Picturing classical and quantum Bayesian inference

Bob Coecke and Robert W. Spekkens

Oxford University Computing Laboratory and Perimeter Institute

Abstract

We introduce a graphical framework for Bayesian inference that is sufficiently general to accommodate not just the standard case but also recent proposals for a theory of quantum Bayesian inference wherein one considers mixed quantum states rather than probability distributions as representative of degrees of belief. The diagrammatic framework is stated in the graphical language of symmetric monoidal categories and of compact structures and Frobenius structures therein, in which Bayesian inversion boils down to transposition with respect to an appropriate compact structure. In the case of quantum-like calculi, the latter will be non-commutative. We identify a graphical property that characterizes classical Bayesian inference. The abstract classical Bayesian graphical calculi also allow to model relations among classical entropies, and reason about these. We generalize conditional independence to this very general setting and also generalize some standard results. Finally, given any dagger compact category, we construct a 'quantum-like' theory of inference. This result is of importance in the light of an existing completeness theorem for dagger compact categories.

Contents

1 Introduction

In this paper we introduce a graphical calculus and corresponding axiomatics in terms of monoidal categories for a very general notion of Bayesian inference. It enables one to reason at a highly abstract level, about theories more general than 'classical' Bayesian inference, including earlier proposals for quantum Bayesian inference by Leifer [27] and by Leifer and Poulin [28]. The graphical language exploits the two-dimensional diagrammatic representation to distinguish givens and conclusions. Bayesian inversion is diagrammatic transposition in terms of the *compact structures* [22, 23]. *Frobenius structures* [9] will be our vehicle for expressing notions such as *conditionalization* and relations of *conditional independence*. 'Classical' Bayesian inference is characterized in terms of a condition of commutativity for the Frobenius structure and therefore this structure is key to expressing *Bayesian updating* in the specific case of classical Bayesian inference.

An abstract representation of Bayesian inference allows one to identify which aspects of the standard probability calculus are merely conventional. For instance, in the context of R. T. Cox's derivation of the rules of classical Bayesian inference [16], the standard assumption that one's degree of belief about a proposition a ought to be represented by a number $p(a)$ between 0 and 1 and that one *multiplies* a conditional probability with a marginal to get the joint probability, i.e. $p(a,b) = p(a|b)p(b)$ is seen to be a consequence of a choice of convention. One could equally well represent this degree of belief by any bijective function of $p(a$ such as $s(a) = -\log p(a)$, in which case $s(a,b) = s(a|b) + s(b)$ and one replaces the standard form of Bayes' rule, $p(a|b) = p(b|a)p(a)/p(b)$, with its "entropic" form $s(a|b) = s(b|a) + s(a) - s(b)$. The abstract approach taken in this work finds a similar result and thereby contributes to the project of extracting the elements of Bayesian inference that are independent of convention.

Our graphical representation of Bayesian inference is also likely to have a close connection with the theory of Bayesian networks, and therefore may shed light on quantum analogues of these [28]. This has practical interest in the field of quantum information theory as quantum analogues of belief propagation algorithms are a natural avenue to quantum error correction schemes. As an example of this connection, the quantum analogue of Bayes' rule has the same form as the approximate reversal channel of Barnum and Knill [6]. Furthermore, given that Bayesian networks provide a powerful tool for inferring something about the causal relations that hold among propositions from the relations of conditional independence that exist in their correlations [32], we also hope that our graphical calculus might ultimately help to infer causal relations from quantum correlations and shed light on the quantum violation of Bell's notion of local causality.

Finally, there has been a great deal of interest recently about general probabilistic theories that are distinct from both classical probability theory and quantum theory, e.g. [8, 4, 7]. By considering a broad landscape of theories, one can hope to identify which aspects of quantum theory are shared with all operational probabilistic theories and which are unique to it. The framework we develop here provides a novel way of attacking this problem. By considering quantum theory as a theory of Bayesian inference, one is led to question which aspects of the theory are shared by all theories of Bayesian inference (insofar as one can define such a set) and which are unique to it.

The logic of categorical graphical languages. A pedestrian introduction to the graphical calculi for symmetric monoidal categories is in [12] and a comprehensive survey on these kinds of results is in [36]. These graphical languages trace back to Penrose's work in the early 70's.

Compact categories show up in a range of areas of mathematical physics including knot theory and the Temperley-Lieb algebra [38, 41] and the theory of quantum groups [37]. *Dagger compact categories* have recently been exploited by Abramsky and Coecke in quantum information theory [1] and in proposals for quantum gravity [2]. Frobenius structures trace back to Ferdinand Georg Frobenius' work on the representation theory of finite groups. They provide a very concise presentation of topological quantum field theories [3, 24], they provide a bridge between classical and linear logic [30], and allow diagrammatic axiomatization of quantum observables and C*-algebras [15, 40]. Similarly, they allow to distinguish between classical and quantum states [13].

To know how much one can actually prove in a diagrammatic language one relies on the correspondence between graphical languages and certain kinds of monoidal categories, for example:

Theorem 1.1 (Joyal-Street 1991 [21])**.** *An equation follows from the axioms of symmetric monoidal categories if and only if it can be derived in the corresponding graphical language.*

Theorem 1.2 (Kelly-Laplaza 1980, Selinger 2007 [23, 34])**.** *An equation follows from the axioms of (dagger) compact categories if and only if it can be derived in the corresponding graphical language.*

If one knows to which categorical structure a certain graphical calculus corresponds, one may ask the question wether there exist complete models of these.[1] We are aware of two results of this nature:

Theorem 1.3 (Hasegawa-Hofmann-Plotkin 2008 [19])**.** *An equation follows from the axioms of traced monoidal categories if and only if it holds in finite dimensional vector spaces.*

Theorem 1.4 (Selinger 2010 [35])**.** *An equation follows from the axioms of dagger compact categories if and only if it holds in finite dimensional Hilbert spaces.*

Theorem 1.4 is a highly surprising and powerful theorem which states that an important set of equational statements in quantum theory holds if and only if it can be derived in the graphical calculus. This result is moreover not only relevant for quantum mechanics related theories, but also classical probabilistic ones, since the latter can be represented in the category of Hilbert spaces, linear maps and the tensor product by means of Frobenius structures [13]. Unfortunately, there are no completeness result yet of the above kind directly involving Frobenius structures.

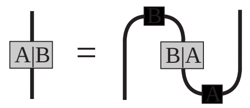
Results like Theorem 1.4 are obviously also important in the context of automated reasoning, and important steps towards automated reasoning with compact structures and Frobenius structures have already been made [17, 18]. The developments in this paper make these tools available to the study of (generalized) Bayesian inference.

Given the importance of dagger compact categories in the light of Theorem 1.4, we construct a class of theories of quantum-like Bayesian inference, one for each dagger compact category, for which the concrete non-commutative Frobenius structures arise from the underlying commutative compact structures.

Structure of the paper. In Section 2 we review compact structures, compact categories and dagger compact categories, dagger Frobenius structures therein, the interaction of the latter with compact structures, and the graphical calculus of all of these. In Section 3 we define general Bayesian graphical calculi, and also define the restricted case of classical Bayesian graphical calculi. We provide an example of a classical Bayesian graphical calculus and show that it is in fact canonical. We show how entropies provide a model of classical Bayesian graphical calculus. We also provide an example of

[1] That is, which enable to embed the corresponding free such categories, and hence, which are such that an equational statement holds in all models if and only if it is a consequence of the axioms of the categorical structure.

a non-classical one, namely the one introduced by Leifer [27, 28], to illustrate the generality of the framework. Typically, while for classical Bayesian graphical calculi the Frobenius multiplications will act commutatively, for non-classical ones it will act non-commutatively. We also observe that the key structural component of Bayesian graphical calculi, the Frobenius comultiplication, is in fact a logical broadcasting operation (a map from one object to a pair of these such that the final state has both marginals equal to the initial state). Section 4 relies on compact structures to obtain a graphical presentation where givens are inputs and conclusions are outputs. In this setting, Bayesian inversion is nothing but transposition, and looks as follows:



In Section 5 we define and study the important concept of conditional independence. We provide a simple example of how an assumption of conditional independence leads to a generalized formula for pooling multiple states of belief, and briefly discuss the semi-graphoid axioms and the connection between our work and Bayesian networks. Finally, in Section 6, in contrast to Section 3 where Bayesian calculi are defined axiomatically, we provide an explicit construction for any dagger compact category of a non-commutative Frobenius multiplication, of which composition of density operators is a special case. This also gives an explicit presentation of the non-completely positive Frobenius comultiplication which plays the role of a logical broadcasting operation.

2 Background: dagger Frobenius structures and compact structures

In this paper we work within the graphical language of symmetric monoidal categories (SMCs). In such a graphical calculus associativity and unit natural isomorphisms are always strict, that is:

$$(A \otimes B) \otimes C = A \otimes (B \otimes C) \qquad \text{and} \qquad A \otimes \mathrm{I} = A = \mathrm{I} \otimes A\,. \tag{1}$$

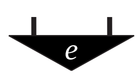
General morphisms (or operations) $f : A \to B$, which we interpret as 'processes', are represented as:

(2)

while points (or elements) $e : \mathrm{I} \to A$, which we interpret as 'states', are represented as:

(3)

To emphasize that a state $e : \mathrm{I} \to A \otimes B$ is bipartite we represent it as:

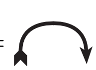


(4)

Composition and tensoring are respectively represented as:

(5)

A *compact structure* on an object A consists of another object A^* together with a pair of morphisms:

$\eta_A =$  $: \mathrm{I} \to A^* \otimes A$ $\qquad$ $\epsilon_A =$  $: A \otimes A^* \to \mathrm{I}\,,$ (6)

sometimes referred to as 'cups' and 'caps', which satisfy the 'yanking' equations:

 (7)

Hence we depict A by an upward arrow and A^* by a downward one. We call A^* the *dual* of A.

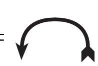
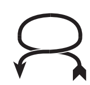
A category $\mathbf{C}$ is a *compact category* (CC) [22, 23] if each object comes with a compact structure, which interact in a coherent manner [22, 23]. For a number of reasons, including 'planarity' of the graphical representation of compact structures on compound objects, one usually adopts the convention that duals are (strictly) contravariant with respect to the tensor, that is,

$$(A \otimes B)^* = B^* \otimes A^* \qquad \text{and} \qquad \mathrm{I}^* = \mathrm{I}\,. \tag{8}$$

Cups and caps on a compound object $A \otimes B$ then become:

 (9)

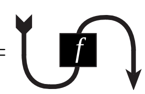
When we moreover have that $A^{**} = A$ then the direction of arrows clearly distinguishes between 'no $*$' and '$*$'. In this case, coherence[2] requires us to set

$\eta_{A^*} =$  $=$  $= \sigma_{A^*,A} \circ \eta_A \qquad \epsilon_{A^*} =$  $=$  $= \epsilon_A \circ \sigma_{A^*,A}\,,$ (10)

where $\sigma_{A,B} : A \otimes B \to B \otimes A$ is the morphism that simply swaps the objects A and B. We refer to such a CC as *strict*. In this paper all CCs will be strict.

Remark 2.1. The over/under-crossings of wires in the pictures have no formal meaning (cf. braiding), but only serve to make pictures more readable.

In any CC each morphism $f : A \to B$ has a *transpose*

$f^T := (1_{A^*} \otimes \epsilon_B) \circ (1_{A^*} \otimes f \otimes 1_{B^*}) \circ (\eta_A \otimes 1_{B^*}) =$  $: B^* \to A^*\,.$ (11)

Contravariance of $(-)^*$ on objects implies that

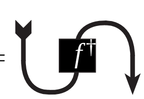
$$(f \otimes g)^T = g^T \otimes f^T\,. \tag{12}$$

A CC is a *dagger compact category* (dCC) [1, 34] if it comes with a contravariant dagger functor

$$(-)^\dagger : \mathbf{C}^{op} \to \mathbf{C}$$

that coherently preserves the compact structure. An ordinary category that comes with such a functor is called a *dagger category* (dC).

We call the composite of the transposed and the dagger the *conjugate*. Explicitly, for a morphism $f : A \to B$ its conjugate is

$\bar{f} := (f^\dagger)^T = (1_{B^*} \otimes \epsilon_A) \circ (1_{B^*} \otimes f^\dagger \otimes 1_{A^*}) \circ (\eta_B \otimes 1_{A^*}) =$  $: A^* \to B^*\,.$ (13)

[2] This means that structural morphisms of the same type are equal. Here that is, if by means of composing and tensoring symmetry, (identities), cups and caps one can obtain morphisms $f, g : A \to B$ of the same type, then these have to be equal.

Coherence now requires that $\bar{\eta}_A = \eta_A$ and $\bar{\epsilon}_A = \epsilon_A$. In the graphical calculus this condition can be derived from the interpretation of the dagger as flipping pictures upside-down.

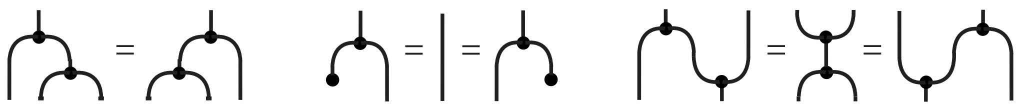
A *dagger Frobenius structure* [9, 15] on an object A consists of an (internal) multiplication

$m =$  $: A \otimes A \to A$ (14)

which is associative, has a two-sided unit

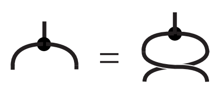
$u =$  $: \mathrm{I} \to A\,,$ (15)

and satisfies the dagger Frobenius law. Diagrammatically these are, respectively,



(16)

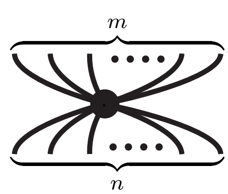
The morphism $\delta := m^\dagger : A \to A \otimes A$ is called a *comultiplication* and $\epsilon := u^\dagger : A \to \mathrm{I}$ its *counit*. A dagger Frobenius structure is *commutative* when we have

$m =$  $= m \circ \sigma_{A,A}\,.$ (17)

A dagger Frobenius structure admits an elegant diagrammatic calculus in terms of 'spiders' [24, 25, 11]. More precisely, one can show that any morphism

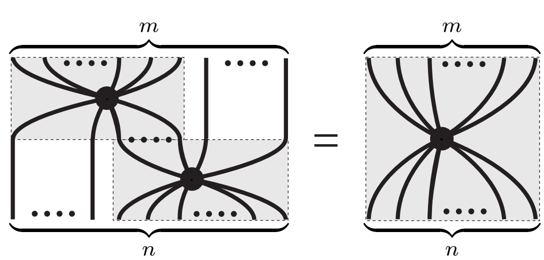
$$f : \underbrace{A \otimes \ldots \otimes A}_{n} \to \underbrace{A \otimes \ldots \otimes A}_{m} \tag{18}$$

obtained by composing and tensoring $m, u, m^\dagger, u^\dagger, 1_A$ (and also σ in the case that the multiplication is commutative), and of which the diagrammatic representation is connected, only depends on n and m. We represent this unique morphism of that type as:



(19)

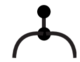
and it is then also immediately clear that these 'spiders' compose as follows:



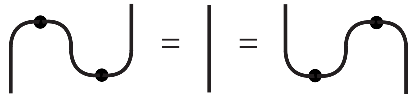
(20)

This composition rule encapsulates all of the properties of a dagger Frobenius structure in Eq. (16), and is below referred to as the *spider theorem*.

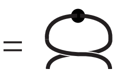
Each Frobenius structure induces a *self-dual* (i.e. $A = A^*$) compact structure

$$\eta^{Frob} =$$  $$:=$$  $$= m^\dagger \circ u : I \to A \otimes A \qquad \epsilon^{Frob} =$$  $$:=$$  $$= u^\dagger \circ m : A \otimes A \to I\,, \tag{21}$$

for which we have:

 (22)

Because of the self-duality, we can omit the arrows in these diagrams, all of which would point upward. The dot in the cups and caps (where the arrows would change direction if we were to include them) denotes the fact that the compact structure is self-dual. By self-duality we also have $A^{**} = A$, so coherence requires that the compact structure satisfies (cf. Eqs. (10)):

$$\eta^{Frob} =$$  $$=$$  $$= \sigma_{A,A} \circ \eta^{Frob} \qquad \epsilon^{Frob} =$$  $$=$$  $$= \epsilon^{Frob} \circ \sigma_{A,A}\,. \tag{23}$$

We call such a compact structure *commutative*. Obviously, in the case of a commutative Frobenius multiplication, the induced compact structure is always commutative.

For such a self-dual compact structure the convention of Eq. (8) cannot be maintained since it leads to $A \otimes B = (A \otimes B)^* = B^* \otimes A^* = B \otimes A$, which is easily seen to cause a collapse of the structure. Hence cups and caps on compound objects now have to be denoted as:

 (24)

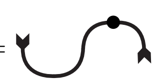
These arise from the *canonical dagger Frobenius structure on* $A \otimes B$ given one on both A and B:

 (25)

Sometimes one wishes to restore a proper dual compact structure. Following [14], the manner to do this is to introduce a *dualizer* $d : A \to A^*$ which turns the self-dual compact structure into a non-self-dual one, specifically,

$$(d \otimes I_A) \circ \eta^{Frob} = \eta_A.$$

Such a dualizer arises as

$$d :=$$  $$= (1_{A^*} \otimes \epsilon^{Frob}) \circ (\eta_A \otimes 1_A) =$$  $$: A \to A^*\,. \tag{26}$$

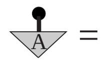
A commutative dagger Frobenius structure is moreover *normalized* if $m \circ m^\dagger = 1_A$. It was shown that for the dCC $\mathbf{FdHilb}$ of finite dimensional Hilbert spaces, linear maps, with the tensor product as the monoidal tensor, and with the linear algebraic adjoint as the dagger, normalized commutative dagger Frobenius structures are in bijective correspondence with orthonormal bases [15] . Therefore we will refer to normalized commutative dagger Frobenius structures as *classical structures*.

3 Bayesian graphical calculus

3.1 Definition

Consider a dSMC $\mathbf{C}$ in which each object comes with a dagger Frobenius structure.

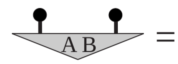
BC1 For every object $A \in |\mathbf{C}|$, we assume the existence of a *normalized state*, that is, a point which when composed with the counit yields the morphism $1_{\mathrm{I}} : \mathrm{I} \to \mathrm{I}$ (the identity morphism on the trivial object), which we depict by the 'empty picture':



$: \mathrm{I} \to \mathrm{I}\,.$ (27)

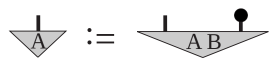
A normalized state for a composite object $A \otimes B \in |\mathbf{C}|$,

$: \mathrm{I} \to A \otimes B$ such that



$: \mathrm{I} \to \mathrm{I}\,,$ (28)

will be called a *joint state*. Note that the composition of a joint state on $A \otimes B$ with the counit on B is a state on A, which we call the *marginal state* on A,



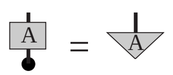
$: \mathrm{I} \to A\,.$ (29)

For most of this article, we will be concerned with just a single joint state on a set of objects together with its marginals. Consequently, it is adequate for our purposes to label states by the object on which they are defined. On the few occasions in which it will be necessary to refer to two different states on a single object, we will distinguish these by a prime.

BC2 For every object $A \in |\mathbf{C}|$, we assume the existence of a *modifier*, that is, an endomorphism

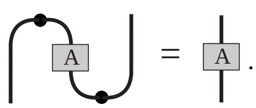
$: A \to A$ (30)

which is such that



(31)

and which is also *self-transposed*, that is,



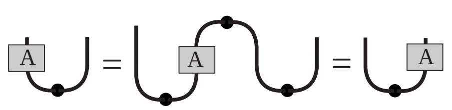
(32)

These modifiers are calculus-specific. We give concrete examples in Sections 3.4 and 3.2 of how one can construct modifiers in terms of marginal states and the Frobenius multiplication.

Proposition 3.1. *Since modifiers are self-transposed they can move along cups and caps:*

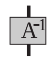
(33)

Proof: By the definition of the transpose and Eq. (22) we have:

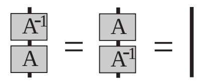


□

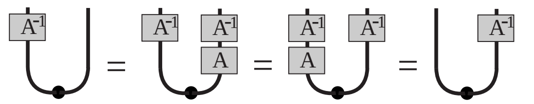
Definition 3.2. The *inverse of a modifier* : $A \to A$ is a process

 : $A \to A$ (34)

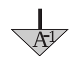
such that

 (35)

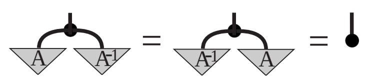
Transpose invariance of modifiers also implies that their inverses can move along cups and caps:



Definition 3.3. The *Frobenius inverse* of relative to a Frobenius multiplication is a point

 : $\mathrm{I} \to A$ (36)

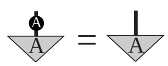
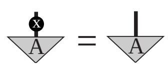
such that

 (37)

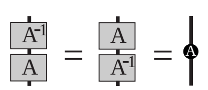
As it is the case for inverses standardly, these Frobenius inverses are easily seen to be unique.

In our key examples, there will be marginal states and associated modifiers that do not have a Frobenius inverse or an inverse modifier respectively. It turns out, however, that it suffices to have a more general notion of inverse, namely inverses relative to a support.

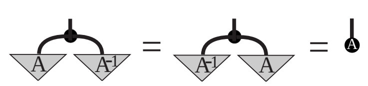
Definition 3.4. A *support* for : $\mathrm{I} \to A$ is a self-adjoint idempotent : $A \to A$ which is such that:

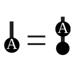
1.  , and
2.  for another self-adjoint idempotent : $A \to A$ implies .

We say that : $A \to A$ is the *inverse* to : $A \to A$ *relative to this support* if we have that

 (38)

and that : $\mathrm{I} \to A$ is the *Frobenius inverse* to : $\mathrm{I} \to A$ *relative to this support* if we have that

 (39)

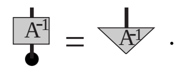
where  .

We can always take the support of a joint state on a composite object to be the tensor product of the supports of its marginals.

Below all inverses are to be understood in this generalized sense i.e. relative to a suitable support. One can also incorporate the support within the Frobenius structure by taking as the identity.

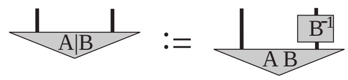
The reason we don't need to indicate the support explicitly in the current work is that we will restrict our attention to a single joint state together with the marginals and conditional states it defines and as such, we will never have need to consider states having different supports on the same object.

BC3 We assume that each state admits of a Frobenius inverse relative to its support and each modifier admits an ordinary inverse relative to its support such that the latter is the modifier associated with the former:



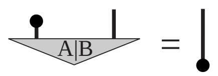
(40)

Definition 3.5. For every joint state on a pair of objects, we can define a *conditional state* to be the point

 $: \mathrm{I} \to A \otimes B\,.$

(41)

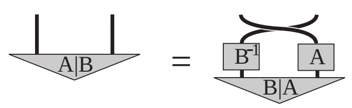
A conditional state is such that if we compose the conditioned object (the one on the left of the conditional bar) with the co-unit, we obtain the unit on the conditioning object (the one of the right of the conditional bar)



(42)

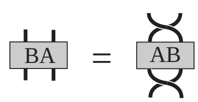
Definition 3.6. We call a graphical calculus with ingredients **BC1**, **BC2**, **BC3** a *Bayesian graphical calculus*.

This definition is motivated by the fact that with notions of joint states, marginal states, conditional states, modifiers and inverses, we have the minimal amount of structure required to describe basic concepts of Bayesian inference. For example, Bayes' rule depicts as:



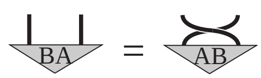
(43)

We can straightforwardly extend the above to multiple variables $A, B, C, \ldots$. When setting:



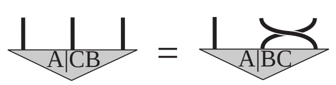
(44)

it straightforwardly follows that:



(45)

and that



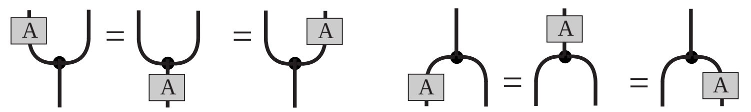
(46)

Many important concepts can now be defined at this high level of generality, most notably, conditional independence (cf. Section 5 below), and many results can be derived graphically, e.g. pooling (cf. Section 5.3 below).

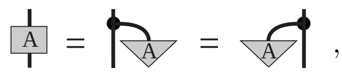
3.2 Classical Bayesian graphical calculus

Definition 3.7. A Bayesian graphical calculus is called *classical* if it satisfies the following equivalent conditions:

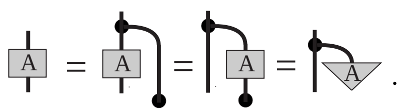
(a) modifiers can move through the Frobenius structure:


(47)

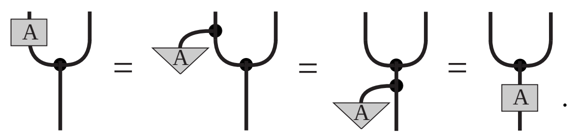
(b) modifiers are of the form:


(48)

Proof: We first show that condition (a) implies condition (b). By the spider theorem (more specifically, the counit law in Eq. (16)), Eq. (47) and Eq.(31),



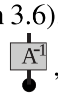
The second equality in condition (b) is proven with the mirror image of this argument. Condition (b) implies condition (a) since again by the spider theorem (more specifically, the Frobenius law in Eq. (16)),

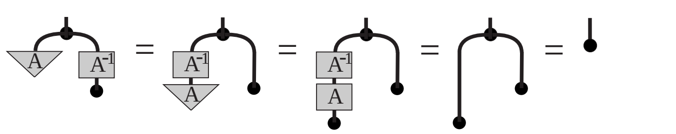


The second equality in condition (a) is again proven with the mirror image of this argument. □

So in classical Bayesian graphical calculi, in addition to moving along cups and caps (cf. Proposition 3.1), modifiers can move through the Frobenius structure, and hence, by the spider theorem, in a classical Bayesian graphical calculus modifiers can move through arbitrary spiders.

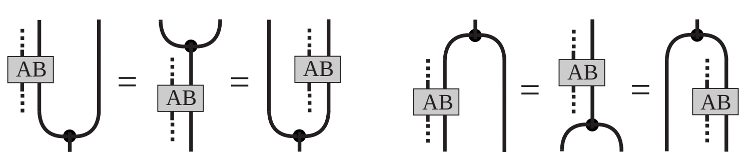
Note that the conditions in Eq. (47) and Eq. (48) hold for states and modifiers of composite objects using the Frobenius structure for the latter.

Note that the modifiers in a classical Bayesian graphical calculus are automatically self-transposed (cf. Definition 3.6). In addition, the consistency condition on inverses in Eq. (40), that is, the equivalence of  and , is automatically satisfied because

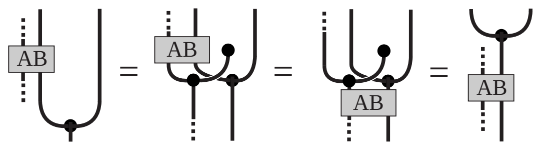


It is useful to consider some of the features of such a calculus.

Proposition 3.8. *In a classical Bayesian graphical calculus, modifiers on composite objects move through the Frobenius structure of one of the objects:*

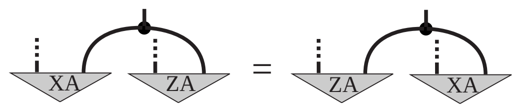

(49)

Proof: We have:



by Eq. (16) and Eq. (47) applied to $A \otimes B$. The other equalities are proven similarly. □

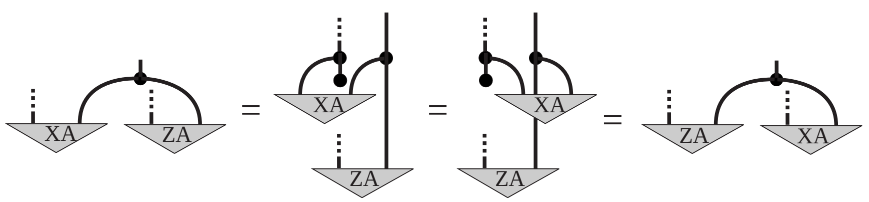
Proposition 3.9. *In a classical Bayesian graphical calculus, the Frobenius multiplication always acts commutatively on states, that is:*



(50)

Multiplication is also commutative if one or both of the states are replaced by conditional states.

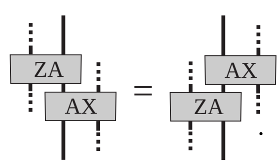
Proof: By the spider theorem and Eq. (48) we have



□

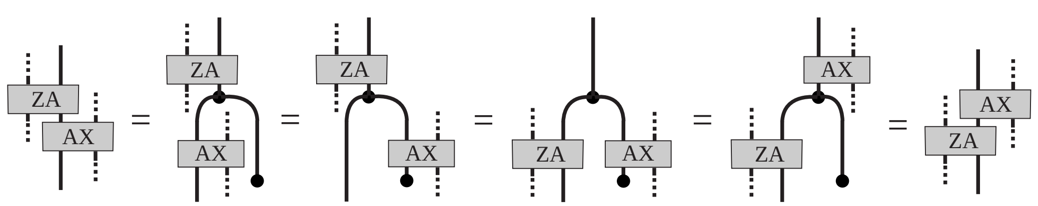
Note that it could however still be possible that the Frobenius structure itself is not commutative, but just acts commutative on the joint and marginal probabilities under consideration. E.g. the special case of Example 3.17 (which we present shortly) when all relevant density operators commute.

Proposition 3.10. *In a classical Bayesian graphical calculus, composition of modifiers on an object is commutative:*



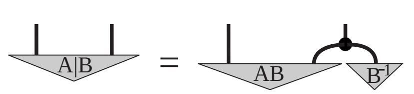
(51)

Proof: By Eqs. (16), (33), (47) we have:



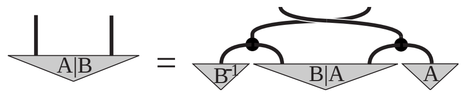
□

For a classical Bayesian calculus, conditional states have the form:



(52)

and Bayes' theorem, Eq. (43), has the form



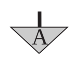
$$. \tag{53}$$

By virtue of the multiplicative commutativity, the order in which the states are 'Frobenius-multiplied' doesn't matter (unlike the quantum generalization, as we will see).

This is an abstract characterization of classical Bayesian inference. We now present a couple of concrete realizations of this calculus. We shall thereby see how the abstract characterization avoids the conventional elements of the concrete realizations.

3.3 Representations of the classical Bayesian graphical calculus

Example 3.11. Standard probability theory. Standard probability theory constitutes a special case of a classical Bayesian calculus. The objects are natural numbers and the morphisms from n to m are the $m \times n$ positive-valued matrices (consequently the points are column vectors and their daggers are row vectors). Composition is matrix product, and the tensor product is the matrix tensor product. It follows that we have



$$: \mathrm{I} \to A = (p_1, p_2, \ldots, p_n)\,, \tag{54}$$

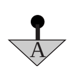
The unit is

$$: \mathrm{I} \to A = (1, 1, \ldots, 1)\,, \tag{55}$$

which implies that the co-unit must be

$$: A \to \mathrm{I} = (1, 1, \ldots, 1)^T \tag{56}$$

where T denotes matrix transposition. The counit acting on a point gives the sum of the coefficients of the associated vector

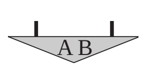


$$= (1, 1, \ldots, 1)^T (p_1, p_2, \ldots, p_n) = \sum_{j=1}^{n} p_j. \tag{57}$$

It follows that a normalized state in the Bayesian graphical calculus (cf. condition BC1) here corresponds to a positive vector with coefficients that sum to 1:

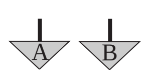
$$: \mathrm{I} \to A = (p_1, p_2, \ldots, p_n) \text{ such that } \sum_{j=1}^{n} p_j = 1. \tag{58}$$

In other words, normalized states for an object are probability distributions over the set $\{1, \ldots, n\}$. Normalized states on a composite object (nm) are simply probability distributions over the set $\{1, \ldots, nm\}$,



$$: \mathrm{I} \to A \otimes B = (p_{1,1}, p_{1,2}, \ldots, p_{n,m}) \text{ such that } \sum_{i=1}^{n} \sum_{j=1}^{m} p_{i,j} = 1. \tag{59}$$

when the joint state is a tensor product of a state $(p_1, p_2, \ldots, p_n)$ on A and a state $(q_1, q_2, \ldots, q_m)$ on B,



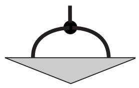
$$: \mathrm{I} \to A \otimes B = (p_i q_j | i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\})\,, \tag{60}$$

we say that it is *uncorrelated* .

The Frobenius multiplication is the $n \times n^2$ matrix $M := (M^{(1)}, M^{(2)}, \ldots, M^{(n)})$ where $M^{(k)}$ is the $n \times n$ matrix which is zero everywhere except at the kth diagonal element, where it is one,

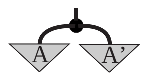
$$: A \otimes A \to A = M := \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 & & 0 & 0 & \ldots & 0 \\ \vdots & & & & \vdots & & & & \ldots & \vdots & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & & 0 & 0 & \ldots & 1 \end{pmatrix}. \quad (61)$$

Therefore, composing an arbitrary point on $A \otimes A$ with the Frobenius multiplication yields



$$= M(p_{i,i'}|i, i' \in \{1, \ldots, n\}) = (p_{i,i}|i \in \{1, \ldots, n\}). \quad (62)$$

If the point on $A \otimes A$ is a product of a state $(p_1, p_2, \ldots, p_n)$ on A and a different state $(p'_1, p'_2, \ldots, p'_n)$ also on A, then composing with the Frobenius multiplication yields



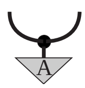
$$= M(p_i p'_{i'}|i, i' \in \{1, \ldots, n\}) = (p_i p'_i|i \in \{1, \ldots, n\}). \quad (63)$$

This is simply the component-wise product of the input vectors.

It follows from the above that the Frobenius co-multiplication is the $n^2 \times n$ matrix which is the matrix transpose of M.

$$: A \to A \otimes A = M^T. \quad (64)$$

Composing an arbitrary state on A with the comultiplication yields



$$= M(p_i|i \in \{1, \ldots, n\}) = (p_i \delta_{i,i'}|i, i' \in \{1, \ldots, n\}). \quad (65)$$

where

$$\delta_{i,i'} := \begin{cases} 1 & \text{if} \quad i = i' \\ 0 & \text{if} \quad i \neq i' \end{cases} \quad (66)$$

is the Kronecker delta. The comultiplication can therefore be understood as a classical *broadcasting map* [5] (see Section 3.5 below).

It is tedious but straightforward to verify that these definitions of unit, co-unit, multiplication and co-multiplication yield a Frobenius structure.

We now demonstrate that it is a representation of the classical Bayesian graphical calculus.

The marginal state on A of a joint state on $A \otimes B$ associated with probability distribution

$$(p_{i,j}|i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\})$$

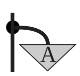
is simply the marginal distribution on A, that is,

:=

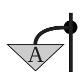
$$: \mathrm{I} \to A = (p_i = \sum_j p_{i,j}|i \in \{1, \ldots, n\}). \quad (67)$$

If one defines the modifier associated with a state $(p_1, \ldots, p_n)$ on A through Eq. (48), then it is represented by the $n \times n$ matrix

=

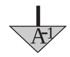


=



$$: A \to A = M[(p_1, p_2, \ldots, p_n) \otimes I] = \begin{pmatrix} p_1 & 0 & \ldots & 0 \\ 0 & p_2 & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & p_n \end{pmatrix} \quad (68)$$

The Frobenius inverse of a marginal state $(p_1, \ldots, p_n)$ on A is the vector



$$: \mathrm{I} \to A = (r_1, \ldots, r_n) \tag{69}$$

where

$$r_i := \begin{cases} p_i^{-1} & \text{if} \quad p_i \neq 0 \\ 0 & \text{if} \quad p_i = 0 \end{cases} . \tag{70}$$

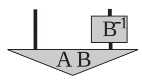
Furthermore, one easily verifies that the inverse of the modifier in Eq. (68) is simply the matrix $diag(r_1, r_2, \ldots, r_n)$, so that property BC3 is indeed satisfied.

It follows that the conditional state on $A \otimes B$ that arises from the joint state on $A \otimes B$ is simply the ordered set of conditional probability distributions that arise from the joint distribution

$$(p_{i,j} | i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}) ,$$

that is,

$:=$



$$: \mathrm{I} \to A \otimes B = (p_{i|j} | i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}) \tag{71}$$

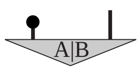
where

$$p_{i|j} := \begin{cases} p_{i,j} p_j^{-1} & \text{if} \quad p_j \neq 0 \\ 0 & \text{if} \quad p_j = 0 \end{cases} , \tag{72}$$

is a probability distribution over $i \in \{1, \ldots, n\}$, labeled by $j \in \{1, \ldots, m\}$. Note that $p_j := \sum_i p_{i,j}$ in this expression, which ensures normalization

$$\sum_{i=1}^{n} p_{i|j} = 1 \text{ for all } j \in \{1, \ldots, m\}. \tag{73}$$

Consequently the composition of the conditional state with the co-unit on A is indeed the unit on B,



$=$

$$= [(1, 1, \ldots, 1) \otimes I](p_{i|j} | i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}) \tag{74}$$

$$= (\sum_i p_{i|j} | j \in \{1, \ldots, m\}) = (1, 1, \ldots, 1) . \tag{75}$$

In a slight abuse of notation, we can use A, B and C not only to denote the objects in our category but also to denote random variables associated with these. For instance, we take A to denote the random variable taking values from the set $\{1, \ldots, n\}$ where n is the natural number associated with the categorical object A. We can also follow a standard notation and write

$$p(A) := (p_a | a \in \{1, \ldots, n\}) , \tag{76}$$

$$p(A, B) := (p_{a,b} | a \in \{1, \ldots, n\}, b \in \{1, \ldots, m\}) , \tag{77}$$

$$p(A|B) := (p_{a|b} | a \in \{1, \ldots, n\}, b \in \{1, \ldots, m\}) , \tag{78}$$

etcetera. We can then write many equations in a simple form. For instance, the Bayes' rule for classical Bayesian graphical calculi as in Eq. (53), takes the form

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} , \tag{79}$$

where this is understood to be an equality that holds component by component.

Example 3.12. Alternative representations. Here everything is defined as it was before – objects are natural numbers, morphisms are positive-valued matrices, composition is the matrix product and tensor product is the matrix tensor product – except that the underlying notions of scalar addition and multiplication are modified. The new operations, denoted by $\boxplus$ and $\boxdot$ respectively, can be defined for an arbitrary pair s, t of scalars as follows. For any function f that is bijective and hence invertible on the positive reals, they are

$$s \boxplus t = f(f^{-1}(s) + f^{-1}(t)), \qquad s \boxdot t = f(f^{-1}(s)f^{-1}(t)). \tag{80}$$

One easily verifies that these two operations obey the distributive law:

$$s \boxdot (t_1 \boxplus t_2) = (s \boxdot t_1) \boxplus (s \boxdot t_2)\,.$$

The unit for the new notion of addition, denoted $0_{\boxplus}$ and satisfying $s \boxplus 0_{\boxplus} = s$ for all s, is

$$0_{\boxplus} = f(0), \tag{81}$$

while the unit for the new notion of multiplication, denoted $1_{\boxdot}$ and satisfying $s \boxdot 1_{\boxdot} = s$ for all s, is

$$1_{\boxdot} = f(1). \tag{82}$$

The new product of two matrices M and N, denoted $M \circledcirc N$, is defined accordingly:

$$[M \circledcirc N]_{ij} = \boxplus_k([M]_{ik} \boxdot [N]_{kj}), \tag{83}$$

as is the new tensor product of two matrices, M and N, denoted $M \boxtimes N$,

$$[M \boxtimes N]_{ik,jl} = [M]_{ij} \boxdot [N]_{kl}. \tag{84}$$

The Frobenius multiplication, co-multiplication, unit and co-unit are defined as before, but with the scalars 0 and 1 replaced by $0_{\boxplus}$ and $1_{\boxdot}$. By construction, for every monotonic function f, we obtain a representation of a Bayesian graphical calculus.

It is useful to consider an example of this sort of alternative to the standard probability representation.

Example 3.13. The negative logarithm of probability representation. Consider the case where the monotonic function f is the negative natural logarithm (the generalization to an arbitrary base is straightforward),

$$f(s) = -\ln s, \qquad f^{-1}(s) = e^{-s}, \tag{85}$$

so that

$$s \boxplus t = -\ln(e^{-s} + e^{-t}), \qquad s \boxdot t = s + t. \tag{86}$$

We then have [3]

$$\begin{aligned}
[M \circledcirc N]_{ij} &= -\ln[\sum_k e^{-([M]_{ik}+[N]_{kj})}], &(88)\\
[M \boxtimes N]_{ik,jl} &= [M]_{ij} + [N]_{kl}, &(89)\\
0_{\boxplus} &= \infty, &(90)\\
1_{\boxdot} &= 0\,. &(91)
\end{aligned}$$

[3] As an aside, there is often a subtelty concerning inverse. The new multiplicative inverse of a scalar s, denoted $s^{\boxminus 1}$, must satisfy $s \boxdot s^{\boxminus 1} = 1_{\boxdot}$. It follows that

$$s^{\boxminus 1} = -s. \tag{87}$$

However, the new additive inverse of a scalar s, denoted $\boxminus s$ must satisfy $s \boxplus \boxminus s = 0_{\boxplus}$, which implies that $\boxminus s = s - \ln(-1)$, which is undefined. Consequently, there are no additive inverses in this new calculus.

Now consider a state $(s_1, s_2, \ldots, s_n)$. For it to be normalized, it must satisfy the condition

$$(1_\boxdot, 1_\boxdot, \ldots, 1_\boxdot)^T \circledcirc (s_1, s_2, \ldots, s_n) = 1_\boxdot, \tag{92}$$

which implies that

$$-\ln[\sum_k e^{-s_k}] = 0 \tag{93}$$

$$\sum_k e^{-s_k} = 1. \tag{94}$$

It follows that the components of the vector $(s_1, s_2, \ldots, s_n)$ are the negative logarithms of the components of a probability distribution $(p_1, p_2, \ldots, p_n)$,

$$\forall k : s_k = -\ln p_k. \tag{95}$$

In this new calculus, an impossible value of k (one for which $p_k = 0$) is represented by $s_k = \infty$, while a certain value (one for which $p_k = 1$) is represented by $s_k = 0$.

We can represent these vectors as $s(A), s(A, B), s(A|B)$ and so forth. We find that we have

$$s(A|B) = s(A, B) - s(B), \tag{96}$$

which is understood component-wise, that is,

$$s(A|B) := (s_{a|b} | a \in \{1, \ldots, n\}, b \in \{1, \ldots, m\}) \tag{97}$$

where

$$s_{a|b} := -\ln p_{a|b}. \tag{98}$$

The Bayes' rule takes the form

$$s(A|B) = s(B|A) + s(A) - s(B). \tag{99}$$

One has a choice in representing degrees of belief. It can be done with probabilities, but it can also be accomplished with negative logarithms of probabilities, or indeed any monotonic function of probabilities. It is a matter of convention only which is chosen. An argument to this effect was made by R. T. Cox in the context of an axiomatization of Bayesian inference [16]. We have supported Cox's conclusion by demonstrating that an abstract graphical characterization of Bayesian inference shows certain aspects of the standard probability calculus to be merely conventional.

Finally, note that by taking the usual inner product of the vector $s(A) := (s_1, s_2, \ldots, s_n)$ of negative logarithms of probabilities with the vector $p(A) := (p_1, p_2, \ldots, p_n)$ of probabilities, one obtains the *Shannon entropy* of the probability distribution $p(A)$, denoted $S(A)$,

$$S(A) := \sum_k p_k s_k = -\sum_k p_k \ln p_k. \tag{100}$$

One can similarly obtain the joint entropy as

$$S(A, B) := \sum_{i,j} p_{i,j} s_{i,j} = -\sum_{i,j} p_{i,j} \ln p_{i,j}, \tag{101}$$

and the conditional entropy as

$$S(A|B) := \sum_{i,j} p_{i,j} s_{i|j} = -\sum_{i,j} p_{i,j} \ln p_{i|j}. \tag{102}$$

Noting the the marginal entropy can also be obtained by averaging over the joint distribution,

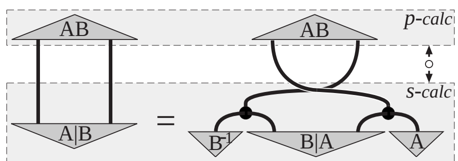
$$\sum_{i,j} p_{i,j} s_i = \sum_i p_i s_i = S(A), \tag{103}$$

it follows that any expression that holds among joints, marginals and conditionals for negative logarithms of distributions (i.e. among $s_{i,j}$, s_i, $s_{i|j}$ etcetera) also holds among the joint, marginal and conditional *entropies*. For instance, Bayes' rule in terms of negative logarithms of probabilites, Eq. (104), implies the analogous relation among entropies

$$S(A|B) = S(B|A) + S(A) - S(B). \tag{104}$$

Thus the classical Bayesian graphical calculus has the power to represent relations among classical entropies.

In more abstract terms one realizes this by considering the p- and the s-calculi as two distinct composition and tensor structures on morphisms, above denoted by $(\circ, \otimes)$ and $(\boxdot, \boxtimes)$, were $\otimes$ and $\boxtimes$ do coincide on objects. One then post-composes both sides in Eq. (53), realized in the s-calculus, with the normalized joint state of the p-calculus by means of the $\circ$-composition. That is,



. (105)

In other words, a 'p-operation'

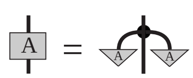
$$p(A, B)^T \circ - : \mathbf{C}(A \otimes B, \mathrm{I}) \to \mathbf{C}(\mathrm{I}, \mathrm{I})$$

is applied to both sides of an equation between s-terms in $\mathbf{C}(A \otimes B, \mathrm{I})$. Since such a p-operation can be applied to both sides of any equation between s-terms in classical Bayesian calculus, such an equation always results in a corresponding statement about classical entropies.

3.4 $Q_{1/2}$-calculus

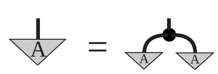
Particular cases of Bayesian graphical calculi arise by choosing a specific construction of the modifiers.

Definition 3.14. A Bayesian graphical calculus is a $Q_{1/2}$*-calculus* when modifiers are of the form:



. (106)

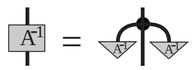
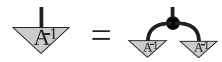
In this definition we introduced points  : $\mathrm{I} \to A$ that are distinguished from the marginal states by being denoted by smaller triangles. By Eq. (31) and the spider theorem, these must obey:



,

that is, they are the *square roots of marginal probabilities* relative to the multiplication operation of the Frobenius structure. Again by the spider theorem we also have the following lemma:

Lemma 3.15. *If Eqs. (106) and (31) hold, and if*  *has an inverse* *, then*

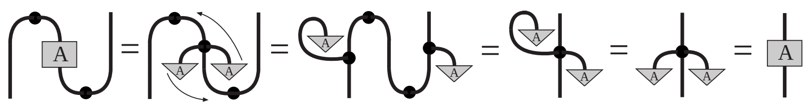
 *and*  , (107)

and hence, the consistency condition Eq. (40) is also automatically satisfied.

Also that modifiers are self-transposed now comes for free:

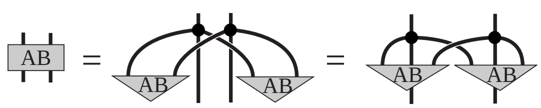
Lemma 3.16. *Modifiers of the form Eq. (106) are automatically self-transposed.*

Proof: We have:



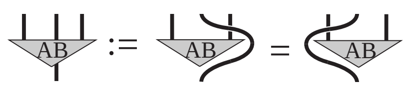
where the 2nd and 3rd step use the spider theorem, and the 4th one uses commutativity of the caps. □

In terms of the canonical dagger Frobenius structure on $B \otimes C$ we have:



(108)

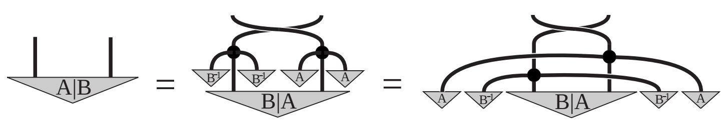
where:

 , (109)

which follows by naturality of symmetry.

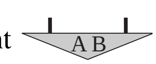
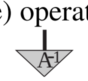
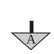
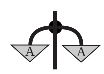
We will assume the existence of inverses of the square roots of marginals for $Q_{1/2}$-calculi, and consequently, by Lemma 3.15, inverses of the marginals themselves will also exist in $Q_{1/2}$-calculi.

For $Q_{1/2}$-calculi the Bayesian update law Eq. (43) becomes:

 . (110)

In the final expression of Eq. (110), the order of the two small triangles on the left could be reversed because they are not connected to each other by a spider. The same is true of the two small triangles on the right.

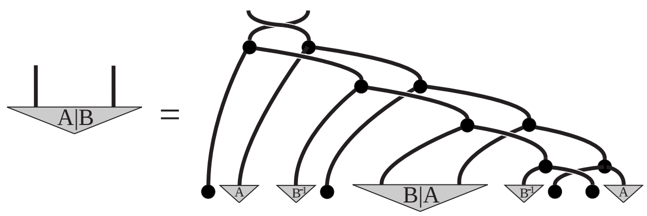
Example 3.17. The conditional density operator calculus. We explicitly construct a $Q_{1/2}$-calculus for density operators in Section 6. To give the reader a more concrete handle of the ongoing graphical axiomatization, here we already present the resulting rules which translate the graphical language to density operators.

We take the point  to be a density operator $\rho(A) : A \to A$ and the point  to be the joint density operator $\rho(A, B) : A \otimes B \to A \otimes B$. We take the Frobenius multiplication  to be the (non-commutative) operator product $- \circ -$ of density operators, and hence the identity operator 1_A is its unit . Hence,  is the inverse density operator $\rho(A)^{-1}$,  is the square-root density operator $\sqrt{\rho(A)}$, and the modifier  =  is the completely positive map $\sqrt{\rho(A)} \circ - \circ \sqrt{\rho(A)}$.

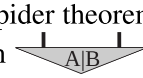
The consistency condition Eq.(40) is now also clearly satisfied. The trace is  (which is indeed the adjoint to the unit when taken in a suitable manner [34]) so marginals arise by tracing out a system on a joint density operator. The point  is Leifer's conditional density operator [27, 28], that is, a positive operator $\rho(A|B) : A \otimes B \to A \otimes B$ such that $\mathrm{Tr}_A[\rho(A|B)] = 1_B$. Note that the commutation of the compact structure, Eq. (23), corresponds to the cyclic property of the trace, i.e. $tr(\rho_A \circ \rho'_A) = tr(\rho'_A \circ \rho_A)$. Applying this translation to the diagrammatic Eq.(110), we obtain the Bayesian update rule as an identity between operators,

$$\rho(A|B) = \sigma_{B,A} \circ (1_B \otimes \sqrt{\rho(A)}) \circ (\sqrt{\rho(B)}^{-1} \otimes 1_A) \circ \rho(B|A) \circ (\sqrt{\rho(B)}^{-1} \otimes 1_A) \circ (1_B \otimes \sqrt{\rho(A)}) \circ \sigma_{A,B} \tag{111}$$

where $\rho(B|A) : B \otimes A \to B \otimes A$ is the conditional density operator associated with . To see that this is indeed the translation, it is useful to consider the diagram that incorporates each element of Eq (111) explicitly, namely,



, (112)

then note that the latter can be reduced to Eq. (110) by application of the spider theorem. The fact that we require a swap map is due to our diagrammatic convention to interpret in  the left wire as A and the right wire as B while in  it is the other way around. One can simplify Eq. (111) to

$$\rho(A|B) = \sqrt{\rho(A)}\sqrt{\rho(B)}^{-1}\rho(B|A)\sqrt{\rho(B)}^{-1}\sqrt{\rho(A)} \tag{113}$$

by leaving implicit the identity operators, the product symbols, and the swap operations. This form makes the equivalence with the diagrammatic expression Eq. (110) more evident. This *quantum Bayes rule* was introduced in this form in [29].

3.5 A Frobenius comultiplication as a logical broadcasting operation

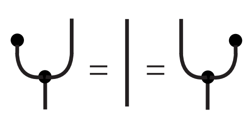
By a broadcasting operation we mean any operation

$$\delta : A \to A \otimes A \tag{114}$$

acting on a space of density operators and satisfying

$$(tr_A \otimes 1_A) \circ \delta = 1_A = (1_A \otimes tr_A) \circ \delta\,. \tag{115}$$

A Frobenius comultiplication on density operators for which the operator trace $tr_A : A \to \mathrm{I}$ is its counit satisfies Eq. (115) by counitality:


(116)

By the no-broadcasting theorem [5] it then also follows that such a Frobenius comultiplication is necessarily non-physical, i.e. it cannot be a completely positive map.

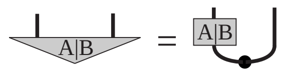
In Section 6 we give an explicit presentation of this logical broadcasting operation as an operation acting on density matrices which fails to be completely positive.

4 Inferential presentation of Bayesian graphical calculus

Above, we represented both joint and conditional states by the same triangles, only distinguishing them in terms of their labeling. We will now rely on the compact structure induced by the Frobenius structure to clearly distinguish between givens (objects on the right of the conditional bar "|" in our notation) and conclusions (objects on the left of the conditional) by representing the first as inputs (appearing at the bottom of the diagram) and the latter as outputs (appearing at the top). We do so by defining the following process, which we call a *conditional process*:

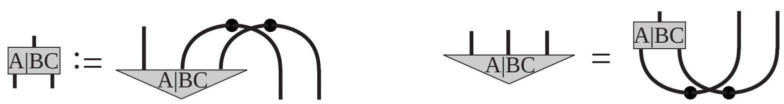
(117)

We can recover the conditional state from the conditional process by acting it upon the cup of the compact structure:



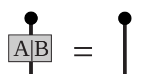
(118)

(In the context of the conditional density operator calculus, this isomorphism between conditional processes and conditional states corresponds to the version of the Choi-Jamiolkowski isomorphism described in [27].) For multiple givens we set:



(119)

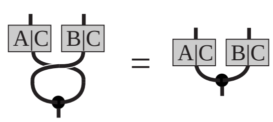
The normalization condition for conditional states, Eq. (42), is expressed in terms of conditional processes as



(120)

Using this dictionary, results that were previously expressed in terms of states may be expressed in terms of conditional processes. For instance, the commutativity of multiplication of conditional states in the classical Bayesian graphical calculus, described in Prop. (50) is equivalent to the commutativity of comultiplication of conditional processes.

Proposition 4.1. *In a classical Bayesian graphical calculus:*



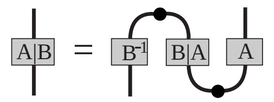
(121)

Proof: Follows from the fact that in a classical Bayesian graphical calculus, the Frobenius structure on states is commutative, Eq. (50), and from the definition of conditional processes in terms of conditional states, Eq. (117). □

We shall refer to the diagrammatic representation of an expression wherein every conditional state is replaced by its isomorphic process as the *inferential presentation* because by reading the diagram from bottom to top one follows a chain of inferences.

Note that one should not interpret the morphisms in a Bayesian graphical calculus as transformations of a physical system, but as the steps of a computation that a theorist might make in reasoning about the physical system. It is useful to emphasize this point. The classical Bayesian graphical calculus *does not* model the evolution of random variables undergoing stochastic maps but rather the mathematical operations (i.e. the belief propagation algorithm) that a statistician applies in drawing conclusions about one random variable from information about another. Similarly, a *quantum* Bayesian graphical calculus does not model the evolution of density operators under completely positive maps (in contrast to the graphical calculi that have been introduced in other works e.g. [10]), but rather the mathematical operations that a quantum theorist applies in a quantum analogue of a belief propagation algorithm.

Bayes' rule for a general Bayesian calculus, described in Eq. (43), has a particularly nice form in the inferential presentation. We simply replace the conditional states in Eq. (43) with their associated modifiers using Eqs. (117) and (118) to obtain:



(122)

This form can be simplified further. One easily verifies that the morphisms

(123)

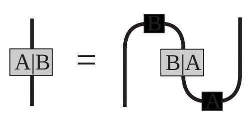
define another compact structure on A, which we will refer to as the *modified compact structure*. Note that, like the original compact structure, it is commutative and self-dual.

This modified compact structure simplifies diagrams considerably. For instance, the isomorphism between conditional processes and conditional states, Eqs. (117) and (118), can be expressed elegantly in terms of conditional processes and *joint* states using the new compact structure, as follows:

(124)

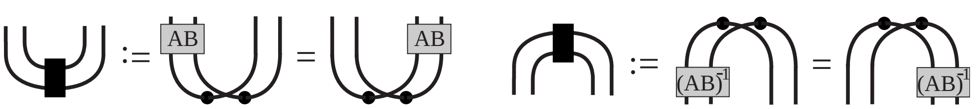
We do not decorate the black box of the modified compact structure with the label of the modifier, since this label can be inferred from the object to which the black box is connected within an inferential scenario.

The modified compact structure also provides a very simple formulation of Bayes' rule for general Bayesian calculi. It is simply the statement that A|B is the *modified transpose* of B|A :



(125)

It is straightforward to generalize these results to an arbitrary number of objects. For simplicity, we consider pairs of objects; the general case is analogous. A modifier containing a pair of object labels is simply the modifier defined by the joint state for those labels. One then introduces a modified compact structure for a pair of objects in a manner analogous to Eq. (123), namely,



(126)

Our diagrammatic convention is that the objects on the left of the modifier are in the same order as the objects on the right. In other words, we 'hide' the crossing of wires within the black box. This convention maintains the diagrams as planar as possible and in the cases where non-commutativity plays a role, it minimizes the number of swap operations one must display simultaneously.

It follows, for instance, that a conditional process of the form 

can be expressed in terms of the joint state $ABCD$ using the modified compact structure on CD:

$=$

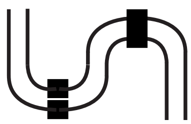
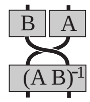

. (127)

Remark 4.2. The canonical natural isomorphism $u_{A,B}$ –cf. [23]§6– in the diagram

$$\begin{array}{ccc} \mathrm{I} & \xrightarrow{\eta_B} & B^{(*)} \otimes B \\ \downarrow{\scriptstyle \eta_{A\otimes B}} & & \downarrow{\scriptstyle 1_{B^{(*)}} \otimes \eta_A \otimes 1_B} \\ A^{(*)} \otimes B^{(*)} \otimes A \otimes B & \xrightarrow[u_{A,B} \otimes 1_{A\otimes B}]{} & B^{(*)} \otimes A^{(*)} \otimes A \otimes B \end{array}$$

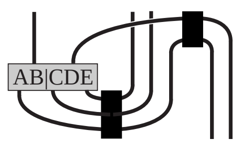
is crucially non-trivial –i.e. not just $\sigma_{A,B}$– for the modified compact structure. It is

$$u_{A,B} \;=\; (1_{B\otimes A} \otimes \epsilon_{A\otimes B}) \circ (((1_B \otimes \eta_A \otimes 1_B) \circ \eta_B) \otimes 1_{A\otimes B}) \tag{128}$$

$=$

$=$


$: A \otimes B \to B \otimes A\,.$ (129)

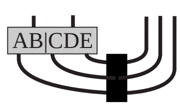
Remark 4.3 (generalized transposition)**.** The *transposition rule* in Eq. (125) can be generalized to arbitrary numbers of objects, but this requires some caution. For instance, suppose one wants to express the conditional process $ACD|BE$, which we call the *target conditional,*in terms of the conditional process $AB|CDE$, which we call the *source conditional*. It is done as follows:

$=$


(130)

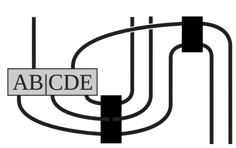
The general prescription for how to act upon the source conditional with the modified compact structure to obtain the target conditional is as follows (we illustrate with our example):

(1) one transposes all inputs into outputs:



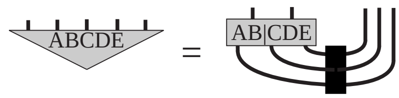
(131)

(2) one transposes those of the outputs which initially were inputs and that one wants to retain as inputs back to inputs, together with those of the initial outputs one wants to transpose into inputs:



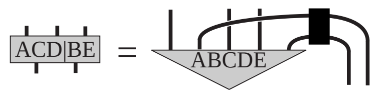
(132)

To see that this is the correct prescription, note simply that if one expresses the joint state in terms of the source conditional using the modified compact structure



(133)

and one expresses the target conditional in terms of the joint state using the modified compact structure



(134)

then by substituting Eq. (133) into Eq. (134), one obtains Eq. (130).

5 Conditional independence

5.1 Definition

One of the most important notions in the theory of Bayesian inference is that of conditional independence. In classical probability theory, a set of random variables X and another set Y are said to be conditionally independent given a third set Z if the following equivalent conditions hold:

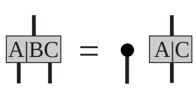
(a) $p(X|Y,Z) = p(X|Z)$

(b) $p(Y|X,Z) = p(Y|Z)$

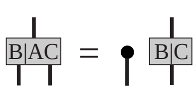
(c) $p(X,Y|Z) = p(X|Z)p(Y|Z)$

In the general Bayesian graphical calculus, there are analogues of each of these conditions, but they are no longer equivalent. We therefore distinguish two pairs of notions of conditional independence. The first pair are the analogues of Eqs. (a) and (b) respectively, while the second pair constitute analogues of Eq. (c) where one differs from the other by an interchange of the roles of A and B (which yields a different condition due to the non-commutativity of the Frobenius structure):

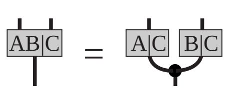
CI1$_L$



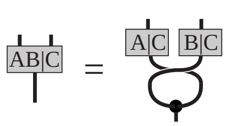
CI1$_R$



CI2$_L$



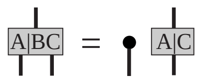
CI2$_R$



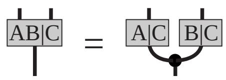
5.2 Results

Proposition 5.1. *In a Bayesian graphical calculus, if any two of the following three equalities hold then the third one also holds:*

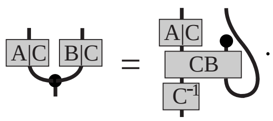
CI1$_L$



CI2$_L$



F$_L$

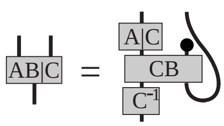


And similarly for the case where one interchanges A and B (where a condition F_R is defined in the obvious way).

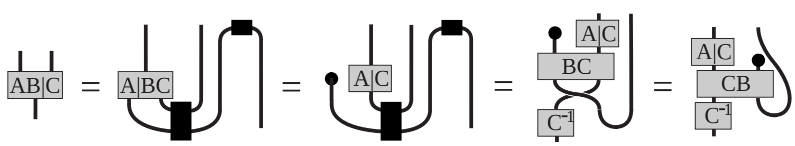
To see this, we make use of the following lemma:

Lemma 5.2. *The condition* **CI1**$_L$ *is equivalent to*

CI1$'_L$



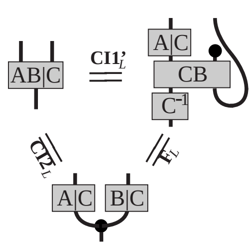
Proof: Using Bayesian inversion together with Eq. (44) and CI1$_L$, we have:



(135)

□

Proof: [proposition 5.1] Since:

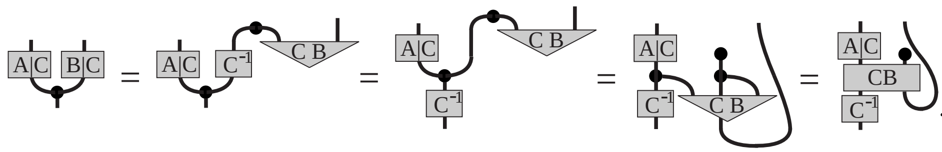


validity of any two of these equalities implies that the third also holds. The analogous equalities hold if one interchanges A and B □

It is straightforward to recover the classical notion of conditional independence, as follows.

Proposition 5.3. *In a classical Bayesian graphical calculus, the four notions of conditional independence, CI1$_L$, CI1$_R$, CI2$_L$, CI2$_R$, are all equivalent.*

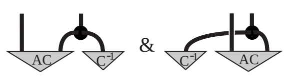
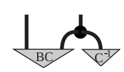
Proof: First, note that the equality F_L always holds in a classical Bayesian graphical calculus. This is proven using Eqs. (47) and (48) and the spider theorem:



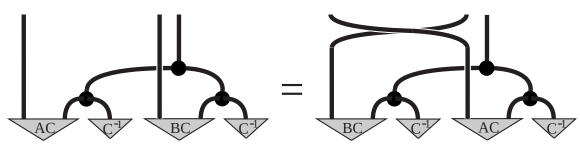
Similarly, one can prove the equality F_R, wherein A and B are interchanged relative to F_L. Given these equalities, Prop. (5.1) implies that $\mathrm{CI1}_L$ is equivalent to $\mathrm{CI2}_L$ and that $\mathrm{CI1}_R$ is equivalent to $\mathrm{CI2}_R$. Finally, the commutativity of the comultiplication of conditional processes, Prop. (4.1), implies that $\mathrm{CI2}_L$ and $\mathrm{CI2}_R$ are equivalent. Consequently, all four conditions are equivalent. □

What is more difficult is to recover a quantum notion of conditional independence. An open question is whether specifying that the form of the modifiers is as given in Eq. (106) is sufficient to prove everything that can be proven within the conditional density operator calculus. In particular, it is not clear how to derive that $\mathrm{CI2}_L$ and $\mathrm{CI2}_R$ are equivalent.

Example 5.4. In [28], by relying on results in [20], which in turns rely on a Theorem by Uhlmann [39], it was established that in the case of Example 3.17 $\mathbf{CI1}_L$ implies $\mathbf{F}_L$, and hence, by Prop. (5.1]), $\mathbf{CI1}_L$ is equivalent to the pair $\mathbf{CI2}_L$ and $\mathbf{F}_L$. The analogue holds if we interchange A and B. It would be interesting to establish whether there is a weakening to our definition of classical Bayesian graphical calculus which also establishes this. Note here also that the assumption made in [28, Thm 3.8] to derive $\mathbf{CI2}_L$ from $\mathbf{CI1}_L$ translates in graphical language to the condition that

 commute with  (136)

relative to the Frobenius multiplication, that is, for example:

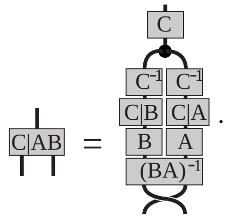
 (137)

By a tedious calculation, it can be shown that these conditions imply the weaker condition $\mathbf{F}_L$, which suffices for this purpose.

5.3 Example application: generalized pooling

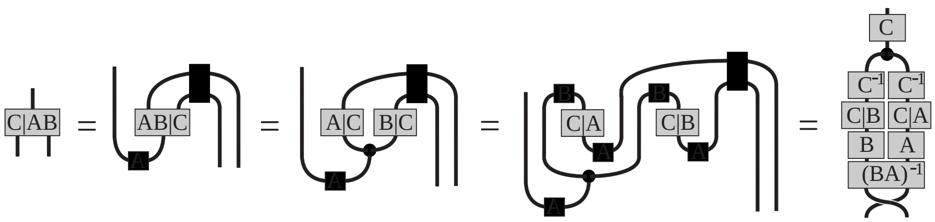
A simple example of what one can derive from the notion of conditional independence, we consider the problem of pooling. Here, one seeks to assign a conditional state to C given A, B and the question is whether this state can be expressed in terms of a conditional state for C given A and a conditional state for C given B. In the classical case (which we shall describe below), a sufficient condition is that for this to be possible is that A and B are conditionally independent given C. We here consider an analogue for general Bayesian graphical calculi.

Proposition 5.5. *If A and B are conditionally independent relative to C, in the sense of* $\mathbf{CI2}_L$, *then we* *have*



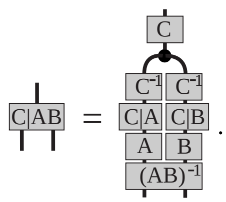
(138)

Proof:



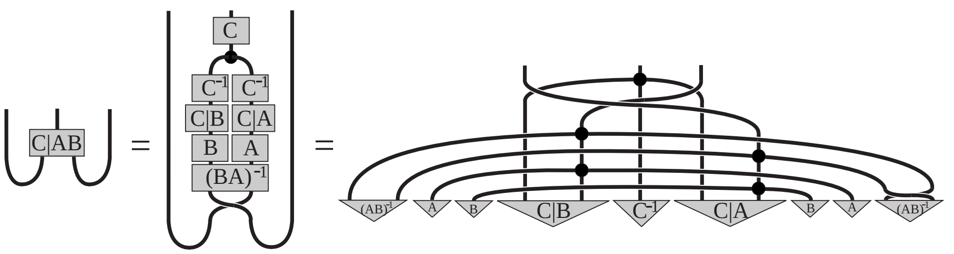
□

The case where A and B are conditionally independent relative to C in the sense of $\mathbf{CI2}_R$ differs by a swap:



(139)

Example 5.6. For $Q_{1/2}$-calculi, when expressing Eq. (138) in terms of conditional states rather than in the inferential form we obtain:



(140)

For density operators, Eq. (140) is equivalent to

$$\rho(C|AB) = \sqrt{\rho(A,B)}^{-1}\sqrt{\rho(A)}\sqrt{\rho(B)}\rho(C|B)\,\rho(C)^{-1}\rho(C|A)\sqrt{\rho(B)}\sqrt{\rho(A)}\sqrt{\rho(A,B)}^{-1}. \quad (141)$$

For classical probability distributions, we obtain

$$P(C|AB) = \frac{P(A)\,P(B)}{P(A,B)} \cdot \frac{P(C|A)\,P(C|B)}{P(C)}\,. \quad (142)$$

This result is known as the *pooling formula* because if A and B are conditionally independent given C, the posterior $P(C|AB)$ can be reconstructed from the posteriors $P(C|A)$ and $P(C|B)$ and the prior $P(C)$ (the dependence on A and B is inferred from normalization). As such, it is sufficient to "pool" the information contained in the two posteriors. Eq. (141) generalizes this to a *quantum pooling formula*, and Eq. (138) generalizes this further, to arbitrary Bayesian calculi.

5.4 The semi-graphoid axioms

One of the reasons for identifying relationships of conditional independence among objects is to have the ability to describe their mutual dependencies without providing a full specification of their joint state. Thus, it is useful to consider what implications hold among statements of conditional independencies. These conditions are well known in the classical case as the semi-graphoid axioms [31]. Let U, W, X and Y denote sets of random variables and let $X \cup Y$ denote the set-theoretic union of X and Y. In a standard notation, $I(U, W \mid X)$ is taken to express the statement that the variables in U and the variables in W are conditionally independent given X. The semi-graphoid axioms, which are easily derived from the definition (cf. 5.1) of conditional independence, are:

1. Symmetry: $I(U,W|X) \Rightarrow I(W,U|X)$
2. Decomposition: $I(U,W \cup Y|X) \Rightarrow I(U,W|X)$
3. Weak Union: $I(U,W \cup Y|X) \Rightarrow I(U,W|X \cup Y)$
4. Contraction: $I(U,W|X)$ and $I(U,Y|X \cup W) \Rightarrow I(U,W \cup Y|X)$

The semi-graphoid axioms are important because their satisfaction implies the possibility of a representation of (certain facts about) the mutual dependencies of sets of random variables in terms of a directed acyclic graph known as a Bayesian network.

It is interesting to explore the extent to which these axioms hold true for a general Bayesian graphical calculus when objects play the role of sets of random variables, tensor product plays the role of set-theoretic union, and $I(A, B \mid C)$ expresses the statement that the ordered pair of objects A, B are conditionally independent given C. Because we have four distinct notions of conditional independence in a general Bayesian graphical calculus, one can ask about the satisfaction of the axioms for any of these. As it turns out, few of the axioms hold for any of the notions of conditional independence in a general Bayesian graphical calculus. We leave for future work the question of what additional ingredients are required of a Bayesian graphical for the axioms to be satisfied. We note, however, that they are all satisfied by the classical Bayesian graphical calculus. In this sense, our formalism for classical Bayesian inference is at least as powerful as the graphoid axiomatization.

Significantly, Leifer and Poulin have shown in Ref. [28] that the conditional density operator calculus satisfies the semi-graphoid axioms, so that one may apply the tools of Bayesian networks to quantum belief propagation. Consequently, finding axiomatic graphical conditions implying the semi-graphoid axioms will presumably go hand-in-hand with finding an axiomatic graphical characterization of quantum Bayesian inference.

If the semi-graphoid axioms are satisfied within a Bayesian graphical calculus, the topology of our graphical representation of a set of correlations will reproduce the topology of the Bayesian network (with objects being mapped to nodes, and morphisms being mapped to sets of directed edges). It is our hope that by understanding how Bayesian networks can be embedded within the diagrammatic calculus of dCCs, a bridge might be built between these two fields such that insights from one might be adapted to the other.

6 Bayesian graphical calculi for arbitrary dagger compact categories

6.1 A graphical concretely non-commutative dagger Frobenius structure

We now provide a class of models, one for every dCC, each coming with a canonical non-commutative Frobenius structure that can be used to construct graphical Bayesian calculi, for example $Q_{1/2}$-calculi.

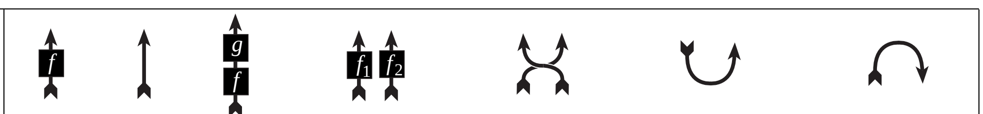
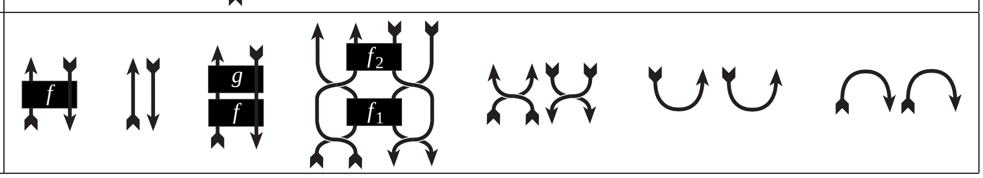
These include the conditional density operator calculus of Example 3.17 as a special case, namely the one that arises for the dCC $\mathbf{FdHilb}$. The diagrammatic presentation of mixed quantum states and completely positive maps in terms of dCCs is due to Selinger [34]. But here we cannot restrict ourselves to completely positive maps, since, as shown above in Section 3.5, the Frobenius comultiplication cannot be a completely positive map. In this context, the concrete graphical form of this non-completely positive map which we provide in this section will be insightful.

Definition 6.1. Given a dCC $\mathbf{C}$ we define another dagger category $\mathrm{D}(\mathbf{C})$ as follows:

- $|\mathrm{D}(\mathbf{C})| := |\mathbf{C}|$ i.e. the set of objects is the same for the two dCCs;
- $\mathrm{D}(\mathbf{C})(A, B) := \mathbf{C}(A \otimes A^*, B \otimes B^*)$ i.e. every morphism from $A \otimes A^*$ to $B \otimes B^*$ in $\mathbf{C}$, is a morphism from A to B in $\mathrm{D}(\mathbf{C})$;
- composition and dagger are inherited from $\mathbf{C}$ via the embedding

$$E : \mathrm{D}(\mathbf{C}) \hookrightarrow \mathbf{C} :: \begin{cases} A \mapsto A \otimes A^* \\ f \mapsto f \end{cases} . \tag{143}$$

Since $\mathrm{D}(\mathbf{C})$ is a dCC in its own right it comes with its own graphical language. It is useful to see how various elements of $\mathrm{D}(\mathbf{C})$ are represented both in the graphical language of $\mathrm{D}(\mathbf{C})$ and in the graphical language of $\mathbf{C}$. Some examples are provided in the table below. The first three columns depict morphisms on a single object: a general morphism, identity, and composition of two morphisms. Note that in the graphical language of $\mathbf{C}$ we adopt the convention that the dual objects will be represented by wires *to the right* of the primal objects.

$\mathrm{D}(\mathbf{C})$	
$\mathbf{C}$	

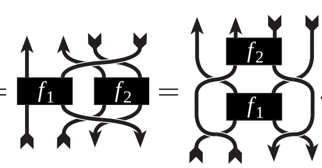
We now consider tensor products.

Definition 6.2. For

$$f_i \in \mathrm{D}(\mathbf{C})(A_i, B_i) := \mathbf{C}(A_i \otimes A_i^*, B_i \otimes B_i^*)\,, \tag{144}$$

we define a tensor $\otimes_{\mathrm{D}}$ on $\mathrm{D}(\mathbf{C})$ as

$$f_1 \otimes_{\mathrm{D}} f_2 := (1_{B_1} \otimes \sigma_{B_1^*, B_2 \otimes B_2^*}) \circ (f_1 \otimes f_2) \circ (1_{A_1} \otimes \sigma_{A_2 \otimes A_2^*, A_1^*}) = \quad = \quad . \tag{145}$$



Proposition 6.3. *Recalling that $\bar{f}$ is the* conjugate *to f i.e. the transpose of $f^\dagger$, an SMC-structure and compactness arises on $\mathrm{D}(\mathbf{C})$ from the SMC-structure and compactness of $\mathbf{C}$ via the functor*

$$F : \mathbf{C} \to \mathrm{D}(\mathbf{C}) :: \begin{cases} A \mapsto A \\ f \mapsto f \otimes \bar{f} \end{cases} \tag{146}$$

which maps the tensor $\otimes$ of $\mathbf{C}$ on the tensor $\otimes_{\mathrm{D}}$ of $\mathrm{D}(\mathbf{C})$.

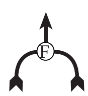
Proof: This is a trivial generalization of Theorem 4.20 in [34]. □

We recall that (cf. Eq. (8)) in the graphical language of $\mathbf{C}$ we adopt another useful convention: for the case where there is more than one object, the wires for the dual objects (in addition to appearing on the right) will appear *in the opposite order* to those of the primal objects.

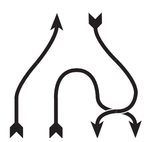
The table above presents some additional examples of elements of $\mathrm{D}(\mathbf{C})$ represented both in the graphical language of $\mathrm{D}(\mathbf{C})$ and in the graphical language of $\mathbf{C}$, in particular, the last four columns depict a tensor product of morphisms, the swap (symmetry), and the cups and caps of the compact structure.

Notation 6.4. To avoid confusion, below all 1's, $\otimes$'s, σ's, ϵ's and η's refer to the dCC $\mathbf{C}$, except for when explicitly stated otherwise. We write $f : A \to_{\mathrm{D}(\mathbf{C})} B$ for a morphism made up of these components to stipulate its type in the dCC $\mathrm{D}(\mathbf{C})$.

Proposition 6.5. *For every object $A \in |\mathrm{D}(\mathbf{C})|$ the morphism*

$$\mathcal{F} =$$  $$: A \otimes_{\mathrm{D}} A \to_{\mathrm{D}(\mathbf{C})} A$$

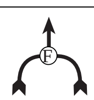
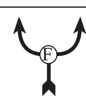
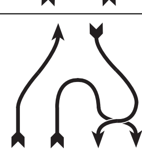
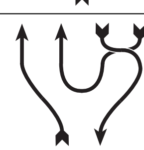
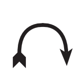
defined by

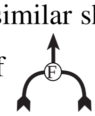
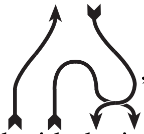
$$\mathcal{F} := (1_A \otimes \epsilon_A \otimes 1_{A^*}) \circ (1_{A\otimes A} \otimes \sigma_{A^*,A^*}) =$$  $$: A \otimes A \otimes A^* \otimes A^* \to_{\mathbf{C}} A \otimes A^* \tag{147}$$

is the multiplication of a dagger Frobenius structure with unit $\eta_{A^*} =$  *$: \mathrm{I} \to_{\mathrm{D}(\mathbf{C})} A$ defined by*

$$\eta_{A^*} =$$  $$: \mathrm{I} \to_{\mathbf{C}} A \otimes A^* \,. \tag{148}$$

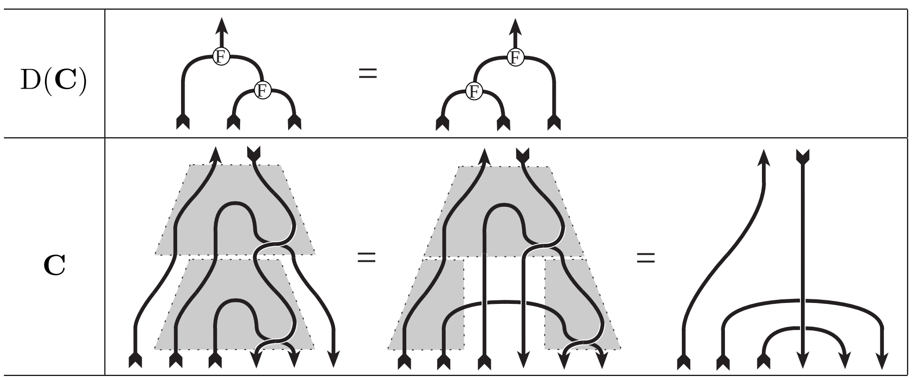
The following table depicts the multiplication, its unit, the comultiplication and its counit of the dagger Frobenius structure in the respective graphical languages of $\mathrm{D}(\mathbf{C})$ and $\mathbf{C}$.

$\mathrm{D}(\mathbf{C})$				
$\mathbf{C}$				

Because the two graphical representations of the multiplication have a similar shape, it is easy to misinterpret the mapping between these. By our convention, the left leg of  is *not* associated with the left pair of legs of , but rather with the outermost pair of legs, while the right leg of the former is associated with the innermost pair of the latter. It is useful to imagine a central left-right partition for the diagrams in **C** which divides the primal objects on the left from the oppositely-ordered dual objects on the right. The shape of the diagram in $\mathrm{D}(\mathbf{C})$ should be compared with the right-hand side of the diagram in **C**.

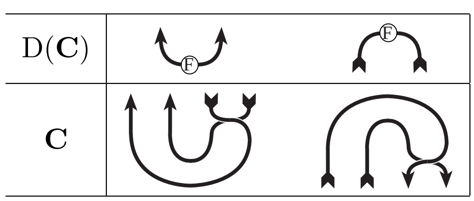
Note also that in the graphical language of $\mathrm{D}(\mathbf{C})$ we use a dot decorated by an 'F' to denote the Frobenius structure just defined. We do so to distinguish it from a Frobenius structure native to **C** (although we will not need to make use of such a structure in this article).

Proof: We must verify that $\mathcal{F}$ is associative and satisfies the dagger Frobenius law, and that η_{A^*} is indeed a two-sided unit. Representing $\mathcal{F}$ and η_{A^*} in the graphical language of $\mathrm{D}(\mathbf{C})$, these properties are given diagrammatically as Eq. (16). The tedious but straightforward proof proceeds by recasting each identity within the graphical language of **C** and verifying graph isomorphism for each. For example, associativity of the multiplication is verified as follows:

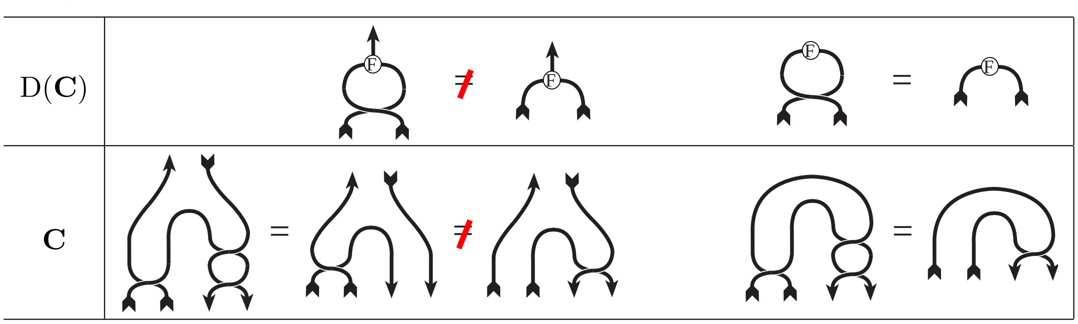


The other properties are verified similarly. □

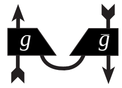
The above also illustrates how a non-commutative Frobenius multiplication can be constructed from commutative compact structures. The dagger Frobenius structure $\mathcal{F}$ induces a self-dual compact structure, which depicts as follows:



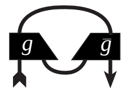
While the Frobenius multiplication is typically non-commutative (except in the degenerate case that $\sigma_{A,A} = 1_{A,A}$, which forces **C** to be trivial) the induced compact structure is always commutative:



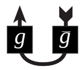
Definition 6.6 (Selinger [34])**.** A morphism $f : A \to_{\mathrm{D}(\mathbf{C})} B$ in $\mathrm{D}(\mathbf{C})$ is *completely positive* if its embedding in $\mathbf{C}$ is of the form:

$$f = (g \otimes \bar{g}) \circ (1_A \otimes \eta_{C^*} \otimes 1_{A^*}) =$$  $$: A \otimes A^* \to_{\mathbf{C}} B \otimes B^*, \tag{149}$$

for some morphism $g : A \otimes C \to_{\mathbf{C}} B$. It is *normalized* if we moreover have:

$$\epsilon_B \circ f =$$  $$=$$  $$= \epsilon_A\,. \tag{150}$$

More specifically, a point $e : \mathrm{I} \to_{\mathrm{D}} (\mathbf{C})A$ in $\mathrm{D}(\mathbf{C})$ is a *mixed state* if its embedding in $\mathbf{C}$ is of the form:

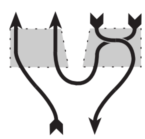
$$e = (g \otimes \bar{g}) \circ \eta_{C^*} =$$  $$: \mathrm{I} \otimes \mathrm{I}^* \to_{\mathbf{C}} B \otimes B^* \tag{151}$$

for some morphism $g : C \to_{\mathbf{C}} A$ in $\mathbf{C}$. It is *normalized* if we moreover have:

$$\epsilon_A \circ e =$$  $$= \quad . \tag{152}$$

Example 6.7. In **FdHilb** the concepts introduced in Definition 6.6 coincide with the usual ones; we explicitly establish this connection in the following section.

It is now easy to see that the failure of complete positivity in the case of the $\mathcal{F}$-comultiplication (cf. Section 3.5) is due to the lack of symmetry between the left and the right side of the picture:



(153)

This asymmetry is also what causes it to be non-commutative.

Example 6.8. Given a normalized mixed state $e_{A\ldots Z} : \mathrm{I} \to_{\mathrm{D}(\mathbf{C})} A \otimes \ldots \otimes Z$ in any such category $\mathrm{D}(\mathbf{C})$, the specified Frobenius structure allows one build a $\mathrm{Q}_{1/2}$-calculus (provided the category has the appropriate inverses and square-roots) wherein the mixed state plays the role of the joint state.

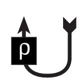
6.2 From operator presentation to $\mathrm{D}(\mathbf{C})$-presentation

At the convenience of the reader who is familiar with operator theory we now provide an explicit translation of typical operator theory concepts to the diagrammatic category $\mathrm{D}(\mathbf{C})$.

By an *operator* we mean an endomorphisms in $\rho \in \mathbf{C}(A, A)$. Such an operator ρ is *positive* if it is of the form:

$$\rho = g \circ g^\dagger =$$  .

Proposition 6.9. *For any object* $A \in |\mathbf{C}| = |\mathrm{D}(\mathbf{C})|$, *operators* $\mathbf{C}(A, A)$ *are in bijective correspondence with morphisms* $\mathrm{D}(\mathbf{C})(\mathrm{I}, A)$ *via the isomorphism*

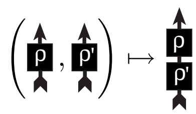
$$\xi_A : \mathbf{C}(A, A) \to \mathrm{D}(\mathbf{C})(\mathrm{I}, A) :: \rho =$$  $$\mapsto (\rho \otimes 1_{A^*}) \circ \eta_{A^*} =$$  $$. \tag{154}$$

Along this isomorphism, the positive operators in $\mathbf{C}(A, A)$ *are in bijective correspondence with the mixed states in* $\mathrm{D}(\mathbf{C})(\mathrm{I}, A)$.

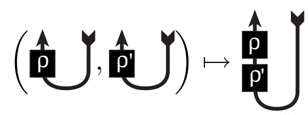
Proof: That this map is a bijection follows easily from Definition 6.1 of $\mathrm{D}(\mathbf{C})(\mathrm{I}, A)$ and the yanking equations (7), and that positive operators are in correspondence with mixed quantum states follows from the definitions of the latter, Eqs. (151) and (6.2), and from the definition of the conjugate, Eq. (13), together with the yanking equations. □

The following proposition expresses how operations on operators in $\mathbf{C}$ relate to operations on the corresponding points in $\mathrm{D}(\mathbf{C})$ along the isomorphism ξ, most notably:

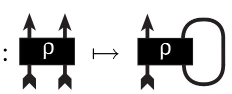
- that composition of operators in $\mathbf{C}$

$$- \circ - : \mathbf{C}(A, A) \times \mathbf{C}(A, A) \to \mathbf{C}(A, A) ::$$  (155)

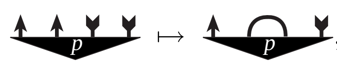
corresponds to tensor product of the corresponding points in $\mathrm{D}(\mathbf{C})$ composed with the non-commutative dagger Frobenius multiplication $\mathcal{F}$,

$$\mathcal{F} \circ (- \otimes_{\mathrm{D}(\mathbf{C})} -) : \mathrm{D}(\mathbf{C})(I, A) \times \mathrm{D}(\mathbf{C})(I, A) \to \mathrm{D}(\mathbf{C})(I, A) ::$$  (156)

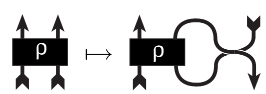
We also show that the *partial trace* of operators in $\mathbf{C}$,

$$tr_B : \mathbf{C}(A \otimes B, A \otimes B) \to \mathbf{C}(A, A) ::$$  (157)

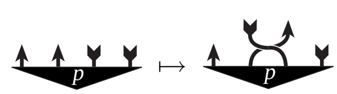
corresponds in $\mathrm{D}(\mathbf{C})$ to

$$tr_B^{\mathbf{D}} : \mathrm{D}(\mathbf{C})(\mathrm{I}, A \otimes B) \to \mathrm{D}(\mathbf{C})(\mathrm{I}, A) ::$$  (158)

and that the *partial transpose* of operators in $\mathbf{C}$,

$$T_B : \mathbf{C}(A \otimes B, A \otimes B) \to \mathbf{C}(A \otimes B^*, A \otimes B^*) ::$$  (159)

corresponds in $\mathrm{D}(\mathbf{C})$ to

$$T_B^{\mathrm{D}} : \mathrm{D}(\mathbf{C})(\mathrm{I}, A \otimes B) \to \mathrm{D}(\mathbf{C})(\mathrm{I}, A \otimes B^*) ::$$  (160)

Proposition 6.10. *(i) The following diagram commutes:*

$$\begin{array}{ccc} \mathbf{C}(A, A) \times \mathbf{C}(A, A) & \xrightarrow{\xi_A \times \xi_A} & \mathrm{D}(\mathbf{C})(\mathrm{I}, A) \times \mathrm{D}(\mathbf{C})(\mathrm{I}, A) \\ \Big\downarrow {\scriptstyle - \circ -} & & \Big\downarrow {\scriptstyle \mathcal{F} \circ (- \otimes_{\mathrm{D}} -)} \\ \mathbf{C}(A, A) & \xrightarrow[\xi_A]{} & \mathrm{D}(\mathbf{C})(\mathrm{I}, A) \end{array} \quad , \tag{161}$$

and (ii) when setting $tr_B^{\mathbf{D}}$ *as in Eq. (158), then the following diagram commutes:*

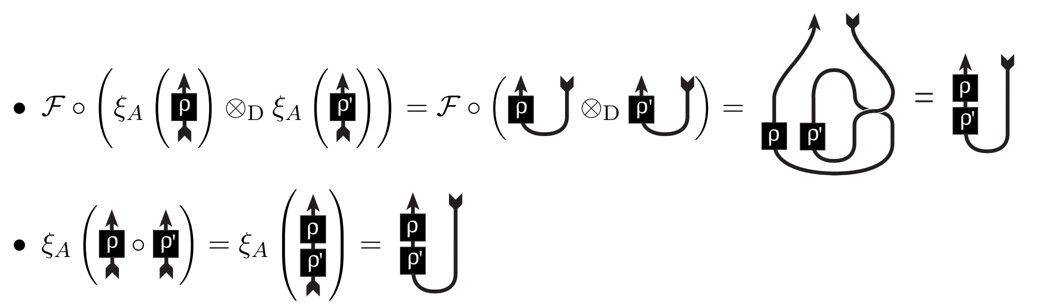
$$\begin{array}{ccc} \mathbf{C}(A\otimes B, A\otimes B) & \xrightarrow{\xi_{A\otimes B}} & \mathrm{D}(\mathbf{C})(\mathrm{I}, A\otimes B) \\ \Big\downarrow tr_B & & \Big\downarrow tr_B^{\mathbf{D}} \\ \mathbf{C}(A, A) & \xrightarrow[\xi_A]{} & \mathrm{D}(\mathbf{C})(\mathrm{I}, A) \end{array} , \tag{162}$$

and (iii) when setting T_B^{D} *as in Eq. (159) then the following diagram commutes:*

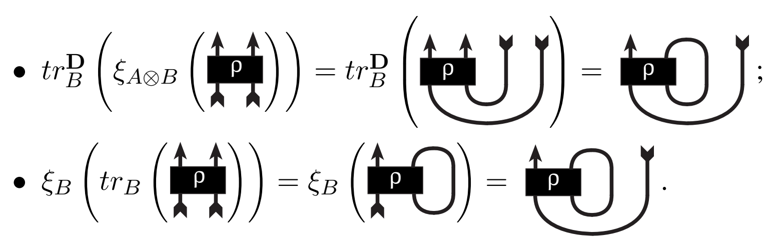
$$\begin{array}{ccc} \mathbf{C}(A\otimes B, A\otimes B) & \xrightarrow{\xi_{A\otimes B}} & \mathrm{D}(\mathbf{C})(\mathrm{I}, A\otimes B) \\ \Big\downarrow T_B & & \Big\downarrow T_B^{\mathrm{D}} \\ \mathbf{C}(A\otimes B^*, A\otimes B^*) & \xrightarrow[\xi_{A\otimes B^*}]{} & \mathrm{D}(\mathbf{C})(\mathrm{I}, A\otimes B^*) \end{array} \tag{163}$$

Proof: We have:

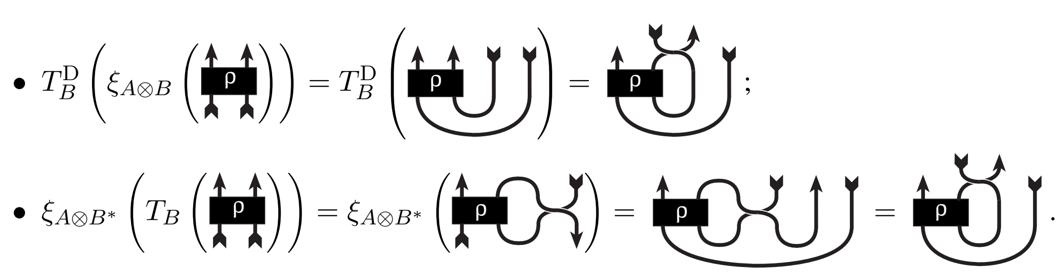
(i)



(ii)

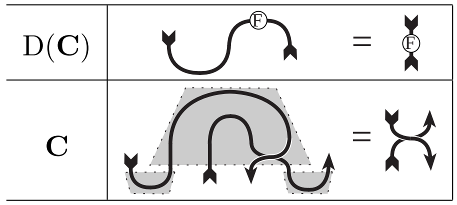


(iii)



□

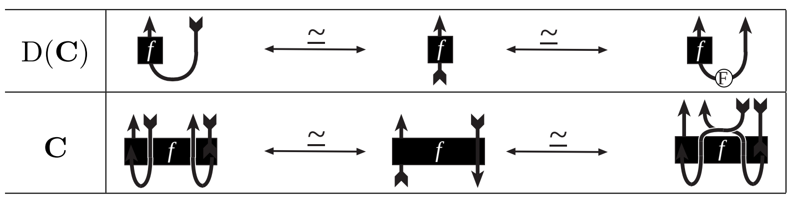
Hence the caps of the compact structure in $\mathbf{C}$ provides the partial trace, which in $\mathrm{D}(\mathbf{C})$ becomes the counit of the Frobenius multiplication, while symmetry in $\mathbf{C}$ (the swap of an object with its dual) provides the partial transpose, which in $\mathrm{D}(\mathbf{C})$ turns out to become the *dualizer* of Eq. (26), as follows:



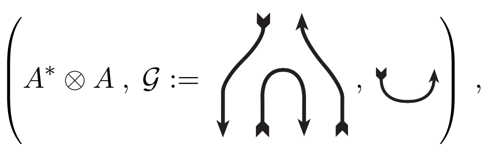
It is again clear from the form of this dualizer that it is not completely positive.

Definition 6.11. [34] If $\mathbf{C}$ is any dCC then we define $\mathrm{CPM}(\mathbf{C})$ to be the sub-dCC of $\mathrm{D}(\mathbf{C})$ which has the same objects as $\mathbf{C}$ and which has completely positive maps as morphisms.

The beauty of both $\mathrm{D}(\mathbf{C})$ and $\mathrm{CPM}(\mathbf{C})$ is that (density) operators become points rather than operations, and that completely positive maps, rather than being mappings from (density) operators to (density) operators, become morphisms. Similarly, the Choi-Jamiolkowski isomorphism takes a particularly elegant form in $\mathrm{D}(\mathbf{C})$ and $\mathrm{CPM}(\mathbf{C})$, in that it becomes a bijective correspondence between elements and morphisms. Actually, since there are two compact structures on $\mathrm{D}(\mathbf{C})$ and $\mathrm{CPM}(\mathbf{C})$, one can consider two slightly distinct Choi-Jamiolkowski isomorphisms:



Remark 6.12. A similar presentation of the *internal endomorphism monoid* in arbitrary dCCs has already appeared in the literature e.g. [26, 40], that is, a presentation as an object together with a non-commutative Frobenius structure which captures composition of endomorphisms, namely:



where $\mathcal{G}$ is now easily seen to be a dagger Frobenius structure within $\mathbf{C}$ itself, that is, in particular, with respect to the $\otimes$-tensor. While the $\otimes$-tensor and form of the Frobenius multiplication $\mathcal{G}$ are simpler to manipulate, the $\otimes_{\mathrm{D}}$-tensor is essential for $\mathrm{D}(\mathbf{C})$ (or $\mathrm{CPM}(\mathbf{C})$) to be closed under tensoring [34], and the particular form of $\mathcal{F}$ is essential for it to be an internal dagger Frobenius structure within $\mathrm{D}(\mathbf{C})$.

7 Acknowledgement

The authors thank Matt Leifer for discussions and for feedback on a draft of this article, in particular, the suggestion to explore the graphoid axioms. RWS acknowledges support from the Perimeter Institute, which is funded by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. BC acknowledges support from EPSRC ARF EP/D072786/1, ONR Grant N00014-09-1-0248, EU FP6 STREP QICS and Perimeter Institute who hosted him as a Long Term Visiting Scientist.

References

[1] S. Abramsky and B. Coecke (2004) *A categorical semantics of quantum protocols*. In: Proceedings of 19th IEEE conference on Logic in Computer Science, pages 415–425. IEEE Press. arXiv:quant-ph/0402130.

[2] J. C. Baez (2006) *Quantum quandaries: a category-theoretic perspective*. In: The Structural Foundations of Quantum Gravity, D. Rickles, S. French and J. T. Saatsi (Eds), pages 240–266. Oxford University Press. `arXiv:quant-ph/0404040`

[3] J. C. Baez and J. Dolan (1995) *Higher-dimensional algebra and topological quantum field theory*. Journal of Mathematical Physics **36**, 60736105. arXiv:q-alg/9503002

[4] H. Barnum, J. Barrett, M. Leifer and A. Wilce (2006) *Cloning and broadcasting in generic probabilistic theories*. arXiv:quant-ph/0611295

[5] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher (1996) *Non-commuting mixed states cannot be broadcast*. Physical Review Letters **76**, 2818–2821. `arXiv:quant-ph/9511010`

[6] H. Barnum and E. Knill (2002) *Reversing quantum dynamics with near-optimal quantum and classical fidelity*. Journal of Mathematical Physics **43**, 2097–2106.

[7] H. Barnum and A. Wilce (2009) *Ordered linear spaces and categories as frameworks for information-processing characterizations of quantum and classical theory*. arXiv:0908.2354

[8] J. Barrett (2007) *Information processing in general probabilistic theories*. Physical Review A **75**, 032304. arXiv:quant-ph/0508211

[9] A. Carboni and R. F. C. Walters (1987) *Cartesian bicategories I*. Journal of Pure and Applied Algebra **49**, 11–32.

[10] B. Coecke (2010) *Quantum picturalism*. Contemporary Physics **51**, 59–83. arXiv:0908.1787

[11] B. Coecke and E. O. Paquette (2006) *POVMs and Naimark's theorem without sums*. Electronic Notes in Theoretical Computer Science **210**, 131–152. `arXiv:quant-ph/0608072`

[12] B. Coecke and E. O. Paquette (2009) *Categories for the practising physicist*. In: New Structures for Physics, B. Coecke, Ed. Springer lecture Notes in Physics. To appear. `arXiv:0905.3010`

[13] B. Coecke, E. O. Paquette and D. Pavlovic (2009) *Classical and quantum structuralism*. To appear in Semantic Techniques for Quantum Computation, I. Mackie and S. Gay (eds), Cambridge University Press. `arXiv:0904.1997`

[14] B. Coecke, E. O. Paquette and S. Perdrix (2008) *Bases in diagrammatic quantum protocols*. Electronic Notes in Theoretical Computer Science **218**, 131–152. `arXiv:0808.1037`

[15] B. Coecke, D. Pavlovic, and J. Vicary (2008) *A new description of orthogonal bases*. Mathematical Structures in Computer Science. To appear. `arXiv:0810.0812`

[16] R. T. Cox (1946) *Probability, frequency, and reasonable expectation*. American Journal of Physics **14**, 1–13.

[17] L. Dixon and R. Duncan (2009) *Graphical reasoning in compact closed categories for quantum computation*. Annals of Mathematics and Artificial Intelligence **56**, 23–42.

[18] L. Dixon, R. W. Duncan and A. Kissinger, http://dream.inf.ed.ac.uk/projects/quantomatic/.

[19] M. Hasegawa, M. Hofmann and G. Plotkin (2008) *Finite dimensional vector spaces are complete for traced symmetric monoidal categories*. Lecture Notes in Computer Science **4800**, pages 367–385, Springer-Verlag.

[20] P. Hayden, R. Jozsa, D. Petz and A. Winter (2004) *Structure of states which satisfy strong subadditivity of quantum entropy with equality*. Communications in Mathematical Physics **246**, 359–374.

[21] A. Joyal and R. Street (1991) *The Geometry of tensor calculus* I. Advances in Mathematics **88**, 55–112.

[22] G. M. Kelly (1972) *Many-variable functorial calculus*. In: Coherence in Categories, G. M. Kelly, M. L. Laplaza, G. Lewis and S. Mac Lane (eds), pages 66–105. Lecture notes in Mathematics **281**, Springer-Verlach.

[23] G. M. Kelly and M. L. Laplaza (1980) *Coherence for compact closed categories*. Journal of Pure and Applied Algebra **19**, 193–213.

[24] J. Kock (2003) *Frobenius Algebras and 2D Topological Quantum Field Theories*. Cambridge University Press.

[25] S. Lack (2004) *Composing PROPs*. Theory and Applications of Categories **13**, 147–163.

[26] A. D. Lauda (2006) *Frobenius algebras and ambidextrous adjunctions*. Theory and Applications of Categories **16**, 84–122. `arXiv:math.CT/0502550`

[27] M. S. Leifer (2006) *Quantum dynamics as an analog of conditional probability*. Physical Review A **74**, 042310. `arXiv:quant-ph/0606022`

[28] M. S. Leifer and D. Poulin (2008) *Quantum graphical models and belief propagation*. Annals of Physics **323**, 1899–1946. `arXiv:0708.1337`

[29] M. S. Leifer and R. W. Spekkens (2008) *Quantum analogues of Bayes' theorem, sufficient statistics and the pooling problem*, in preparation (2009).

[30] P.-A. Melliés (2009) *Categorical semantics of linear logic*. `www.pps.jussieu.fr/~mellies/papers/panorama.pdf`

[31] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Francisco, 1988.

[32] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2000.

[33] R. Penrose (1971) *Applications of negative dimensional tensors*. In: Combinatorial Mathematics and its Applications, pages 221–244. Academic Press.

[34] P. Selinger (2007) *Dagger compact categories and completely positive maps*. Electronic Notes in Theoretical Computer Science **170**, 139–163.

[35] P. Selinger (2010) *Finite dimensional Hilbert spaces are complete for dagger compact closed categories*. Electronic Notes in Theoretical Computer Science, to appear.

[36] P. Selinger (2010) *A survey of graphical languages for monoidal categories*. In: New Structures for Physics, B. Coecke (ed), pages 275–337. Lecture Notes in Physics, Springer-Verlag. arXiv:0908.3347

[37] R. Street (2007) *Quantum Groups: A Path to Current Algebra*. Cambridge University Press.

[38] V. Turaev (1994) *Quantum Invariants of Knots and 3-Manifolds*, de Gruyter.

[39] A. Uhlmann (1977) *Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory*. Communications in Mathematical Physics **54**, 21–32.

[40] J. Vicary (2008) *Categorical formulation of quantum algebras*. Communications in Mathematical Physics. To appear. `arXiv:0805.0432`

[41] D. N. Yetter (2001) *Functorial Knot Theory. Categories of Tangles, Coherence, Categorical Deformations, and Topological Invariants*. World Scientific.