D}(4)$ are pictured right.*



$\mathcal{D}(2)$ $\qquad$ $\mathcal{D}(3)$ $\qquad$ $\mathcal{D}(4)$

2. Any real vector space V is a convex set with structure map $\gamma\colon \mathcal{D}(V) \to V$ given by, $\gamma(\varphi) = \sum_{v \in V} \varphi(v) \cdot v$, for $\varphi \in \mathcal{D}(V)$. 3. Obviously a convex subset of a convex space is again a convex set. 4. A convex set which is isomorphic to a convex subset of a real vector space is called ***representable****. For every set X the space $\mathcal{D}(X)$ is representable since $\mathcal{D}(X)$ is a subset of the real vector space of functions from X to $\mathbb{R}$.*

In the remainder of this section we concentrate on coproducts of convex sets. Each category of algebras for a monad on **Sets** is cocomplete, by a theorem of Linton, see *e.g.* [BW85, § 9.3, Prop. 4]. This applies in particular to the category $\mathbf{Conv} = \mathcal{EM}(\mathcal{D})$, see Proposition 1. Hence we know that coproducts $+$ exist in **Conv**, but the problem is that the abstract construction of such coproducts of algebras uses a coequaliser in the category of algebras. Our aim is to get a more concrete description. We proceed by first describing the coproduct $X_\bullet = X + 1$ in **Conv**, where 1 is the final one-element convex set $1 = \{\bullet\}$.

Elements of this 'lift' $X_\bullet = X + 1$ can be thought of as being either λx for $\lambda(0,1]$ and $x \in X$, or the special element $\bullet$. This lift construction will be useful to construct the coproduct of convex sets later on.

Definition 3 *Let X be a convex set, via $\alpha\colon \mathcal{D}(X) \to X$. Define the set*

$$X_\bullet \;=\; \{\, (\lambda, x) \in [0,1] \times (\, X \cup \{\bullet\}\,) \colon\; \lambda = 0 \;\textit{ iff }\; x = \bullet \,\}.$$

We will often write $(0, e)$ even when e is an expression that does not make sense. In that case, by $(0, e)$ we mean $(0, \bullet)$. For example, $(0, \frac{1}{0}) = (0, \bullet)$. Given $(\lambda, x) \in X_\bullet$, we call λ the weight *of (λ, x) and denote it as $|(\lambda, x)| = \lambda$.*

Now, we may define a convex structure $\beta\colon \mathcal{D}(X_\bullet) \to X_\bullet$ succinctly:

$$\beta(\,\rho_1\,|(\lambda_1, x_1)\rangle \,+\, \cdots \,+\, \rho_n\,|(\lambda_n, x_n)\rangle\,) \;=\; (\,\zeta,\, \alpha(\tfrac{\rho_1\lambda_1}{\zeta}\,|x_1\rangle + \cdots + \tfrac{\rho_n\lambda_n}{\zeta}\,|x_n\rangle)\,),$$

where $\zeta = \lambda_1\rho_1 + \cdots + \lambda_n\rho_n$. Given an affine map $f\colon X \to Y$, define $f_\bullet\colon X_\bullet \to Y_\bullet$ by $f_\bullet(\lambda, x) = (\lambda, f(x))$ where $f(\bullet) := \bullet$.

Lemma 4 *This* $(X_\bullet, \beta)$ *is a convex set and it is the coproduct* $X+1$ *in* **Conv**.

Proof. The equation $\beta \circ \eta = \mathrm{id}$ is easy: for $(x, \lambda) \in X_\bullet$,

$$\beta(\eta(\lambda, x)) \;=\; \beta(\,|(\lambda, x)\rangle\,) \;=\; (\,\lambda,\, \alpha(\tfrac{\lambda}{\lambda}\,|x\rangle)\,) \;=\; (\,\lambda,\, \alpha(|x\rangle)\,) = (\lambda, x).$$

Verification of the μ-equation is left to the reader. There are obvious coprojections $\kappa_1 \colon X \to X_\bullet$ and $\kappa_2 \colon 1 \to X_\bullet$ given by $\kappa_1(x) = (1, x)$ and $\kappa_2(\bullet) = (0, \bullet)$. Given any convex set Y with $\gamma \colon \mathcal{D}(Y) \to Y$ together with affine maps $c_1 : X \to Y$ and $c_2 \colon 1 \to Y$, we can define a unique affine map $h \colon X_\bullet \to Y$ by $h(\lambda, x) = \gamma(\,\lambda\,|c_1(x)\rangle + (1-\lambda)\,|c_2(\bullet)\rangle\,)$. When $x = \bullet$ (and so $\lambda = 0$) we interpret $h(\lambda, x) = \gamma(|c_2(\bullet)\rangle)$. □

This lifted convex set $X_\bullet$ provides a simple description of coproducts.

Proposition 5 *The coproduct of two convex sets* X *and* Y *can be identified with the convex subset of* $X_\bullet \times Y_\bullet$ *of pairs whose weights sum to one. That is:*

$$X + Y \;\cong\; \{\; (x, y) \in X_\bullet \times Y_\bullet \colon\; |x| + |y| = 1 \;\}$$

The convex structure on this subset is inherited from the product $X_\bullet \times Y_\bullet$. *The first coprojection is given by* $\kappa_1(x) = \langle\,(1, x), (0, \bullet)\,\rangle$, *and there is a similar expression for* κ_2. *The cotuple is* $[f, g]((\lambda, x), (\rho, y)) = \lambda f(x) + \rho g(y)$. □

There is a similar description for the coproduct of n convex sets. E.g., for $n = 3$,

$$X + Y + Z \;=\; \{\; (x, y, z) \;\in\; X_\bullet \times Y_\bullet \times Z_\bullet \colon\; |x| + |y| + |z| = 1 \;\}.$$

From now on we shall use this concrete description for the coproduct $+$ in **Conv**. By the way, the initial object in **Conv** is simply the empty set, $\emptyset$.

3 The cancellation property for convex sets

The cancellation property that will be defined next plays an important role in the theory of convex sets. This section collects several equivalent descriptions from the literature, and adds one new equivalent property, expressed in terms of 'jointly monicity', see Theorem 8 (4) below. Crucially, this property is part of the axiomatisation proposed in [Jac14], and its equivalence to cancellation is the main contribution of this section.

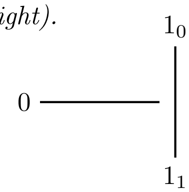
Definition 6 *Let* X *be a convex set. We call* X ***cancellative*** *provided that for all* $x, y_1, y_2 \in X$ *and* $\lambda \in [0, 1]$ *with* $\lambda \neq 1$ *we have*

$$\lambda x + (1-\lambda) y_1 \;=\; \lambda x + (1-\lambda) y_2 \qquad \Longrightarrow \qquad y_1 = y_2.$$

We write **CConv** $\hookrightarrow$ **Conv** *for the full subcategory of cancellative convex sets.*

Representable convex sets — subsets of real vector spaces — clearly satisfy this cancellation property. But not all convex sets do.

Examples 7 *1. If we remove from the unit interval* $[0,1]$ *the point* 1 *and replace it by a copy of the unit interval whose points we will denote by* 1_a *for* $a \in [0,1]$*, we get a convex space we will call* $\dashv$ *(pictured right).*



The convex structure on $\dashv$ *is such that the inclusion* $a \mapsto 1_a$ *is affine and the quotient* $\dashv \to [0,1]$ *which maps* 1_a *to* 1 *and* $[0,1)$ *on itself is affine.*
We have $\frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 1_0 = \frac{1}{2} = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 1_1$*, but* $1_0 \neq 1_1$*.*
Thus $\dashv$ *is not cancellative and hence not representable.*

2. A semilattice L *becomes a convex set if we define* $\sum_i \lambda_i x_i = \bigvee_i x_i$ *for all* $x_i \in L$ *and* $\lambda_i \in (0,1]$ *with* $\sum_i \lambda_i = 1$ *(see [Neu70], §4.5). The semilattice* L *is cancellative as convex set if and only if* $x = y$ *for all* $x, y \in L$*.*

Theorem 8 *For a convex set* X *the following statements are equivalent.*

1. *X is cancellative — see Definition 6;*
2. *X is representable, i.e. isomorphic to a convex subset of a real vector space;*
3. *X is* separated*, in the sense that for all* $x, y \in X$ *if* $f(x) = f(y)$ *for all affine maps* $f\colon X \to \mathbb{R}$*, then* $x = y$*;*
4. *The two maps* $[\kappa_1, \kappa_2, \kappa_2]$*,* $[\kappa_2, \kappa_1, \kappa_2]\colon X + X + X \to X + X$ *are jointly monic in* **Conv***.*

Proof. (3) $\Longrightarrow$ (2) Let $\mathrm{Aff}(X)$ denote the set of affine maps $X \to \mathbb{R}$, and V the vector space of (all) functions $\mathrm{Aff}(X) \to \mathbb{R}$, with pointwise structure. Let $\eta\colon X \to V$ be given by $\eta(x)(f) = f(x)$. We will prove that η is an injective affine map, making X representable.

Let $x_1, \ldots, x_N \in X$ and $\lambda_1, \ldots, \lambda_N \in [0,1]$ with $\sum_n \lambda_n = 1$ be given, and also $f \in \mathrm{Aff}(X)$ be given. Since f is affine, we get that η is affine too:

$$\begin{aligned} \eta(\, \lambda_1 x_1 + \cdots + \lambda_N x_N \,)(f) &= f(\, \lambda_1 x_1 + \cdots + \lambda_N x_N \,) \\ &= \lambda_1 f(x_1) + \cdots + \lambda_N f(x_N) \\ &= \lambda_1 \eta(x_1)(f) \,+\, \cdots \,+\, \lambda_N \eta(x_N)(f) \\ &= (\, \lambda_1 \eta(x_1) + \cdots + \lambda_N \eta(x_N) \,)(f). \end{aligned}$$

Towards injectivity of η, let $x, y \in X$ with $\eta(x) = \eta(y)$ be given. Then for each $f \in \mathrm{Aff}(X)$ we have $f(x) = \eta(x)(f) = \eta(y)(f) = f(y)$. Thus $x = y$ since X is separated.
(2) $\Longrightarrow$ (3) Since X is representable we may assume X is a convex subset of a real vector space V. Let $x, y \in X$ with $x \neq y$ be given. To show that X is separated we must find an affine map $f\colon X \to \mathbb{R}$ such that $f(x) \neq f(y)$.

Since $x \neq y$, we have that $x - y \neq 0$. By Zorn's lemma there is a maximal linearly independent set $\mathcal{B}$ which contains $x - y$. The set $\mathcal{B}$ spans V for if $v \in V$ is not in the span of $\mathcal{B}$ then $\mathcal{B} \cup \{v\}$ is a linearly independent set and $\mathcal{B}$ is not maximal. Thus $\mathcal{B}$ is a base for V. There is a unique linear map $f\colon \mathcal{B} \to \mathbb{R}$ such that $f(x - y) = 1$ and $f(b) = 0$ for all $b \in \mathcal{B}$ with $b \neq x - y$. Note that $f(x) \neq f(y)$. Let $g\colon X \to \mathbb{R}$ be the restriction of f to X. Then g is an affine map and $g(x) = f(x) \neq f(y) = g(y)$. Hence X is separated.

(2) $\Longrightarrow$ (1) is easy.
(1) $\Longrightarrow$ (2) We give an outline of the proof, but leave the key step to Stone (see [Sto49]). Let V be the real vector space of functions from X to $\mathbb{R}$ with finite support. Recall that $\mathcal{D}(X) = \{f \in V\colon \sum_{x \in X} f(x) = 1\}$. So we have a map $\eta_X\colon X \to \mathcal{D}(X) \subseteq V$. Let I be the linear span of

$$\{\ \eta_X(\gamma(f)) - f\colon\ f \in \mathcal{D}(X)\ \} \tag{1}$$

where $\gamma\colon \mathcal{D}(X) \to X$ is the structure map of X. Let $q\colon V \to V/I$ be the quotient map. Then by definition of I, the map $q \circ \eta_X\colon X \longrightarrow V$ is affine. So to show that X is representable it suffices to show that $q \circ \eta_X$ is injective. Let $x, y \in X$ with $q(\eta_X(x)) = q(\eta_X(y))$ be given. We must show that $x = y$. We have $f := \eta_X(x) - \eta_Y(y) \in I$. So f is a linear combination of elements from the set in (1). By the same syntactic argument as in the proof of Theorem 1 of [Sto49] we get that $f = 0$ since X is cancellative, and thus $x = y$.
(1) $\Longrightarrow$ (4) Write $\nabla_1 = [\kappa_1, \kappa_2, \kappa_2]$ and $\nabla_2 = [\kappa_2, \kappa_1, \kappa_2]$. We will prove that ∇_1 and ∇_2 are jointly injective (and thus jointly monic). Let $a, b \in X + X + X$ with $\nabla_1(a) = \nabla_1(b)$ and $\nabla_2(a) = \nabla_2(b)$ be given. We must show that $a = b$. Write $a \equiv (a_1, a_2, a_3)$ and $b \equiv (b_1, b_2, b_3)$ (see Proposition 5). Then we have

$$\nabla_1(a) \ = \ (\, a_1,\, a_2 \oplus a_3\,), \qquad \nabla_2(a) \ = \ (\, a_2,\, a_1 \oplus a_3\,), \tag{2}$$

where $\oplus$ is the partial binary operation on $X_\bullet$ given by

$$(\lambda, x) \oplus (\mu, y) \ = \ (\, \lambda + \mu,\ \tfrac{\lambda}{\lambda+\mu} x \ + \ \tfrac{\mu}{\lambda+\mu} y\,)$$

when $\lambda + \mu \leq 1$, and undefined otherwise. By the equalities from Statement (2) and similar equalities for $\nabla_1(b)$ and $\nabla_2(b)$, we get $a_1 = b_1$, $a_2 \oplus a_3 = b_2 \oplus b_3$, $a_2 = b_2$, and $a_1 \oplus a_3 = b_1 \oplus b_3$. It remains to be shown that $a_3 = b_3$. It is easy to see that $\oplus$ is cancellative since X is cancellative. Thus $a_1 \oplus a_3 = b_1 \oplus b_3$ and $a_1 = b_1$ give us that $b_1 = b_3$. Thus $a = b$.
(4) $\Longrightarrow$ (1) We assume that $\nabla_1,\ \nabla_2\colon X + X + X \to X + X$ (see above) are jointly monic and must prove that X is cancellative. The affine maps from 1 to $X + X + X$ correspond to the (actual) points of $X + X + X$, so it is not hard to see that ∇_1 and ∇_2 are jointly injective. Let $x_1, x_2, y \in X$ and $\lambda \in [0, 1]$ with $\lambda \neq 0$ and $\lambda x_1 + (1-\lambda) y = \lambda x_2 + (1-\lambda) y$ be given. We must show that $x_1 = x_2$.

Write $a_i = (\frac{\lambda}{2-\lambda}, x_i)$ and $b = (\frac{1-\lambda}{2-\lambda}, y)$ (where $i \in \{1, 2\}$). Then $a_i, b \in X_\bullet$. Further, $|a_i| + |b| + |b| = 1$, so $v_i := (b, b, a) \in X + X + X$. Note that

$$a_i \oplus b \ = \ (\ \tfrac{1}{2-\lambda},\ \lambda x_i + (1-\lambda) y\ ).$$

So we see that $a_1 \oplus b = a_2 \oplus b$. We have

$$\begin{aligned} \nabla_1(b, b, a_1) \ &= \ (b, b \oplus a_1) \ = \ (b, b \oplus a_2) \ = \ \nabla_1(b, b, a_2), \\ \nabla_2(b, b, a_1) \ &= \ (b, b \oplus a_1) \ = \ (b, b \oplus a_2) \ = \ \nabla_2(b, b, a_2). \end{aligned}$$

Since ∇_1,∇_2 are jointly injective this entails $a_1 = a_2$. Thus $x_1 = x_2$. $\square$

What we call (cancellative) convex sets appear under various different names in the literature. For instance, cancellative convex sets are called convex structures in [Gud77], convex sets in [Ś74], convex spaces of geometric type in [Fri09], and are the topic of the barycentric calculus of [Sto49]. Convex sets are called semiconvex sets in [Ś74,Flo81], and convex spaces in [Fri09]. The fact that every cancellative convex set is representable as a convex subset of a real vector space was proven by Stone, see Theorem 2 of [Sto49]. The description of convex sets as Eilenberg–Moore algebras is probably due to Świrszcz, see §4.1.3 of [Ś74] (see also [Jac10]), but the (quasi)variety of (cancellative) convex sets was already studied by Neumann [Neu70]. The fact that a convex set is cancellative iff it is separated by functionals was also noted by Gudder, see Theorem 3 of [Gud77]. The separation of points (and subsets) by a functional in a non-cancellative convex set has been studied in detail by Flood [Flo81]. The pathological convex set $\dashv$ (see Ex. 1) appears in [Fri09].

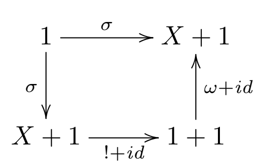
The duality of states and effects in quantum theory, see [HZ12], is formalised categorically in terms of an adjunction between 'effect modules' and convex sets. An effect module is a positive cancellative partial commutative monoid $(E, \ovee, 0)$ with a selected element 1 such that for all a there is a (unique) $a^\perp$ with $a \ovee a^\perp = 1$ *and* with a compatible action of $[0, 1]$. By **EMod**, we denote the category of effect modules with maps that preserve partial addition $\ovee$, scalar multiplication and 1. For details on effect modules we refer to [Jac14], but for the record we should note the following.

Proposition 9 *The adjunction* $\mathbf{EMod}^{op} \leftrightarrows \mathbf{Conv}$ *obtained by "homming into* $[0, 1]$*" restricts to an adjunction* $\mathbf{EMod}^{op} \leftrightarrows \mathbf{CConv}$. □

4 Normalisation

This section introduces a categorical description of normalisation, and illustrates what it means in several examples. As far as we know, this is new. Roughly, normalisation says that each non-zero substate can be written as a scalar product of a unique state.

Definition 10 *Let* $\mathbf{C}$ *be a category with finite coproducts* $(+, 0)$ *and a final object* 1. *We call maps* $1 \to X$ ***states*** *on* X, *and maps* $1 \to X + 1$ ***substates***. *We introduce the property* ***normalisation*** *as follows: for each substate* $\sigma \colon 1 \to X + 1$ *with* $\sigma \neq \kappa_2$ *there is a unique state* $\omega \colon 1 \to X$ *such that* $((\omega \circ !) + id) \circ \sigma = \sigma$. *That is, the diagram to the right commutes. The scalar involved is the map* $(! + \mathrm{id}) \circ \sigma \colon 1 \to 1 + 1$.



(The formulation of normalisation can be simplified a bit in the Kleisli category of the lift monad $(-) + 1$.*)*

Examples 11 *We briefly describe what normalisation means in several categories, and refer to [Jac14] for background information about these categories.*

1. *In the Kleisli category $\mathcal{K}\ell(\mathcal{D})$ of the distribution monad $\mathcal{D}$ a state $1 \to X$ is a distribution $\omega \in \mathcal{D}(X)$, and a substate $1 \to X + 1$ is a* subdistribution *$\sigma \in \mathcal{D}_{\leq 1}(X)$, for which $\sum_x \sigma(x) \leq 1$. If such a σ is not κ_2, that is, if $r = \sum_x \sigma(x) \in [0, 1]$ is not zero, take $\omega(x) = \frac{\varphi(x)}{r}$. Then $\sum_x \omega(x) = 1$.*
2. *Let $\mathbf{Cstar}_{PU}$ be the category of C^*-algebras with positive unital maps. We claim that normalisation holds in the opposite category $\mathbf{Cstar}_{PU}^{\mathsf{op}}$. The opposite is used in this context because C^*-algebras form a category of predicate transformers, corresponding to computations going in the reverse direction. In $\mathbf{Cstar}_{PU}^{\mathsf{op}}$ the complex numbers $\mathbb{C}$ are final, and coproducts are given by $\times$. Thus, let $\sigma \colon \mathscr{A} \times \mathbb{C} \to \mathbb{C}$ be a substate on a C^*-algebra $\mathscr{A}$. If σ is not the second projection, then $r := \sigma(1, 0) \in [0, 1]$ is non-zero. Hence we define $\omega \colon \mathscr{A} \to \mathbb{C}$ as $\omega(a) = \frac{\sigma(a,0)}{r}$. Clearly, ω is positive, linear and $\omega(1) = 1$. (In fact, substates $\mathscr{A} \times \mathbb{C} \to \mathbb{C}$ may be identified with* subunital *positive maps $\omega \colon \mathscr{A} \to \mathbb{C}$, for which $0 \leq \omega(1) \leq 1$. Normalisation rescales such a map ω to $\omega' := \frac{\omega(-)}{\omega(1)}$ with $\omega'(1) = 1$.)*
3. *The same argument can be used in the opposite category $\mathbf{EMod}^{\mathsf{op}}$ of effect modules. Hence $\mathbf{EMod}^{\mathsf{op}}$ also satisfies normalisation.*
4. *Normalisation holds both in* **Conv** *and in* **CConv***, that is, it holds for convex and for cancellative convex sets. This is easy to see using the description $X + 1 = X_\bullet$ from Lemma 4. Indeed, if $\sigma \colon 1 \to X_\bullet$ is not κ_2, then writing $\sigma(\bullet) \equiv (\lambda, a)$ we have $\lambda > 0$ and $a \neq \bullet$. Now take as state $\omega \colon 1 \to X$ with $\omega(1) = a$.*

In the present context we restrict ourselves to effect modules and convex sets over the unit interval $[0, 1]$, and not over some arbitrary effect monoid, like in [Jac14]. Normalisation holds for such effect modules over $[0, 1]$ because we can do division $\frac{s}{r}$ in $[0, 1]$, for $s \leq r$. More generally, it must be axiomatised in effect monoids. That is beyond the scope of the current article.

5 Effecti

The next definition refines the requirements from [Jac14] and introduces the name 'effectus' for the kind of category at hand. The main result is that taking the states of an arbitrary effects yields a functor to cancellative convex sets, which preserves coproducts. This leads to a robust notion, which is illustrated via the state-and-effect triangle associated with an effectus, which now consists of maps of effecti.

Definition 12 *A category* **C** *is called an* ***effectus*** *if:*

1. *it has a final object 1 and finite coproducts $(0, +)$;*
2. *the following diagrams are pullbacks;*

$$
\begin{array}{ccc}
A + X & \xrightarrow{id+g} & A + Y \\
{\scriptstyle f+id}\downarrow & & \downarrow{\scriptstyle f+id} \\
B + X & \xrightarrow[id+g]{} & B + Y
\end{array}
\qquad\qquad
\begin{array}{ccc}
Y & =\!=\!=\!= & Y \\
{\scriptstyle \kappa_1}\downarrow & & \downarrow{\scriptstyle \kappa_1} \\
Y + A & \xrightarrow[id+g]{} & Y + B
\end{array}
$$

3. the maps $[\kappa_1, \kappa_2, \kappa_2],\ [\kappa_2, \kappa_1, \kappa_2]\colon\ X+X+X \to X+X$ *are jointly monic.*

An ***effectus with normalisation*** *is an effectus in which normalisation holds — see Definition 10.*

The main examples of effecti with normalisation — see also Examples 11 — include the Kleisli category $\mathcal{K}\ell(\mathcal{D})$ of the distribution monad $\mathcal{D}$ for discrete probality, but also the Kleisli category $\mathcal{K}\ell(\mathcal{G})$ of the Giry monad for continuous probability (which we don't discuss here). In the quantum setting our main example is the opposite $\mathbf{Cstar}_{\mathrm{PU}}^{\mathsf{op}}$ of the category of C^*-algebras, with positive unital maps.

A **predicate** on an object X in an effectus is an arrow $X \to 1+1$. A **scalar** is an arrow $1 \to 1+1$. A **state** on X is an arrow $1 \to X$. We write $\mathrm{Pred}(X)$ and $\mathrm{Stat}(X)$ for the collections of predicates and states on X, so that the scalars are in $\mathrm{Pred}(1) = \mathrm{Stat}(1+1)$. We shall say that $\mathbf{C}$ is an effectus *over* $[0,1]$ if the set of scalars $\mathrm{Pred}(1)$ in $\mathbf{C}$ is (isomorphic to) $[0,1]$. This is the case in all previously mentioned effecti, see Examples 11. An n**-test** on X is a map $X \to n \cdot 1$, where $n \cdot 1$ is the n-fold copower $1 + \cdots + 1$.

This paper goes beyond [Jac14] in that it considers not only effecti but also their morphisms. This gives a new perspective, see the proposition about the predicate functor below.

Definition 13 *Let* $\mathbf{C}$, $\mathbf{D}$ *be two effecti. A* ***map of effecti*** $\mathbf{C} \to \mathbf{D}$ *is a functor that preserves the final object and the finite coproducts (and as a consequence, preserves the two pullbacks in Definition 12).*

The next result is proven in [Jac14], without using the terminology of effecti.

Proposition 14 *Let* $\mathbf{C}$ *be an effectus over* $[0,1]$*. The assignment* $X \mapsto \mathrm{Pred}(X)$ *forms a functor* $\mathrm{Pred}\colon \mathbf{C} \to \mathbf{EMod}^{op}$*. This functor is a map of effecti.* □

This motivates us to see if there is a corresponding result for states, *i.e.* whether the assignment $X \mapsto \mathrm{Stat}(X)$ is also a map of effecti. This is where the cancellation and normalisation properties come into play.

Proposition 15 *The category* **CConv** *of cancellative convex sets is an effectus with normalisation.*

Proof. It is clear that the one-point convex set 1 is cancellative. It is also easy to see using the description of the coproduct of convex sets from Proposition 5 that the coproduct in **Conv** of two cancellative convex sets is cancellative. So the coproducts + of **Conv** restrict to **CConv**.

Moreover, the jointly monic property holds in **CConv** by Theorem 8, and normalisation holds by Example 11 (4). What remains is showing that the two diagrams in Definition 12 are pullbacks in **CConv**. For this we use the representation of the coproduct of (cancellative) convex sets of Proposition 5.

To show that the diagram on the left in Definition 12 (2) is a pullback in **CConv** it suffices to show that it is a pullback in **Sets**, so let elements $(a, y) \in$

$A+Y$ and $(b,x) \in B+X$ with $(f+\mathrm{id})(a,y) = (\mathrm{id}+g)(b,x)$ be given. We must show that there is a unique $e \in A+X$ with the following property, called $P(e)$.

$$(\mathrm{id}+g)(e) \;=\; (a,y) \qquad \text{and} \qquad (f+\mathrm{id})(e) \;=\; (b,x) \qquad\qquad (P(e))$$

We claim that $P(a,x)$. For this we must first show that $(a,x) \in A+X$, that is, $|a|+|x| = 1$. Note that since $(f_\bullet(a), y) \equiv (f+\mathrm{id})(a,y) = (\mathrm{id}+g)(b,x) \equiv (b, g_\bullet(x))$ we have $f_\bullet(a) = b$ and $g_\bullet(x) = y$. Then $|a| = |f_\bullet(a)| = |b|$. Further, $|b|+|x| = 1$ since $(b,x) \in B+X$. Thus $|a|+|x| = 1$, and $(a,x) \in A+X$. Now, $(\mathrm{id}+g)(a,x) = (a, g_\bullet(x)) = (a,y)$, and similarly we have $(f+\mathrm{id})(a,x) = (b,x)$. Hence $P(a,x)$.

For uniqueness, suppose that $(a',x') \in A+X$ with $P(a',x')$ is given. We must show that $a = a'$ and $x = x'$. We have $(a,y) = (\mathrm{id}+g)(a',x') = (a', g_\bullet(x'))$ and similarly $(b,x) = (f_\bullet(a'), x')$. Thus $a' = a$ and $x = x'$. Hence the diagram on the left is pullback in **CConv**. A similar reasoning works for the diagram on the right in Definition 12. □

Proposition 16 *Let* **C** *be an effectus with normalisation over* $[0,1]$*. The state functor* $\mathrm{Stat}\colon \mathbf{C} \to \mathbf{Conv}$ *preserves coproducts:* $\mathrm{Stat}(X+Y) \cong \mathrm{Stat}(X)+\mathrm{Stat}(Y)$ *for* $X, Y \in \mathbf{C}$.

Proof. For objects $X, Y \in \mathbf{C}$, consider the canonical map:

$$\mathrm{Stat}(X)+\mathrm{Stat}(Y) \xrightarrow{\vartheta := [\,\mathrm{Stat}(\kappa_1),\, \mathrm{Stat}(\kappa_2)\,]} \mathrm{Stat}(X+Y)$$

We have to show that this ϑ is bijective. First, we give a direct expression for ϑ. Let $(x,y) \in \mathrm{Stat}(X)+\mathrm{Stat}(Y)$ be such that $|x|, |y| \in (0,1)$. Then there are a scalar $\lambda\colon 1 \to 1+1$ and states $\hat{x}\colon 1 \to X$ and $\hat{y}\colon 1 \to Y$ such that $(x,y) = \lambda\kappa_1(\hat{x}) + \lambda^\perp\kappa_2(\hat{y})$, where $\lambda^\perp = [\kappa_2,\kappa_1] \circ \lambda = 1-\lambda$. Observe $\vartheta(x,y) = (\hat{x}+\hat{y}) \circ \lambda$.

To prove surjectivity, let $\omega\colon 1 \to X+Y$ be a state. Define a scalar $\lambda = (!+!) \circ \omega\colon 1 \to 1+1$. Define substates $x = (\mathrm{id}+!) \circ \omega\colon 1 \to X+1$ and $y = [\kappa_2 \circ !, \kappa_1] \circ \omega\colon 1 \to Y+1$. For now, suppose that $\lambda \neq \kappa_1$ and $\lambda \neq \kappa_2$, i.e., $x \neq \kappa_2$ and $y \neq \kappa_2$. Then by normalisation, there are states $\hat{x}\colon 1 \to X$ and $\hat{y}\colon 1 \to Y$ such that

$$x = (\hat{x}+\mathrm{id}) \circ (!+\mathrm{id}) \circ x \qquad \text{and} \qquad y = (\hat{y}+\mathrm{id}) \circ (!+\mathrm{id}) \circ y.$$

Define $\sigma := \langle\, (\lambda, \hat{x}),\, (\lambda^\perp, \hat{y})\,\rangle \in \mathrm{Stat}(X)+\mathrm{Stat}(Y)$. We claim that $\vartheta(\sigma) = \omega$. That is, we must show that $(\hat{x}+\hat{y}) \circ \lambda = \omega$. Note that the two maps

$$(\mathrm{id}+!)\colon X+Y \to X+1 \qquad \text{and} \qquad [\kappa_2 \circ !, \kappa_1]\colon X+Y \to Y+1$$

are jointly monic in **C** by the pullback diagram on the left in Definition 12 (2). Thus it suffices to show that

$$\begin{aligned} &(\mathrm{id}+!) \circ (\hat{x}+\hat{y}) \circ \lambda \;=\; (\mathrm{id}+!) \circ \omega \;\equiv\; x \\ \text{and} \qquad &[\kappa_2 \circ !, \kappa_1] \circ (\hat{x}+\hat{y}) \circ \lambda \;=\; [\kappa_2 \circ !, \kappa_1] \circ \omega \;\equiv\; y \end{aligned}$$

We verify the first equality and leave the second equality to the reader.

$$\begin{aligned}(\mathrm{id}+!) \circ (\hat{x} + \hat{y}) \circ \lambda &= (\,(\mathrm{id} \circ \hat{x}) \,+\, (! \circ \hat{y})\,) \circ \lambda \\ &= (\hat{x} + \mathrm{id}) \circ \lambda \\ &= (\hat{x} + \mathrm{id}) \circ (! + \mathrm{id}) \circ (\mathrm{id}+!) \circ \omega && \text{by def. of } \lambda \\ &= (\hat{x} + \mathrm{id}) \circ (! + \mathrm{id}) \circ x && \text{by def. of } x \\ &= x && \text{by def. of } \hat{x}\end{aligned}$$

Suppose $\lambda = \kappa_2$, i.e., $x = \kappa_2$. Then $\lambda^\perp = \kappa_1$, so $y \neq \kappa_2$. Thus there is a unique $\hat{y}$ with $y = (\hat{y} + \mathrm{id}) \circ (! + \mathrm{id}) \circ y = (\hat{y} + \mathrm{id}) \circ \lambda^\perp = (\hat{y} + \mathrm{id}) \circ \kappa_1 = \kappa_1 \circ \hat{y}$. Thus:

$$\begin{aligned}(\mathrm{id}+!) \circ \kappa_2 \circ \hat{y} &= \kappa_2 \circ ! \circ \hat{y} = \kappa_2 = x = (\mathrm{id}+!) \circ \omega \\ [\kappa_2 \circ !, \kappa_1] \circ \kappa_2 \circ \hat{y} &= \kappa_1 \circ \hat{y} = y = [\kappa_2 \circ !, \kappa_1] \circ \omega.\end{aligned}$$

By joint monicity of $(\mathrm{id}+!)$ and $[\kappa_2 \circ !, \kappa_1]$ we derive $\omega = \kappa_2 \circ \hat{y} \equiv \vartheta(\kappa_1(\hat{y}))$. The case for $x = \kappa_1$ is similar. Thus ϑ is surjective.

For injectivity, let $(x, y), (x', y') \in \mathrm{Stat}(X){+}\mathrm{Stat}(Y)$ with $\vartheta(x, y) = \vartheta(x', y')$ be given. Note that $|x'| = (!+!) \circ \vartheta(x', y') = (!+!) \circ \vartheta(x, y) = |x|$. Assume that $|x| \in (0, 1)$. Then there are $\hat{x}, \hat{x}' \colon 1 \to X$ and $\hat{y}, \hat{y}' \colon 1 \to Y$ such that

$$x = (|x|, \hat{x}); \quad y = (|x|^\perp, \hat{y}); \quad x' = (|x|, \hat{x}') \quad \text{and} \quad y' = (|x|^\perp, \hat{y}').$$

Consequently:

$$\begin{aligned}(\hat{x} + \mathrm{id}) \circ |x| &= (\mathrm{id}+!) \circ (\hat{x} + \hat{y}) \circ |x| = (\mathrm{id}+!) \circ \vartheta(x, y) = (\mathrm{id}+!) \circ \vartheta(x', y') \\ &= (\mathrm{id}+!) \circ (\hat{x}' + \hat{y}') \circ |x| = (\hat{x}' + \mathrm{id}) \circ |x|.\end{aligned}$$

It follows that we have two 'normalisations' $\hat{x}, \hat{x}' \colon 1 \to X$ for the substate $\sigma = (\hat{x} + \mathrm{id}) \circ |x| = (\hat{x}' + \mathrm{id}) \circ |x| \colon 1 \to X + 1$:

$$(\hat{x} + \mathrm{id}) \circ (! + \mathrm{id}) \circ \sigma = (\hat{x} + \mathrm{id}) \circ |x| = (\hat{x}' + \mathrm{id}) \circ |x| = (\hat{x}' + \mathrm{id}) \circ (! + \mathrm{id}) \circ \sigma.$$

And thus by the uniqueness in the normalisation assumption, we conclude $\hat{x} = \hat{x}'$. Similarly, $\hat{y} = \hat{y}'$. Hence $(x, y) = (x', y')$. We leave it to the reader to show that $(x, y) = (x', y')$ when $|x| \in \{0, 1\}$. Thus ϑ is injective. □

This preservation of coproducts is an important property for an abstract account of conditional probability, see Section 6 for the discrete case. For C^*-algebras the above result takes the following concrete, familiar form: let ω be a state of the form $\omega \colon \mathscr{A} \times \mathscr{B} \to \mathbb{C}$ — so that ω is a map $1 \to \mathscr{A} + \mathscr{B}$ in $\mathbf{Cstar}_{\mathrm{PU}}^{\mathsf{op}}$. Take $\lambda = \omega(1, 0) \in [0, 1]$. If we exclude the border cases $\lambda = 0$ and $\lambda = 1$, then we can write ω as convex combination $\omega = \lambda(\omega_1 \circ \pi_1) + (1 - \lambda)(\omega_2 \circ \pi_2)$ for states $\omega_1 = \frac{\omega(-,0)}{\lambda} \colon \mathscr{A} \to \mathbb{C}$ and $\omega_2 = \frac{\omega(0,-)}{1-\lambda} \colon \mathscr{B} \to \mathbb{C}$.

Now we obtain the analogue of Proposition 14 for states.

Theorem 17 *Let* $\mathbf{C}$ *be an effectus with normalisation over* $[0, 1]$*. The assignment* $X \mapsto \mathrm{Stat}(X)$ *yields a functor* $\mathrm{Stat} \colon \mathbf{C} \to \mathbf{CConv}$*, which is a map of effecti.*

Proof. Most of this is already clear: the functor Stat preserves $+$ by Proposition 16. It sends the initial object $0 \in \mathbf{C}$ to the set $\mathrm{Stat}(0) = \mathrm{Hom}(1, 0)$. This set must be empty, because otherwise $1 \cong 0$, which trivialises $\mathbf{C}$ and makes it impossible that $\mathbf{C}$ has $[0, 1]$ as its scalars. Also, $\mathrm{Stat}(1) \cong 1$, since there is only one map $1 \to 1$.

What remains to be shown is that each convex set $\mathrm{Stat}(X)$ is cancellative. By Theorem 8 we are done if we can show that the following two maps are jointly monic in the category **Conv**.

$$\mathrm{Stat}(X) + \mathrm{Stat}(X) + \mathrm{Stat}(X) \underset{[\kappa_2,\kappa_1,\kappa_2]}{\overset{[\kappa_1,\kappa_2,\kappa_2]}{\rightrightarrows}} \mathrm{Stat}(X) + \mathrm{Stat}(X)$$

But since the functor $\mathrm{Stat}\colon \mathbf{C} \to \mathbf{Conv}$ preserves coproducts by Proposition 16 this is the same as joint monicity of the maps:

$$\mathrm{Stat}(X + X + X) \underset{\mathrm{Stat}([\kappa_2,\kappa_1,\kappa_2])}{\overset{\mathrm{Stat}([\kappa_1,\kappa_2,\kappa_2])}{\rightrightarrows}} \mathrm{Stat}(X + X)$$

Suppose we have two states $\omega, \omega' \in \mathrm{Stat}(X+X+X)$ with $\mathrm{Stat}([\kappa_1, \kappa_2, \kappa_2])(\omega) = \mathrm{Stat}([\kappa_1, \kappa_2, \kappa_2])(\omega')$ and $\mathrm{Stat}([\kappa_2, \kappa_1, \kappa_2])(\omega) = \mathrm{Stat}([\kappa_2, \kappa_1, \kappa_2])(\omega')$. This means that $\omega, \omega'\colon 1 \to X + X + X$ satisfy $[\kappa_1, \kappa_2, \kappa_2] \circ \omega = [\kappa_1, \kappa_2, \kappa_2] \circ \omega'$ and $[\kappa_2, \kappa_1, \kappa_2] \circ \omega = [\kappa_2, \kappa_1, \kappa_2] \circ \omega'$. By using the joint monicity property in $\mathbf{C}$, see Definition 12 (3), we obtain $\omega = \omega'$. □

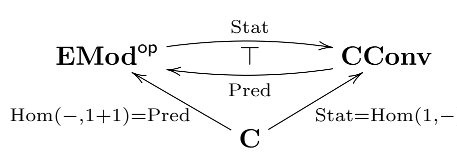
The following observation ties things closer together.

Proposition 18 *The adjunction* $\mathbf{EMod}^{op} \leftrightarrows \mathbf{CConv}$ *from Proposition 9 can be understood in terms of maps of effecti:*

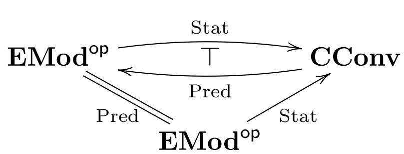
- *the one functor* $\mathbf{EMod}(-, [0, 1])\colon \mathbf{EMod}^{op} \to \mathbf{CConv}$ *is the states functor* $\mathrm{Stat} = \mathbf{EMod}^{op}(1, -)$*, since* $[0, 1]$ *is the initial effect module, and thus the final object* 1 *in* $\mathbf{EMod}^{op}$*;*
- *the other functor* $\mathbf{CConv}(-, [0, 1])\colon \mathbf{CConv} \to \mathbf{EMod}^{op}$ *is the predicate functor* $\mathrm{Pred} = \mathbf{CConv}(-, 1+1)$*, since the sum* $1+1$ *in* $\mathbf{CConv}$ *is* $[0, 1]$. □

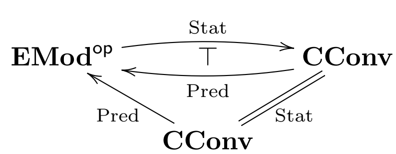
The above series of results culminates in the following.

Corollary 19 *Let* $\mathbf{C}$ *be an effectus over* $[0, 1]$*. Then we obtain a "state-and-effect" triangle shown on the right, where all the arrows are maps of effecti. (Arrows need not commute.)*



As degenerate cases of the triangle we obtain:





6 Conditional probability

An essential ingredient of conditional probability is normalisation, *i.e.* rescaling of probabilities: if we throw a dice, then the probability $P(4)$ of getting 4 is $\frac{1}{6}$. But the conditional probability $P(4 \mid \text{even})$ of getting 4 if we already know that the outcome is even, is $\frac{1}{3}$. This $\frac{1}{3}$ is obtained by rescaling of $\frac{1}{6}$, via division by the probability $\frac{1}{2}$ of obtaining an even outcome. Essentially this is the normalisation mechanism of Definition 10, and the resulting coproduct-preservation of the states functor from Proposition 16, as we will illustrate in the current section. Our general approach to conditional probability applies to both probabilistic and quantum systems. We present it in terms of an effectus with so-called 'instruments'. They are described in great detail in [Jac14], but here we repeat the essentials, for the Kleisli category $\mathcal{K}\ell(\mathcal{D})$ of the distribution monad $\mathcal{D}$. In a later, extended version of this paper the quantum case, using the effectus $\mathbf{Cstar}_{\mathrm{PU}}^{\mathsf{op}}$ of C^*-algebras will be included.

Let $\mathbf{C}$ be an arbitrary effectus. Recall its predicate functor $\mathrm{Pred}\colon \mathbf{C} \to \mathbf{EMod}^{\mathsf{op}}$ which takes the maps $X \to 1+1$ as predicates on X. In case $\mathbf{C} = \mathcal{K}\ell(\mathcal{D})$ we have $\mathrm{Pred}(X) = [0,1]^X$, the fuzzy predicates on X. An *n-test* in an effectus is an n-tuple of predicates $p_1, \ldots, p_n \in \mathrm{Pred}(X)$ with $p_1 \otimes \cdots \otimes p_n = 1$. In $\mathcal{K}\ell(\mathcal{D})$ this translates to predicates $p_i \in [0,1]^X$ with $\sum_i p_i(x) = 1$, for each $x \in X$. An *instrument* for an n-test $\overrightarrow{p}$ is a map $\mathrm{instr}_{\overrightarrow{p}}\colon X \to n \cdot X$ in $\mathbf{C}$, where $n \cdot X = X + \cdots + X$ is the n-fold coproduct. These instruments should satisfy certain requirements, but we skip them here. In $\mathcal{K}\ell(\mathcal{D})$ such an instrument is a map $\mathrm{instr}_{\overrightarrow{p}}\colon X \to \mathcal{D}(n \cdot X)$ defined as:

$$\mathrm{instr}_{\overrightarrow{p}}(x) \;=\; p_1(x)\,|\kappa_1 x\rangle + \cdots + p_n(x)\,|\kappa_n x\rangle\,.$$

We can now introduce the notion of conditional state, via coproduct-preservation.

Definition 20 *Let* $\mathbf{C}$ *be an effectus (over* $[0,1]$*) with normalisation, and with instruments as sketched above. Let* $\omega \in \mathrm{Stat}(X)$ *be a state, and* $\overrightarrow{p} = p_1, \ldots, p_n$ *be an n-test on* X*, of predicates* $p_i \in \mathrm{Pred}(X)$*. By applying the state functor* $\mathrm{Stat}\colon \mathbf{C} \to \mathbf{CConv}$ *we can form the new state:*

$$\omega' \;=\; \mathrm{Stat}\big(\mathrm{instr}_{\overrightarrow{p}}\big)(\omega) \;\in\; \mathrm{Stat}(n \cdot X) \;\overset{Prop.16}{\cong}\; n \cdot \mathrm{Stat}(X) \;\overset{Prop.5}{\subseteq}\; \textstyle\prod_n \mathrm{Stat}(X)\text{.}$$

Hence we write this new state ω' *as a convex combination of what we call* ***conditional*** *states on* X*, written as* $\omega|p_i \in \mathrm{Stat}(X)$*. The probabilities* r_i *in this convex combination can be computed as validity probabilities:*

$$r_i \;=\; \omega \models p_i \;=\; p_i \circ \omega\colon 1 \longrightarrow 1+1.$$

When each r_i *is non-zero, there are* n *such conditional states* $\omega|p_i$*.*

From a Bayesian perspective such a conditional state $\omega|p_i$ can be seen as an update of our state of knowledge, resulting from evidence p_i. This will be illustrated next in a discrete probabilistic example of Bayesian inference. It uses the Kleisli category $\mathcal{K}\ell(\mathcal{D})$ as effectus, in which a state $1 \to X$ in $\mathcal{K}\ell(\mathcal{D})$ corresponds to a distribution $\varphi \in \mathcal{D}(X)$. Conditional states, as defined above, appear as conditional distributions, generalising ordinary conditional probabilities.

Example 21 *Suppose, at an archaeological site, we are investigating a tomb of which we know that it must be from the second century AD, that is, somewhere from the time period 100 – 200. We wish to learn its origin more precisely. During excavation we are especially looking for three kinds of objects* $0, 1, 2$*, of which we know the time of use more precisely, in terms of "prior" distributions. This prior knowledge involves a split of the time period 100 – 200 into four equal subperiods A = 100 – 125, B = 125 – 150, C = 150 – 175, D = 175 – 200. Associated with each object* $i = 0, 1, 2$ *there is a predicate* $p_i \in [0, 1]^{\{A,B,C,D\}}$*, which we write as sequence of probabilities of the form:*

$$p_0 = [0.7, 0.5, 0.2, 0.1] \qquad p_1 = [0.2, 0.2, 0.1, 0.1] \qquad p_2 = [0.1, 0.3, 0.7, 0.8].$$

Predicate p_0 *incorporates the prior knowledge that object* 0 *is with probability* 0.7 *from subperiod* A*, with probability* 0.5 *from subperiod* B*, etc. Notice that these three predicates form a 3-test, since* $p_0 \ovee p_1 \ovee p_2 = 1$*. They can be described jointly as a Kleisli map* $\{A, B, C, D\} \to \mathcal{D}(\{0, 1, 2\})$.

Inference works as follows. Let our current knowledge about the subperiod of origin of the tomb be given as a distribution $\varphi \in \mathcal{D}(\{A, B, C, D\})$*. We can compute* $\varphi' = \mathrm{instr}_{\overrightarrow{p}}(\varphi) \in \mathcal{D}(3 \cdot \{A, B, C, D\})$ *and split* φ' *up into three conditional distributions* $\varphi|p_0, \varphi|p_1, \varphi|p_2 \in \mathcal{D}(\{A, B, C, D\})$*, like in Definition 20. If we find as "evidence" object* i*, then we update our knowledge from* φ *to* $\varphi|p_i$.

If we start from a uniform distribution, and find objects $i_1, \ldots, i_n \in \{0, 1, 2\}$*, then we have as inferred distribution (knowledge)* $\varphi|p_{i_1}|p_{i_2}|\cdots|p_{i_n}$*. For instance, the series of findings* $1, 2, 2, 0, 1, 1, 1, 1$ *yields the consecutive distributions shown in figure 1. Hence period B is most likely. These distributions are computed by a simple Python program that executes the steps of Definition 20. Interestingly, a change in the order of the objects that are found does not affect the final distribution. This is different in the quantum case, where such commutativity is lacking.*

$$0.25\,|A\rangle + 0.25\,|B\rangle + 0.25\,|C\rangle + 0.25\,|D\rangle$$
$$0.33\,|A\rangle + 0.33\,|B\rangle + 0.17\,|C\rangle + 0.17\,|D\rangle$$
$$0.09\,|A\rangle + 0.26\,|B\rangle + 0.30\,|C\rangle + 0.35\,|D\rangle$$
$$0.02\,|A\rangle + 0.14\,|B\rangle + 0.37\,|C\rangle + 0.48\,|D\rangle$$
$$0.05\,|A\rangle + 0.34\,|B\rangle + 0.37\,|C\rangle + 0.24\,|D\rangle$$
$$0.08\,|A\rangle + 0.49\,|B\rangle + 0.26\,|C\rangle + 0.17\,|D\rangle$$
$$0.10\,|A\rangle + 0.62\,|B\rangle + 0.17\,|C\rangle + 0.11\,|D\rangle$$
$$0.11\,|A\rangle + 0.72\,|B\rangle + 0.10\,|C\rangle + 0.06\,|D\rangle$$
$$0.12\,|A\rangle + 0.79\,|B\rangle + 0.05\,|C\rangle + 0.04\,|D\rangle$$

Fig. 1. *inferred distributions*

7 Conclusions

Starting from convex sets, in particular from the cancellation property and a concrete description of coproducts, we have arrived at the notion of effectus as a step towards a categorical axiomatisation of probabilistic and quantum computation. We have proven some 'closure' properties for effecti, among them that the states functor is a map of effecti. The concept of normalisation gave rise to a general notion of conditional state, which we have illustrated in the context of Bayesian inference.

Acknowledgements We would like to thank Kenta Cho and the anonymous referees for their constructive feedback.

References

[AE80] L. Asimow and A. Ellis. *Convexity Theory and its Applications in Functional Analysis*. Academic Press, New York, 1980.

[BW85] M. Barr and C. Wells. *Toposes, triples and theories*, volume 278. Springer-Verlag New York, 1985.

[BW11] H. Barnum and A. Wilce. Information processing in convex operational theories. In B. Coecke, I. Mackie, P. Panangaden, and P. Selinger, editors, *Proceedings of QPL/DCM 2008*, number 270(2) in Elect. Notes in Theor. Comp. Sci., pages 3–15. Elsevier, Amsterdam, 2011.

[Flo81] J. Flood. Semiconvex geometry. *Journal of the Australiam Mathematical Society*, 30:496–510, 1981.

[Fri09] T. Fritz. Convex spaces I: Definition and examples. *arXiv preprint arXiv:0903.5522*, 2009.

[Gud73] S. Gudder. Convex structures and operational quantum mechanics. *Communic. Math. Physics*, 29(3):249–264, 1973.

[Gud77] S. Gudder. Convexity and mixtures. *Siam Review*, 19(2):221–240, 1977.

[HZ12] T. Heinosaari and M. Ziman. The mathematical language of quantum theory: from uncertainty to entanglement. *AMC*, 10:12, 2012.

[Jac10] B. Jacobs. Convexity, duality and effects. In *Theoretical Computer Science*, pages 1–19. Springer, 2010.

[Jac14] B. Jacobs. New directions in categorical logic, for classical, probabilistic and quantum logic. *arXiv preprint arXiv:1205.3940v3*, 2014.

[Koc71a] A. Kock. Bilinearity and cartesian closed monads. *Mathematica Scandinavica*, 29:161–174, 1971.

[Koc71b] A. Kock. Closed categories generated by commutative monads. *Journal of the Australian Mathematical Society*, 12(04):405–424, 1971.

[Neu70] Walter D. Neumann. On the quasivariety of convex subsets of affine spaces. *Archiv der Mathematik*, 21(1):11–16, 1970.

[Ś74] T. Świrszcz. *Monadic functors and categories of convex sets.* PhD thesis, Polish Academy of Sciences, 1974.

[Sto49] M. Stone. Postulates for the barycentric calculus. *Annali di Matematica Pura ed Applicata*, 29(1):25–30, 1949.